

Krzysztof Przybyszewski*, Andrzej Cader*, Alina Marchlewska*

O możliwościach zastosowania wybranych metod wyznaczania efektywności klasycznych i e-learningowych systemów kształcenia

1. Wprowadzenie

Każdy system nauczania można rozpatrywać jako system informacyjny. Dotyczy to zarówno klasycznych systemów kształcenia jak i systemów kształcenia wspomaganych technologiami informatycznymi (systemy kształcenia na odległość, systemy e-kształcenia zwane też systemami kształcenia zdalnego). Nie są to jednak systemy proste – ich złożoność można rozpatrywać w różnych aspektach [30, 33].

Zadaniem systemu kształcenia jest wykształcenie u każdego uczestnika (ucznia, studenta, słuchacza) odpowiednich kompetencji (umiejętności opartych na konkretnej, ustrukturyzowanej i ukierunkowanej wiedzy). Efektywność działania systemu można określić tylko poprzez ocenę wzajemnych relacji między osiągniętymi przez uczestników poziomami zdobytych kompetencji i ponoszonymi przez organizatora kosztami (czasu nauki, zaangażowanych środków materialnych i technicznych itp.).

W celu oceny efektywności działania dowolnego systemu kształcenia (także jego poszczególnych elementów) jako systemu informatycznego należy wyznaczyć zbiór parametrów lub funkcję, na podstawie wartości których można dokonać tej oceny [33]. Jest to jednoznaczne z definicją funkcji informacyjnej dla systemu kształcenia jako systemu informacyjnego.

Zgodnie z definicją systemu informacyjnego [25, 27, 36] funkcja informacyjna jest jednym ze sposobów przedstawiania informacji o elementach systemu charakteryzowanych przez ten sam zbiór cech. W zależności od celu można dobrać odpowiedni zbiór cech. Wartości funkcji informacyjnej dla tego zbioru mogą być przesłankami wniosku lub systematyki elementów systemu. Przykład wniosku o efektywności ekonomicznej systemu kształcenia opartego na odpowiednio skonstruowanej funkcji informacyjnej został opisany w pracy Krusia [18]. Najczęściej jednak wyznaczana jest tzw. efektywność peda-

* Instytut Technologii Informatycznych, Społeczna Wyższa Szkoła Przedsiębiorczości i Zarządzania w Łodzi

gogeniczna systemu. Powinna ona określać stopień dopasowania wyników kształcenia (określanych za pomocą wartości odpowiednich parametrów, np.: ocen, wskaźnika zatrudnienia absolwentów, wyników egzaminu zawodowego czy też wskaźnika dopasowania do przyjętych standardów stanowiska pracy [32]) do wymagań (np.: pracodawców oferujących stanowiska pracy lub szkół wyższego poziomu) lub określonych wcześniej standardów (np.: systemu ECTS opracowanego w trakcie Procesu Bolońskiego¹).

Proces oceny efektywności systemu kształcenia może dotyczyć wielu aspektów systemu. Jeżeli będziemy rozpatrywać systemy kształcenia jako systemy informacyjne i stosować techniki analizy systemowej (a taki sposób wydaje nam się najefektywniejszy), to konieczna jest separacja:

- elementów systemu,
- infrastruktury (lub środowiska przenoszenia informacji),
- procesów,
- otoczenia.

W przypadku każdego z tych obiektów ocena efektywności będzie miała inny wymiar, a metody jej przeprowadzania mogą (choć nie muszą) być różne. Taka sytuacja występuje w przypadku sprawdzianów kompetencyjnych w polskich szkołach podstawowych oraz egzaminów gimnazjalnych i maturalnych. W tych samych raportach umieszczane są: wyniki oceny uczniów, szkół, okręgów (czyli ogólnie mówiąc: elementów systemu lub ich aglomeratów), ale także oceny poprawności i efektywności stosowanych testów (czyli oceny procesu ewaluacji)², co sprawia wrażenie potraktowania podmiotu kształcenia (indywidualnego ucznia), jako przedmiotu poddanego procesowi (produktu). Wydaje nam się to dosyć nieprzyjemnym zabiegiem.

Separacja poszczególnych obiektów składowych jest możliwa w przypadku przyjęcia cybernetycznego modelu nauczania [38] w przypadku zarówno klasycznych systemów nauczania, jak i systemów e-nauczania. Tak naprawdę, taki podział wydaje się nam sztuczny i nieprawidłowy, zważywszy, że co raz częściej można spotkać formę mieszaną kształcenia (tzw. *Blended Learning*), w której część zajęć jest prowadzona sposobem klasycznym, a przy realizacji pozostałej części wykorzystywane są metody e-kształcenia. Fakt ten znalazł nawet odbicie w rozporządzeniach ministerialnych i standardach nauczania, w których ściśle określono dopuszczalną ilość zajęć prowadzonych metodami nauczania zdalnego. Taka forma zajęć nie jest zresztą niczym nowym i wcześniej była wykorzystywana przez nauczycieli jako sposób przekazu dodatkowych materiałów słuchaczom (dawniej w formie tekstowej, następnie w formie elektronicznej i multimedialnej, a wraz z rozwojem sieci tele-

¹ www.bologna-bergen2005.no
www.dfes.gov.uk/bologna

² www.cke.edu.pl/index.php?option=content&task=view&id=136&Itemid=107
www.cke.edu.pl/index.php?option=content&task=view&id=141&Itemid=122
www.cke.edu.pl/index.php?option=content&task=view&id=247&Itemid=147

informatycznych – w formie materiałów umieszczanych na serwerach sieciowych) oraz sposób komunikacji ze słuchaczami poza zajęciami (w formie kół naukowych, korespondencji listownej, a współcześnie w formie listów elektronicznych, blogów oraz z wykorzystaniem różnych funkcji ogólnodostępnych komunikatorów sieciowych).

W centrum naszego zainteresowania znalazła się ocena efektywności pedagogicznej systemów kształcenia (a właściwie jego podstawowych elementów składowych: studenta/ucznia lub ich aglomeratów – grup, szkół), dlatego też nasze rozważania dotyczą przede wszystkim metod dotyczących tych obiektów systemu.

W przypadku nauczania tradycyjnego proces ewaluacji (rozumiany tutaj jako proces przypisywania wskaźnika (np. oceny) odpowiedniego poziomu nabytych przez uczestnika kompetencji) oparty jest na działaniu systemu ekspertowego: nauczyciel posługując się własną wiedzą i ustalonymi regułami, oceniając postępy każdego z uczestników kształcenia oraz proponując dalsze partie materiału czy też sposoby wzbogacania jego umiejętności [8, 21]. Taki też sposób został bezpośrednio zaimplementowany w systemach e-kształcenia. W pierwszym okresie (niestety, trwającym w Polsce po dziś) ewaluację przeprowadzał nauczyciel (w czasie bezpośredniego spotkania lub wykorzystując udostępnione mu narzędzia komunikacyjne). W miarę rozwoju technologii (nie tylko informatycznych) nauczyciel przekazywał swoją wiedzę ekspercką do odpowiedniej bazy reguł, co przy jednoczesnej odpowiedniej konstrukcji testów pozwoliło na coraz większą automatyzację procesu ewaluacji.

Moduły obsługujące bazę reguł (zawierające także zbiór algorytmów wyznaczania ocen końcowych na podstawie ocen cząstkowych) zostały nazwane modułami ewaluacyjnymi. Mogą one jako samodzielne programy wspierać pracę nauczyciela lub też być odrębnymi jednostkami w systemach e-learningowych, w których zazwyczaj są integralną częścią systemów LMS (ang. *Learning Management System*) nadzorujących proces kształcenia w odniesieniu do każdego uczestnika [12, s.: 68]. Moduły ewaluacyjne mogą być dołączane do dowolnego elementu systemu kształcenia, tzn. mogą być połączone z pojedynczym tematem (np. moduły ewaluacji testu) [16], kursem lub przedmiotem (moduły wspomagające ocenę przedmiotową, zaliczeniową lub egzaminacyjną) [26, 30, 31], ale mogą także wspierać (lub samodzielnie dokonywać) ocenę periodyczną (np. semestralną, roczną) lub końcową całego cyklu kształcenia [29].

2. Separacja procesów, elementów i uwarunkowań systemów kształcenia

Rozpatrując cybernetyczny model kształcenia [38], można zauważyć konieczność separacji: ewaluacji samego procesu, ewaluacji (a właściwie oceny stopnia dopasowania do założonych rezultatów) poszczególnych elementów systemu (słuchaczy, nauczycieli) oraz warunków pracy systemu i jego elementów (uwarunkowań).

Od zarania dziejów ludzkości człowiek zdobywał wiedzę na temat otaczającej go rzeczywistości i nabywał umiejętności wykorzystywania narzędzi do zmiany jej elementów samodzielnie. Umiejętność uczenia się zawsze była jednym z elementów inteligencji rozu-

mianej jako zdolność przystosowywania się do życia w zastanych warunkach. Proces uczenia się zawsze był wspomagany przekazem wiedzy lub pokazem umiejętności ze strony osobników bardziej doświadczonych (aspekt decydujący o społecznej naturze człowieka). W miarę powstawania nowych dyscyplin wiedzy i umiejętności konieczna stała się specjalizacja osób przekazujących wiedzę i demonstrujących umiejętności oraz kierowanie wyselekcjonowanych osobników na naukę właśnie do nich. Wybrane jednostki, najmądrzejsze i najbieglejsze w danej dziedzinie działalności człowieka, zostały mistrzami, mentorami, czyli, używając ogólnej nazwy, nauczycielami.

Opieka nauczyciela była sprawowana najpierw indywidualnie (nad jednym uczniem), a później, w miarę wzrostu populacji na danym obszarze, grupowo (nad kilkoma uczniami). Schemat ten dotyczył zarówno dyscyplin praktycznych (np.: wojsko, rzemiosło), jak i działalności intelektualnej, społecznej i politycznej, a nawet artystycznej. Pomimo powiązania nauczyciela z grupą uczniów, to kontakty i przekazy były ściśle spersonalizowane, przy czym nauczyciel mógł dodatkowo wykorzystywać w nauczaniu efekty synergii związane z pracą każdego ucznia w grupie.

Wraz z rozwojem cywilizacyjnym (pismo, druk, środki masowego przekazu, systemy komputerowe, sieci teleinformatyczne, sieci bezprzewodowe, systemy mobilne, systemy rozproszone) zmieniały się kanały komunikacyjne nauczycieli z uczniami oraz układy zależności (np. nauka ucznia mogła być realizowana pod kierunkiem wielu nauczycieli reprezentujących różne ośrodki), ale zasadniczy schemat, podmiot oraz egzekutor procesu uczenia pozostawał niezmienny: uczeń musiał zdobywać wiedzę i rozwijać swoje umiejętności sam, korzystając ze wsparcia nauczyciela oraz wykorzystując efekty synergii wynikające z jego uczestnictwa w grupie.

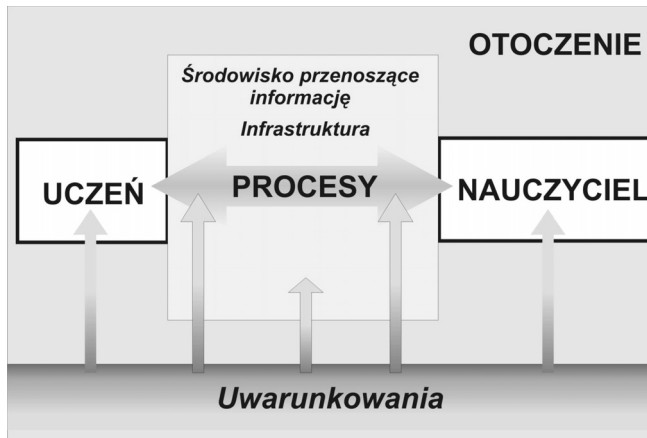
Elementy jednostkowego systemu (układu uczeń-nauczyciel) oraz przebiegające w nim procesy ulegały **granulacji** (np.: uczniowie byli grupowani w zespoły, pokrewne zagadnienia były grupowane wewnątrz dyscyplin nauczania w przedmioty), ale także podlegały **rozproszaniu** (np.: możliwość realizacji studiów interdyscyplinarnych odbywanych w wielu ośrodkach lub rozproszenie wykorzystywanych baz wiedzy i miejsc kształtowania umiejętności czyli laboratorium) oraz **wirtualizacji** (rozumianej jako zastępowanie stanów i bodźców rzeczywistych symulacjami oddziaływującymi bezpośrednio na receptory człowieka).

Warto też zwrócić uwagę na ewolucję procesów realizowanych w systemie jednostkowym (układzie uczeń-nauczyciel). Można wyróżnić trzy zasadnicze procesy:

- 1) przekaz i pokaz (wiedzy, umiejętności),
- 2) stymulacja (rozumiana jako proces ciągłego pobudzania odpowiednimi bodźcami),
- 3) ewaluacja lub ocena (elementów systemu, metod, procesów oraz uwarunkowań).

Ich realizacja (stosowane algorytmy i standardy) w znaczący sposób zależała od stopnia rozwoju cywilizacyjnego.

Schemat jednostkowego układu uczeń – nauczyciel można przedstawić w sposób pokazany na rysunku 1.



Rys. 1. Schemat układu uczeń-nauczyciel
(opracowanie własne)

Zmiany układu polegają jedynie na wymianie zawartości prostokątów symbolizujących: medium komunikacyjne i infrastrukturę oraz otoczenie (w tym także uwarunkowania). Pozostałe elementy i procesy pozostają stałe.

Przy takim układzie ocenie i ewaluacji mogą podlegać poszczególne elementy struktury: uczeń, nauczyciel, medium komunikacyjne oraz infrastruktura, ale także procesy (a właściwie metody ich realizacji). Ocenie i ewaluacji może także podlegać organizacja całego systemu jednostkowego. Możliwa jest także ocena wpływu czynników zewnętrznych (uwarunkowań) na: system, jego elementy oraz procesy w nim zachodzące.

Jak już to stwierdziliśmy w pierwszej części opracowania: w celu oceny efektywności działania dowolnego systemu kształcenia (także jego poszczególnych elementów) jako systemu informatycznego lub też ewaluacji procesów i struktury należy zdefiniować funkcję informacyjną określoną na odpowiedniej dziedzinie (zbiorze cech elementów systemu, zbiorze cech procesów, czy też zbiorze uwarunkowań). Separacja elementów systemu (ucznia, nauczyciela, medium komunikacyjnego i infrastruktury), procesów (komunikacji, przekazów, stymulacji i oceny) oraz uwarunkowań, która została zaproponowana w tej części opracowania, umożliwi wyznaczenie niezależnych funkcji informacyjnych dla odseparowanych elementów systemu (częstkowych funkcji informacyjnych). Procedura umożliwia ewaluację warunkową ograniczoną do odpowiednich obiektów systemu. Ocena efektywności działania całości systemu (ewaluacja zupełna) jest dokonywana na splocie funkcji informacyjnych.

Najczęściej przeprowadzana jest ewaluacja warunkowa systemów kształcenia, których obiekty są granulatami elementów (np.: ocena efektywności działania szkoły, która jest zbiorem grup uczniów oraz grupy nauczycieli lub działania całych okręgów szkolnych będących zbiorem szkół). W takim przypadku można bardzo łatwo określić procedury ewaluacji warunkowej odnoszące się do uczniów lub grup uczniowskich bądź też nauczy-

cieli, po przyjęciu odpowiednio zestandaryzowanej funkcji informacyjnej [30, 34]. Jest to przykład ewaluacji ukierunkowanej przede wszystkim na ucznia (lub nauczyciela) i określającej efektywność pedagogiczną działania systemu kształcenia. Opis procedur porównawczej oceny efektywności pedagogicznej szkół w okręgu szkolnym zawiera praca Pokropka [28].

Przykładem bardzo szczegółowego opisu procedur ewaluacji warunkowej odnoszącej się do procesów zachodzących w otwartych systemach nauczania zdalnego jest publikacja Różewskiego i jego współpracowników [35].

Przykładami ewaluacji wpływu uwarunkowań są prace: Krupskiego i Cadera [17] oraz Strumiłło i innych [37]. Ostatnia praca jest jednocześnie próbą zastosowania wybranych metod do cząstkowej oceny efektywności pedagogicznej wybranego kierunku uczelni wyższej. W takim przypadku nie wystarcza ocena osiągnięć studentów. Jedynym w miarę obiektywnym parametrem, który można zastosować, jest współczynnik dopasowania poziomu umiejętności studentów do wymagań pracodawców. W przytoczonej pracy przedstawiono właśnie jeden ze sposobów szacowania wymagań pracodawców i dostosowania do nich programu wybranych przedmiotów oraz sposobu oceniania wyników studentów.

Bardzo ciekawy sposób ewaluacji postępów studentów i szacowania ich dopasowania do wymogów potencjalnych pracodawców dla systemów e-nauczania oparty na metodzie agentowej przedstawiono w pracach: de Azevedo i Scalabrina [1] oraz Cadera i współpracowników [2].

3. Ocena efektywności pedagogicznej działania systemów kształcących

W klasycznych sposobach oceny efektywności pedagogicznej systemów kształcenia lub jego elementów, jako zbiór dostępnych wartości atrybutów (przeciwdziedzina, obraz, zbiór wartości) najczęściej jest przyjmowana lingwistyczna lub liczbowa skala ocen, a wartości funkcji są przypisywane argumentom zgodnie z przyjętymi przez nauczyciela (bądź gremium pedagogiczne lub zarząd pedagogiczny) algorytmami lub wyznaczane jako oceny końcowe (najczęściej są to średnie z ocen cząstkowych, obliczane z uwzględnieniem odpowiednich wag). Najlepszymi, według nas, opracowaniami przedstawiającymi większość stosowanych algorytmów są prace Niemierki [19–23].

W wielu przypadkach ministerialna skala ocen jest niewystarczająca: brakuje w niej ocen ułamkowych [33]. Taka sytuacja występuje w przypadku obliczania średniej oceny dla grup uczniowskich (klas) lub w przypadku klasyfikacji szkół. Dlatego też w tych przypadkach dokonuje się swoistego „uciąglenia” (zagęszczenia) skali polegającego na przyjęciu reprezentacji skali ocen w przedziale domkniętym $[0; 6]$ liczb wymiernych. Taka reprezentacja liczbowa skali ocen (oznaczona przez nas symbolem $SM_{ciąg}$) była do niedawna najczęściej stosowaną reprezentacją skali ocen.

Uzyskane wartości funkcji informacyjnej służą następnie do systematyki elementów kształcenia (uczniów/studentów, uczestników szkoleń, grup uczniowskich, studenckich lub

kursowych, nauczycieli, szkół i uczelni itp.). Jeżeli wyniki zbierane są w pewnym okresie czasu w odpowiedniej bazie (a właściwie hurtowni) danych [3, 4], to analiza ich zależności od czasu oraz innych czynników, może także zostać wykorzystana do prognozowania wartości funkcji informacyjnej w przyszłości, czyli prognozowaniu efektywności działania badanego elementu lub całego systemu.

Klasyfikacja systemów kształcenia (lub ich elementów) odbywa się poprzez porównywanie średnich wartości zbiorów liczbowych, z uwzględnieniem odpowiednich momentów rozkładów tych zbiorów (testowanie hipotezy o równości zbiorów) [7, 9, 10, 24, 39].

Wymienione wyżej metody klasyfikowania wykorzystują wskaźniki bezwzględne i są możliwe do zastosowania w przypadku przyjęcia jednakowej skali ocen (jednolitego zbioru dostępnych wartości funkcji informacyjnej). W przypadku porównywania wyników kształcenia systemów lub elementów systemu ocenianych z zastosowaniem różnych skal ocen, należy zastosować względne wskaźniki efektywności procesu. Taka sytuacja występuje w przypadku porównywania wyników prowadzonego od siedmiu lat sprawdzianu kompetencyjnego po szkole podstawowej czy też egzaminów gimnazjalnych z ocenami wystawianymi w szkole. Wyniki sprawdzianów i egzaminów podawane są jako ilość zdobytych punktów, natomiast oceny (roczne, końcowe) wystawiane są zgodnie z zarządzeniem ministerialnym, z wykorzystaniem skali ogólnej (reprezentacji lingwistycznej SM_{ling} lub liczbowej SM_{licz}). W przypadku porównywania złożonych systemów kształcących (np.: szkoły, okręgi itp.) lub ich złożonych elementów (np. oddziały szkół) zachodzi konieczność zastosowania mniej precyzyjnych wartości funkcji informacyjnej (np. oceny średniej oddziału czy szkoły), której wartości będą należały do odpowiedniego wymiernego przedziału liczbowego (reprezentacja $SM_{ciąg}$ w przypadku ocen lub odpowiednie przedziały liczbowe dla sprawdzianu kompetencyjnego: przedział punktowy [0, 40]; bądź egzaminu gimnazjalnego: przedział punktowy [0, 50]).

W celu porównania efektywności systemów lub ich elementów stosuje się także metody podziału populacji elementów na podzbiory grupujące elementy o wartościach reprezentacji liczbowej funkcji informacyjnej z odpowiedniego przedziału liczbowego. Najczęściej stosowane są skale znormalizowane: skala staninowa (*standard nine*) i skala zwana w języku polskim standardową piątką lub standardową skalą pięciostopniową (*standard five*). Wartości skali (kwantyle dzielące cały przedział dostępnych wartości funkcji informacyjnej) są konstruowane zgodnie ze standardowymi algorytmami. Wzorowane są na zasadzie podziału całej dziedziny rozkładu normalnego (rozkładu Gaussa–Poissona) na przedziały, o mniej więcej takiej samej szerokości. Określenie „mniej więcej” zostało przez nas użyte celowo: trudno bowiem jest pogrupować rzeczywiste wyniki dokładnie w takie przedziały, jakie są zakładane w warunkach idealnych.

Skala staninowa [38, 39, 6, 40] zakłada podział całego przedziału wynikowego (dostępnych wartości funkcji informacyjnej) na 9 przedziałów. W przedziałach, numerowanych kolejno od 1 do 9, które kolejno zawierają: 4% wyników najniższych, 7% wyników bardzo niskich, 12% wyników niskich, 17% wyników poniżej średnich, 20% wyników średnich oraz analogicznie w górę: 17% wyników powyżej średnich, 12% wyników wyso-

kich, 7% wyników bardzo wysokich i 4% wyników najwyższych. Podane nazwy jednocześnie definiują zmienne w reprezentacji lingwistycznej skali (oznaczamy ją symbolem SM_{ling}^{st}). Numery stanin są wartościami reprezentacji liczbowej skali (oznaczamy ją symbolem SM_{licz}^{st}). Zapis obu reprezentacji można przedstawić w następujący sposób:

$$SM_{ling}^{st} = \{\text{najniższy, bardzo niski, niski, poniżej średniego, średni, powyżej średniego, wysoki, bardzo wysoki, najwyższy}\}$$

oraz

$$SM_{licz}^{st} = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}.$$

Podobnie jak skalę staninową, tworzy się standardową pięciostopniową skalę zwaną też „standardową piątką”. Różnica między skalami polega na tym, że zamiast dziewięciu, tworzy się pięć przedziałów, które kolejno zawierają: 7%, 24%, 38%, 24%, 7% wyników uporządkowanych od wyniku najniższego do najwyższego. Poszczególne stopnie skali charakteryzowane są następująco: niski, poniżej średniego, średni, powyżej średniego, wysoki. Podane nazwy jednocześnie definiują zmienne w reprezentacji lingwistycznej skali (oznaczamy ją symbolem SM_{ling}^{st}). Numery ocen (od 1 do 5) są wartościami reprezentacji liczbowej skali (oznaczamy ją symbolem SM_{licz}^{st}).

Skale znormalizowane mają jedną szczególną cechę: do ich konstrukcji potrzebna jest populacja obiektów elementarnych o dużej liczebności. Taką skalę można skonstruować dla dużej szkoły lub okręgu szkolnego bądź dla całego kraju, co wymaga ujednoczenia metody sprawdzania poziomu kompetencji (ewaluacji) dla całej populacji. To właśnie było przyczyną wprowadzenia jednolitych dla całego kraju sprawdzianów kompetencyjnych w szóstej klasie szkoły podstawowej i egzaminów gimnazjalnych.

Skale znormalizowane są konstruowane dla każdego przypadku klasyfikacji oddzielnie, dla każdego sprawdzianu i egzaminu. Porównania wyników (efektywności działania) elementów lub systemów wymaga każdorazowo wyznaczenia wartości kwantyli dzielących cały zakres skali na przedziały wynikowe (np. staniny) dla każdego przypadku ewaluacji (np. dla każdego sprawdzianu czy egzaminu). Klasyfikacji dokonuje się względem skali zdefiniowanej dla wyższego poziomu hierarchii (np. wyznaczając efektywność działania szkoły bierzemy pod uwagę skalę staninową skonstruowaną dla okręgu lub całego kraju, a dla porównania wyników oddziałów szkolnych, posługujemy się skalą skonstruowaną dla szkoły lub dla okręgu lub też całego kraju).

Poniżej podajemy przykład takiej klasyfikacji klas III dla jednego z gimnazjów dokonanego w 2008 roku, na podstawie wyników egzaminów gimnazjalnych.

Przykład 1

W każdym egzaminie gimnazjalnym uczeń mógł uzyskać 50 punktów. Przedział 50- punktowy został podzielony na 9 przedziałów (stanin) zgodnie z przedstawionymi zasadami konstrukcji skali staninowej. Wyniki egzaminów gimnazjalnych (humanistycznego – oznaczo-

nego przez nas symbolem *EH* i matematyczno-przyrodniczego – oznaczonego przez nas symbolem *EMP*) przeprowadzonych w kwietniu 2008 roku i opublikowane przez Centralną Komisję Egzaminacyjną w postaci skali staninowej³ zebrano w tabeli 1 (skala staninowa do porównywania średnich wyników dla wszystkich gimnazjów w Polsce, skonstruowana na podstawie wyników średnich wszystkich gimnazjów w kraju) oraz w tabeli 2 (skala staninowa do porównywania wyników uczniów skonstruowana na podstawie wyników egzaminów gimnazjalnych wszystkich uczniów w kraju).

Tabela 1

Skala staninowa skonstruowana na podstawie wyników szkół z całego kraju

Stanin	Nazwa	Przedział punktowy dla egzaminu humanistycznego (EH)	Przedział punktowy dla egzaminu matematyczno-przyrodniczego (EMP)
1	najniższy	[00,0; 18,5)	[00,0; 15,3)
2	bardzo niski	[18,5; 25,3)	[15,3; 21,9)
3	niski	[25,3; 27,7)	[21,9; 23,9)
4	poniżej średniego	[27,7; 29,6)	[23,9; 25,6)
5	średni	[29,6; 31,4)	[25,6; 27,4)
6	powyżej średniego	[31,4; 33,2)	[27,4; 29,3)
7	wysoki	[33,2; 35,4)	[29,3; 32,1)
8	bardzo wysoki	[35,4; 39,3)	[32,1; 38,0)
9	najwyższy	[39,3; 50,0]	[38,0; 50,0]

W badanym gimnazjum, do obu egzaminów przystąpiło 148 uczniów.

Dla egzaminu humanistycznego uzyskali oni średni wynik 34,43 punktów, co plasuje gimnazjum w staninie 7 (wynik wysoki). Oznacza to, że wynik gimnazjum jest na pewno lepszy od wyniku 77% szkół w całym kraju i na pewno niższy od wyniku 11% szkół.

Dla egzaminu matematyczno-przyrodniczego uczniowie uzyskali średni wynik 31,25 punktów, co plasuje gimnazjum także w staninie 7 (wynik wysoki). Oznacza to, że wynik gimnazjum jest na pewno lepszy od wyniku 77% szkół w całym kraju i na pewno niższy od wyniku 11% szkół.

W przypadku uczniów można przypisać im wartość funkcji informacyjnej ze skali staninowej dla wszystkich uczniów w kraju (tab. 2). Można także dokonać tego posługując się skalą staninową skonstruowaną dla okręgu, która jest identyczna ze skalą skonstruowaną dla całego kraju⁴ lub skalą skonstruowaną dla badanego gimnazjum (tab. 3, opracowanie własne).

³ www.cke.edu.pl/index.php?option=content&task=view&id=141&Itemid=122

⁴ www.komisja.pl/info.php?i=24

Tabela 2

Skala staninowa skonstruowana na podstawie wyników uczniów z całego kraju

Stanin	Nazwa	Przedział punktowy dla egzaminu humanistycznego (EH)	Przedział punktowy dla egzaminu matematyczno-przyrodniczego (EMP)
1	najniższy	[0; 12]	[0; 10]
2	bardzo niski	[13; 17]	[11; 13]
3	niski	[18; 22]	[14; 17]
4	poniżej średniego	[23; 28]	[18; 22]
5	średni	[29; 34]	[23; 29]
6	powyżej średniego	[35; 38]	[30; 36]
7	wysoki	[39; 42]	[37; 42]
8	bardzo wysoki	[43; 45]	[43; 46]
9	najwyższy	[46; 50]	[47; 50]

Tabela 3

Skala staninowa do porównywania wyników uczniów skonstruowana na podstawie wyników uczniów z badanego gimnazjum

Stanin	Nazwa	Przedział punktowy dla egzaminu humanistycznego (EH)	Przedział punktowy dla egzaminu matematyczno-przyrodniczego (EMP)
1	najniższy	[0; 13]	[0; 13]
2	bardzo niski	[14; 19]	[14; 18]
3	niski	[20; 25]	[19; 22]
4	poniżej średniego	[26; 30]	[23; 29]
5	średni	[31; 35]	[30; 35]
6	powyżej średniego	[36; 39]	[36; 40]
7	wysoki	[40; 42]	[41; 45]
8	bardzo wysoki	[43; 44]	[46; 48]
9	najwyższy	[45; 50]	[49; 50]

Ciekawe spostrzeżenia można uzyskać analizując wartości funkcji informacyjnej dla uczniów przy porównywaniu ich wyników wyznaczanych dla różnych skal (SM_k^{st} – skala staninowa ogólnokrajowa, SM_G^{st} – skala staninowa dla badanego gimnazjum). Przykładowe wyniki dla egzaminu humanistycznego zebrano w tabeli 4.

Tabela 4
Porównanie przykładowych wyników uczniów badanego gimnazjum

Lp.	Liczba punktów uzyskana na egzaminie	$(SM_k^{st\ EH})$		$(SM_G^{st\ EH})$		Różnica $(SM_k^{st\ ZM}) - (SM_G^{st\ ZM})$
		$(SM_{licz}^{st\ ZM})$	$(SM_{ling}^{st\ ZM})$	$(SM_{licz}^{st\ ZM})$	$(SM_{ling}^{st\ ZM})$	
1	7	1	najniższy	1	najniższy	0
2	16	2	bardzo niski	2	bardzo niski	0
3	18	3	niski	2	bardzo niski	1
4	20	3	niski	3	niski	0
5	23	4	poniżej średniego	3	niski	1
6	26	4	poniżej średniego	4	poniżej średniego	0
7	30	5	średni	4	poniżej średniego	1
8	32	5	średni	5	średni	0
9	35	6	powyżej średniego	5	średni	1
10	38	6	powyżej średniego	6	powyżej średniego	0
11	39	7	wysoki	6	powyżej średniego	1
12	42	8	bardzo wysoki	8	bardzo wysoki	0
13	45	8	bardzo wysoki	9	najwyższy	-1

Skala krajowa jest „łagodniejsza” dla uczniów. Nie ma w tym nic dziwnego: przecież gimnazjum jako całość uzyskało wynik *wysoki* (reprezentacja liczbowa wyniku: 7), a więc jego uczniowie muszą plasować się w rankingu ogólnokrajowym wyżej.

Przedstawione metody pozwalają klasyfikować systemy lub ich elementy, ale nie opisują w pełni efektów pedagogicznych przez nie osiągniętych. Przykładem może być sytuacja, w której do dwóch różnych gimnazjów trafiają absolwenci szkół podstawowych o różnym poziomie (określanym na podstawie sprawdzianu kompetencyjnego uczniów klas 6 szkoły podstawowej jako wartości funkcji informacyjnej charakteryzującej ucznia po szkole podstawowej), a w czasie egzaminów gimnazjalnych przeprowadzanych trzy lata później, obie szkoły osiągają zbliżone wyniki.

W takim przypadku należy wyznaczyć wyniki egzaminów gimnazjalnych absolwentów gimnazjum w odniesieniu do odpowiadających wyników sprawdzianu przeprowadzo-

nego dla nich w szkole podstawowej, czyli wartość dodaną procesu edukacyjnego (edukacyjną wartość dodaną, którą oznaczamy symbolem EWD). Procedura, którą proponujemy, w tym przypadku jest dosyć prosta: należy dla odpowiednich skal oceniania (takich samych) określić różnicę między wartościami funkcji informacyjnej charakteryzującej absolwentów gimnazjum (f_i^G) i odpowiadającymi im wartościami funkcji informacyjnej uzyskanymi dla nich jako absolwentów szkół podstawowych (f_i^{SP}). Wynik tego działania pozwala określić wartości nowej funkcji informacyjnej (f_i^{EWD}):

$$f_i^{EWD} = f_i^G - f_i^{SP} \quad (1)$$

dla każdego elementu systemu (ucznia/studenta, nauczyciela itp.). Dla elementu wyższego poziomu można wyznaczyć wartości sumarycznej edukacyjnej wartości dodanej (f^{SEWD}): f^{SEWD} , sumując jej wartości dla wszystkich elementów niższego poziomu (np. w przypadku wyznaczania f^{SEWD} dla gimnazjum należy zsumować wartości EWD dla wszystkich jego uczniów branych pod uwagę). Ponieważ nie można wartości f^{SEWD} porównywać między szkołami, ze względu na różną ich liczebność, trzeba zdefiniować wartość dodaną przypadającą na pojedynczy element systemu (np. ucznia), nazwaną przez nas średnią edukacyjną wartością dodaną (f^{EWD}):

$$EWD_{sr} \equiv \overline{f^{EWD}} = \frac{f^{SEWD}}{N} \quad (2)$$

gdzie N jest ilością wszystkich elementów branych pod uwagę.

Przedstawioną procedurę wyznaczania edukacyjnej wartości dodanej, zapisaną symbolicznie równaniem (2), można zastosować w przypadku dowolnych elementów dowolnego systemu edukacyjnego.

W przykładzie 2 przedstawiliśmy wyniki dla gimnazjum z przykładu 1.

Przykład 2

Do określenia wartości funkcji informacyjnej f_i^{EWD} dla każdego ucznia klasy III badanego gimnazjum konieczna jest znajomość wartości skali staninowej dla egzaminu sprawdzającego przeprowadzonego w 2005 roku. Wartości skali dla całego kraju ($SM_{sp k}^{st}$)⁵ zebrano w tabeli 5, natomiast wartości skali lokalnej dla gimnazjum ($SM_{sp G}^{st}$), opracowane przez autorów, zebrano w tabeli 6.

Średni wynik punktowy ze sprawdzianu dla uczniów, którzy podjęli naukę w badanym gimnazjum w roku szkolnym 2005/2006 i przystąpili do egzaminów gimnazjalnych w roku 2008, wynosił 31,60 punktów, co pozwala przypisać gimnazjum wartość staniny 6 (wynik powyżej średniego): $f^{SP} = 6$.

⁵ www.cke.edu.pl/index.php?option=content&task=view&id=136&Itemid=107

Tabela 5

Skala staninowa wyników sprawdzianu kompetencyjnego przeprowadzonego w 2005 roku, dla uczniów szkół podstawowych w całym kraju (SM_{spk}^{st})

Stanin	Nazwa	Przedział punktowy dla sprawdzianu
1	najniższy	[0; 13]
2	bardzo niski	[14; 18]
3	niski	[19; 23]
4	poniżej średniego	[24; 28]
5	średni	[29; 32]
6	powyżej średniego	[33; 35]
7	wysoki	[36; 37]
8	bardzo wysoki	[38; 38]
9	najwyższy	[39; 40]

Tabela 6

Skala staninowa wyników sprawdzianu kompetencyjnego przeprowadzonego w 2005 roku, dla absolwentów szkół podstawowych, którzy przyszli do badanego gimnazjum (SM_{spG}^{st})

Stanin	Nazwa	Przedział punktowy dla sprawdzianu
1	najniższy	[0; 19]
2	bardzo niski	[20; 22]
3	niski	[23; 28]
4	poniżej średniego	[29; 32]
5	średni	[33; 35]
6	powyżej średniego	[36; 36]
7	wysoki	[37; 38]
8	bardzo wysoki	[39; 39]
9	najwyższy	[40; 40]

Średni wynik punktowy z egzaminu humanistycznego dla uczniów, którzy podjęli naukę w badanym gimnazjum w roku szkolnym 2005/2006 i przystąpili do egzaminów gimnazjalnych w roku 2008, wynosił 31,43 punktów, co pozwala przypisać szkole wartość staniny 5 (wynik średni): $f^{G EH} = 5$.

Obliczona edukacyjna wartość dodana dla egzaminu humanistycznego według skali dla całego kraju wynosi -1 ($f^{G\ EH\ EWD} = -1$).

Średni wynik punktowy z egzaminu matematyczno-przyrodniczego dla uczniów, którzy podjęli naukę w badanym gimnazjum w roku szkolnym 2005/2006 i przystąpili do egzaminów gimnazjalnych w roku 2008, wynosił 31,25 punktu, co pozwala przypisać szkole wartość staniny 7 (wynik wysoki): $f^{G\ EMP} = 7$.

Obliczona edukacyjna wartość dodana dla egzaminu matematyczno-przyrodniczego według skali staninowej dla szkół z całego kraju wynosi 1 ($f^{G\ EMP\ EWD} = 1$).

Wyniki obliczeń dla średnich wyników punktowych gimnazjum i skali staninowej skonstruowanej dla wszystkich szkół w kraju, przedstawiono w tabeli 7.

Dla obu egzaminów obliczono także sumaryczną edukacyjną wartość dodaną sumując wartości dodane dla wszystkich uczniów biorących udział w egzaminach gimnazjalnych ($f^{SEWD\ EH}$ oraz $f^{SEWD\ EMP}$) oraz odpowiadające im średnie edukacyjne wartości dodane ($f^{SEWD\ EH}$ oraz $f^{SEWD\ EMP}$). Obliczone wartości zebrano w tabeli 8.

Tabela 7

Wyniki obliczeń dla średnich wyników punktowych gimnazjum i skali staninowej skonstruowanej dla wszystkich szkół w kraju

	Średni wynik punktowy sprawdzianu	f^{sp}	Średni wynik punktowy egzaminu	f^G	f^{EWD}
Egzamin humanistyczny	31,60	6	31,43	5	- 1
Egzamin matematyczno-przyrodniczy			31,25	7	+ 1

Tabela 8

Wyniki obliczeń średniej edukacyjnej wartości dodanej dla badanego gimnazjum liczone według globalnej skali staninowej wyznaczonej dla uczniów z całego kraju

	f^{SEWD}	$\overline{f^{EWD}}$
Egzamin humanistyczny	- 81	- 0,55
Egzamin matematyczno-przyrodniczy	+ 21	+ 0,14

Różnica między wartościami funkcji informacyjnych $\overline{f^{EWD}}$ i f^{EWD} wynika, tak jak w przykładzie 1, z różnicy zastosowanej skali: w przypadku wartości funkcji $\overline{f^{EWD}}$ brana jest pod uwagę skala ogólnokrajowa dla uczniów, natomiast w przypadku wartości funkcji f^{EWD} brana jest pod uwagę skala dla uczniów, którzy rozpoczęli naukę w badanym gimnazjum w 2005 roku.

Uważamy, że wyniki uzyskane poprzez obliczenie średniej edukacyjnej wartości dodanej (f^{EWD}) mogą być mylące z uwagi na stosowanie niejednorodnych skal w trakcie obliczeń (wartości funkcji informacyjnych dla poszczególnych uczniów są obliczane według skali dla całego kraju, a wynik jest interpretowany w skali jednej szkoły). Jediną możliwością uniknięcia tego błędu jest zastosowanie jednolitej skali, czyli obliczanie jej wartości dla skal skonstruowanych dla gimnazjum. Dlatego też przeprowadzono obliczenia tym sposobem. Stosując taki algorytm obliczeń, uzyskano wyniki zebrane w tabeli 9.

Tabela 9

Wyniki obliczeń średniej edukacyjnej wartości dodanej dla badanego gimnazjum liczone według lokalnej skali staninowej gimnazjum

	f_i^{SEWD}	$\overline{f^{EWD}}$
Egzamin humanistyczny	+ 21	+ 0,14
Egzamin matematyczno-przyrodniczy	+ 20	+ 0,14

Dla uzupełnienia całości obliczeń, w tabeli 10, zebrano przykładowe wyniki obliczeń edukacyjnej wartości dodanej dla uczniów, którzy byli rozpatrywani także w przykładzie 1.

Dla tej grupy uczniów sumaryczna edukacyjna wartość dodana wynosi +4 ($f_i^{SEWDEH} = +4$), a średnia edukacyjna wartość dodana jest równa +0,31 ($f_{i\ sr}^{EWD\ EH} = +0,31$).

Tabela 10

Przykładowe wyniki obliczeń edukacyjnej wartości dodanej dla uczniów, którzy byli rozpatrywani także w przykładzie 1, z zastosowaniem lokalnej skali staninowej badanego gimnazjum

Lp.	Liczba punktów uzyskana na sprawdzianie	$SM_{sp\ licz}^{st}$	Liczba punktów uzyskana na egzaminie humanistycznym	$SM_{G\ licz}^{st}$	$f_{i\ sr}^{EWD\ EH}$
1	29	4	7	1	- 3
2	16	1	16	2	+ 1
3	30	4	18	2	- 2
4	22	2	20	3	+ 1
5	26	3	23	3	0
6	25	3	26	4	+ 1
7	36	6	30	4	- 2
8	28	3	32	5	+ 2
9	37	7	35	5	- 2
10	32	4	38	6	+ 2
11	36	6	39	6	0
12	36	6	42	8	+ 2
13	34	5	45	9	+ 4

Przedstawiane wyniki mają charakter przykładowy, dlatego też nie prezentujemy ich większej ilości. Badania zostały w całości wykonane na prośbę Dyrekcji badanego gimnazjum. Zarząd pedagogiczny szkoły chciał w ten sposób pozyskać informacje pozwalające oceniać pracę nauczycieli oraz pozwalające na konstrukcję rozsądnego planu rozwojowego szkoły.

4. Porównanie sposobów oceny efektywności działania systemów kształcących

Przedstawiona powyżej metoda oceny efektywności pedagogicznej gimnazjum różni się nieznacznie od metody proponowanej przez Centralną Komisję Egzaminacyjną (CKE) [5, 6, 13, 14, 28], szczególnie w odniesieniu do wyznaczania edukacyjnej wartości dodanej (EWD).

W naszej metodzie EWD jest wyznaczana z różnic między poziomem nabytych umiejętności i zdobytej wiedzy na zakończenie nauki w gimnazjum (egzamin gimnazjalne) i w szkole podstawowej (sprawdzian kompetencyjny). Oparta jest na założeniach: rzetelności i obiektywności pomiaru dydaktycznego w przypadku egzaminów i sprawdzianów zewnętrznych oraz niezależności i obiektywności wyznaczania poziomu umiejętności i zdobytej wiedzy metodą przyporządkowania wartości funkcji informacyjnej w postaci stanin. Nie wymaga dodatkowych obliczeń i szacunków. Algorytm naszej metody można zapisać w sposób następujący:

1. Wprowadzenie wyników egzaminów/testów wszystkich uczniów/studentów danej szkoły (okręgu lub innego zespołu – ogólnie: systemu kształcenia) biorących udział w egzaminie.
2. Wyznaczenie przedziałów staninowych dla wyników uzyskanych przed przystąpieniem do kształcenia w danej jednostce. W przypadku gimnazjum będą to wyniki sprawdzianu kompetencyjnego przeprowadzonego trzy lata wcześniej.
3. Przyporządkowanie odpowiedniej reprezentacji (liczbowej lub słownej) oceny początkowej (wejściowej) każdemu uczestnikowi (np.: $SM_{sp\ i}^{st}$ lub f_i^{SP}). Odpowiada to określeniu sygnału wejściowego w systemach automatyki i sterowania.
4. Wyznaczenie przedziałów staninowych dla wyników uzyskanych na egzaminach kończących edukację w danej jednostce. W przypadku gimnazjum będą to wyniki egzaminów.
5. Przyporządkowanie odpowiedniej reprezentacji (liczbowej lub słownej) oceny końcowej (wyjściowej) każdemu uczestnikowi (np.: SM_G^st lub f_i^G). Odpowiada to określeniu sygnału wyjściowego w systemach automatyki i sterowania.
6. Obliczenie różnic poziomów w postaci edukacyjnej wartości dodanej EWD_i jako różnicy między wartościami oceny końcowej (wyjściem) i oceny początkowej (wejściem) (np.: $f_i^G - f_i^{SP} = f_i^{EWD}$) dla każdego uczestnika. Odpowiada to wyznaczeniu parametru wzmocnienia elementu w systemach automatyki i sterowania.
7. Obliczenie sumarycznej efektywności jednostki (systemu) jako sumy wszystkich

EWD_i uczestników procesu (np.: $EWD \equiv f^{SEWD} = \sum_{i=1}^N f_i^{EWD}$) lub średniej efektywno-

ści jednostki (np.: $\overline{EWD} \equiv \overline{f^{EWD}} = \frac{f^{SEWD}}{N}$). Jest to równoznaczne z wyznaczeniem wartości funkcji informacyjnej dla jednostki i odpowiada wyznaczeniu parametru wzmocnienia układu elementów połączonych równolegle w systemach automatyki i sterowania.

Zaproponowana metoda wyznaczania edukacyjnej wartości dodanej jest możliwa do zastosowania dla dowolnego systemu kształcenia (poczynając od systemu jednostkowego uczeń-nauczyciel, a kończąc na ogólnokrajowym systemie oceny danego poziomu nauczania), także dla systemów e-nauczania i uczelni wyższych. Warunkiem jej zastosowania jest odpowiednie zdefiniowanie funkcji informacyjnej na wejściu i wyjściu z systemu. Po wprowadzeniu ujednoczonych i zewnętrznych: sprawdzianów kompetencyjnych, egzaminów gimnazjalnych i maturalnych, jest to bardzo proste w przypadku szkół podstawowych, gimnazjum i szkół ponadgimnazjalnych (taka zresztą myśl przewodnia przyświecała inicjatorom i realizatorom wprowadzenia jednolitych, zewnętrznych sprawdzianów i egzaminów). Natomiast w przypadku szkół wyższych bardzo trudno jest zdefiniować wyjściową funkcję informacyjną z uwagi na brak zewnętrznych, niezależnych i jednolitych egzaminów końcowych. Próbą rozwiązania tego problemu było wprowadzenie jednolitych, zewnętrznych egzaminów dla wybranych grup zawodowych (np.: prawników, urzędników państwowych, lekarzy), lecz w większości przypadków próby zakończyły się niepowodzeniem.

Proponowana metoda umożliwi także obliczenie edukacyjnej wartości dodanej na poziomie danej instytucji kształcącej (w prezentowanych przykładach wykorzystywane są staniny wyznaczone dla danego gimnazjum), a także względem poziomu wyższego (w prezentowanych przykładach wykorzystywane są staniny wyznaczone dla całego kraju).

W metodzie zalecanej przez CKE i stosowanej w przypadku sprawdzianów kompetencyjnych, egzaminów gimnazjalnych i maturalnych, edukacyjna wartość dodana dla każdego ucznia/studenta (f_i^G) wyznaczana jest jako różnica między punktami rzeczywistego wyniku egzaminu końcowego i wartościami oczekiwanymi tych punktów. Wartość oczekiwana punktów uzyskanych na egzaminie jest wyznaczana na podstawie estymacji zależności liniowej (jedno lub wieloparametrycznej) punktów z egzaminu końcowego dla danego poziomu kształcenia od punktów z egzaminu (lub sprawdzianu) końcowego dla niższego poziomu kształcenia [6, 13, 14]. W prezentowanych przykładach byłaby to zależność punktów uzyskanych przez ucznia w egzaminie gimnazjalnym od punktów uzyskanych przez niego w sprawdzianie, który składał 3 lata wcześniej. Sumaryczna edukacyjna wartość dodana (f^{SEWD}) jest wyznaczana jako suma różnic punktów, natomiast średnia edukacyjna wartość dodana ($\overline{f^{EWD}}$) jest średnią z tych różnic. Wartości są podawane w punktach i trudno je porównywać z poziomami wykształcenia podanymi w skali staninowej.

Wydaje nam się, że ten sposób obliczania efektywności systemów kształcenia jest zbyt skomplikowany i traktuje jednostkę przedmiotowo (uczeń rozpatrywany jest jako produkt jednostkowy w procesie technologicznym kształcenia w systemie). Niezrozumiała

jest także procedura estymacji wartości oczekiwanej punktów z egzaminu końcowego: dlaczego wartości te zostały potraktowane jako wartości wejściowej funkcji informacyjnej, skoro są obliczane dla końca procesu, a proces nie podlega sprzężeniu zwrotnemu? Wydaje nam się, że wynika to z bezpośredniego przełożenia metody szacowania wartości dodanej ze stosowanych nauk ekonomicznych [5, 13, 14], w których jest wykorzystywana do szacowania opłacalności produkcji. Po co jednak szacować w ten sposób wartości wejściowej funkcji informacyjnej w systemach kształcenia, skoro odpowiednie jej wartości, w tym samym układzie (czyli skali staninowej) można w prosty i bezpośredni sposób wyznaczyć na podstawie wyników egzaminów (sprawdzianów) przeprowadzanych we wcześniejszym etapie kształcenia? Określają przecież one poziom wykształcenia jednostki osiągnięty po ukończeniu wcześniejszego etapu kształcenia, a właśnie takie wartości powinna przyjmować wejściowa funkcja informacyjna.

Co prawda, zwolennicy metody propagowanej przez CKE uważają, że w trakcie estymacji wartości oczekiwanej punktów z egzaminu końcowego można uwzględnić różne czynniki wpływające na proces kształcenia [15], takie jak: status rodzinny, uwarunkowania społeczne, warunki nauczania itp.; ale wydaje nam się to niezbyt oczywiste i sensowne. Zgodnie z opisaną wcześniej zasadą separacji obiektów systemu kształcenia, wpływ tych czynników można określać innymi, niezależnymi metodami, a na pewno nie powinny to być uwzględniane przy wyznaczaniu edukacyjnej efektywności systemu.

Warto też zwrócić uwagę na fakt, że zaproponowana przez nas metoda umożliwia oszacowanie indywidualnej edukacyjnej wartości dodanej dla każdego ucznia/studenta, co nie jest możliwe w przypadku metody polecanej przez CKE [15 s. 70]. Dlatego też nasza metoda ma szerszy zakres zastosowań: może posłużyć na przykład do wyznaczenia efektywności pedagogicznej nauczyciela.

W celu porównania wyników zastosowania obu metod, w tabelach: 11 i 12 zebrano rezultatu obliczeń wykonanych obydwoma metodami. W tabeli 11 zestawiono wyniki obliczeń średniej efektywności edukacyjnej, a w tabeli 12 wartości EWD dla wybranych uczniów (tych samych co w przykładach 1 i 2).

Tabela 11

Porównanie rezultatów obliczeń średniej efektywności edukacyjnej rozpatrywanego gimnazjum zgodnie metodą zalecaną przez CKE ($\overline{f_{CKP}^{EWD}}$) i naszą metodą ($\overline{f^{EWD}}$)

	$\overline{f_{CKP}^{EWD}}$	$\overline{f^{EWD}}$
Egzamin humanistyczny	- 1,43	- 0,55
Egzamin matematyczno-przyrodniczy	+ 2,21	+ 0,14

Wartości zamieszczone w tabeli 11 wyraźnie wskazują na takie same tendencje, lecz wyniki osiągane zaproponowaną przez nas metodą bardziej odzwierciedlają sytuację gimnazjum opisaną w przykładzie 1.

W tabeli 12 umieściliśmy także różnice punktowe dla poszczególnych uczniów, choć, zgodnie z zaleceniami CKE nie powinniśmy tego robić. Jak wynika z porównania wartości z dwóch ostatnich kolumn, rzeczywiście nie umożliwiają nawet oszacowania zmiany poziomu wykształcenia indywidualnego uczestnika procesu.

Tabela 12

Przykładowe wyniki obliczeń edukacyjnej wartości dodanej egzaminu humanistycznego dla uczniów, którzy byli rozpatrywani także w przykładach 1 i 2, z zastosowaniem metody zalecanej przez CKE ($f_{iCKE}^{EWD\ EH}$) i naszej ($f_i^{EWD\ EH}$)

Lp.	Liczba punktów uzyskana na sprawdzianie	Liczba punktów uzyskana na egzaminie humanistycznym	$f_{iCKE}^{EWD\ EH}$	$f_i^{EWD\ EH}$
1	29	7	- 24,21	- 3
2	16	16	- 3,93	+ 1
3	30	18	- 11,23	- 2
4	22	20	- 4,89	+ 1
5	26	23	- 5,23	0
6	25	26	- 1,26	+ 1
7	36	30	- 8,76	- 2
8	28	32	+ 4,59	+ 2
9	37	35	- 4,89	- 2
10	32	38	+ 6,44	+ 2
11	36	39	+ 0,25	0
12	36	42	+6,04	+ 2
13	34	45	+11,27	+ 4

5. Wnioski

W celu oceny efektywności działania dowolnego systemu kształcenia (także jego poszczególnych elementów) jako systemu informatycznego należy wyznaczyć zbiór parametrów lub funkcję, na podstawie wartości których można dokonać tej oceny. Jest to jednoznaczne z definicją funkcji informacyjnej dla systemu kształcenia jako systemu informacyjnego i stosuje się w przypadku każdego systemu kształcenia, zarówno klasycznego, jak i wspomaganego technologiami komputerowymi (e-kształcenie). Separacja: elementów systemu (uczni, nauczyciela, medium komunikacyjnego i infrastruktury), procesów (komunikacji, przekazów, stymulacji i oceny) oraz uwarunkowań, która została zaproponowa-

na w tym opracowaniu, umożliwiła wyznaczenie niezależnych funkcji informacyjnych dla odseparowanych obiektów systemu (cząstkowych funkcji informacyjnych). Procedura umożliwia ewaluację warunkową ograniczoną do odpowiednich obiektów systemu.

Opierając się na zasadzie separacji obiektów systemu kształcenia, zaproponowaliśmy sposób wyznaczania edukacyjnej efektywności systemów kształcenia wykorzystująca zmodyfikowaną metodę edukacyjnej wartości dodanej. Zaproponowana metoda ma szerszy zakres zastosowań w porównaniu z metodą klasyczną: pozwala między innymi wyznaczać efektywność systemu w odniesieniu do jednostkowego elementu systemu (ucznia/studenta, nauczyciela) oraz jest możliwa do zastosowania na dowolnym poziomie kształcenia, w tym także w szkołach wyższych. Metoda jest prostsza obliczeniowo i nie wymaga dodatkowych estymacji. Pozwala także na jednoznaczne określenie wartości wejściowej i wyjściowej funkcji informacyjnej każdego elementu systemu kształcenia.

Rezultaty zastosowania nowego sposobu wyznaczania efektywności edukacyjnej systemów kształcenia porównaliśmy z wynikami zastosowania klasycznej metody polecanej przez Centralną Komisję Egzaminacyjną. Wyniki analizy ewidentnie potwierdzają przewagę nowej metody nad klasyczną.

Dalsze prace prowadzone są w kierunku określenia sposobów wnioskowania o efektywności pedagogicznej systemu kształcenia i prognozowania tego efektu dla tak zdefiniowanej funkcji informacyjnej. Prognozowanie efektów powinno być oparte na naszej zmodyfikowanej metodzie edukacyjnej wartości dodanej EWD oraz wykorzystywać wybrane metody sztucznej inteligencji w celu osiągnięcia efektu zrównoważonego i trwałego rozwoju systemu kształcenia.

Podziękowania

Autor pracy pragnie szczególnie serdecznie podziękować Dyrekcji 26 Gimnazjum Publicznego w Łodzi za udostępnienie danych, bez których nie byłoby możliwe przeprowadzenie obliczeń i uzyskanie wyników do analizy porównawczej.

Literatura

- [1] de Azevedo H.J.S., Scalabrin E.E., *A Human Online Learning Environment Using Intelligent Agents*. [w:] O.L. Fuhua (red.). *Distributed Learning Environments with Intelligent Software Agents*. Idea Group Inc. 2004.
- [2] Cader A., Przybyszewski K., Marchlewska A., *Systemy agentowe w e-learningu*. [w:] A. Cader *et al.* (red.). *Wybrane zagadnienia inżynierii wiedzy*. Wyd. SWSPiZ, Łódź, 2008, 9–32.
- [3] Cader A., *The Management System of Distance Education*. [w:] *Environmental Mechanics, Methods of Computer Science and Simulations*, t. 2: *Methods of Computer Science and Simulations*. Społom Press Institute of Mathematical Modeling, Lwów, 2004, 5–9.
- [4] Cader A., *Model danych dla systemu zarządzania nauczaniem zdalnym*. *Automatyka (półrocznik AGH)*, 3 (8), 2004, 635–642.

- [5] Czarnotta-Mączyńska J., Firsruk M., Lipska M., Lisiecka Z., *Analiza i interpretacja wyników oceniania i egzaminowania*. Teoria i praktyka egzaminowania, 1, Wydział Badań i Ewaluacji CKE, Warszawa 2007; także: www.cke.edu.pl/images/stories/badania/abc_stat_2a.pdf.
- [6] Dolata R. (red.), *Edukacyjna wartość dodana jako metoda oceny efektywności nauczania na podstawie egzaminów zewnętrznych*. CKE, Warszawa, 2007.
- [7] Ferguson G.A., Takane Y., *Analiza statystyczna w psychologii i pedagogice*. PWN, Warszawa, 2004.
- [8] Grandbastien M., *Teaching expertise is at the core of ITS Research*. International Journal of Artificial Intelligence in Education, 10, 1999, 335–349.
- [9] Guilford J.P., *Fundamental Statistics in Psychology and Education*. McGraw-Hill Book Company, New York, 1942.
- [10] Guilford J.P., *Podstawowe metody statystyczne w psychologii i pedagogice*. PWN, Warszawa, 1964.
- [11] Hornowska E., *Testy psychologiczne. Teoria i praktyka*. Wydawnictwo Naukowe Scholar, Warszawa, 2003.
- [12] Hyla M., *Przewodnik po e-learningu*. ABC Wolters Kluwer Polska, Kraków, 2007.
- [13] Jakubowski M., *Wpływ rzetelności pomiaru umiejętności uczniów na edukacyjną wartość dodaną*. [w:] *Uczenie się i egzamin w oczach nauczyciela*. Materiały XIV Konferencja Diagnostyki Edukacyjnej, Opole, 2008, 283–298.
- [14] Jakubowski M., *Metody szacowania edukacyjnej wartości dodanej*. [w:] *Edukacyjna wartość dodana*, cz. 2, Biuletyn Badawczy CKE, nr 14, Warszawa, 2007, 7–20.
- [15] Jakubowski M., *Metody szacowania edukacyjnej wartości dodanej*. [w:] *Edukacyjna wartość dodana*, Biuletyn Badawczy CKE, nr 8, Warszawa, 2006, 67–82.
- [16] Kolibabka M., Przybyszewski K., *Problemy ewaluacji w wirtualnym laboratorium edukacyjnym*. [w:] D. Rutkowska i inni (red.), *Selected Problems of Computer Science*. AOW EXIT, Warszawa, 2005, 591–602.
- [17] Krupski M., Cader A., *Rola kształcenia zdalnego w rozwoju społeczeństwa informacyjnego*. Automatyka (półrocznik AGH), 3 (10), 2006, 641–653.
- [18] Kruś L., *Problemy konstrukcji komputerowych systemów wspomaganie decyzji*. [w:] R. Kulikowski *et al.* (red.), *Systemowo-komputerowe wspomaganie zarządzania wiedzą*. AOW Exit, Warszawa, 2006, 141–151.
- [19] Niemierko B., *Jaki pomiar dydaktyczny jest nam potrzebny?* [w:] B. Niemierko i inni (red.), *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wydawnictwo WSHE, Łódź, 2003.
- [20] Niemierko B., *Ocenianie szkolne bez tajemnic*. WSiP, Warszawa, 2002.
- [21] Niemierko B., *Między oceną szkolną i dydaktyką. Bliżej dydaktyki*. WSiP, Warszawa, 1997.
- [22] Niemierko B. (red.), *Diagnostyka edukacyjna*, Wydawnictwo UG, Gdańsk, 1994.
- [23] Niemierko B., *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki obliczeniowe*. WSiP, Warszawa, 1975.
- [24] Niemierko B., Szyling G. (red.), *Holistyczne i analityczne metody diagnostyki edukacyjne. Perspektywy informatyczne egzaminów szkolnych*. Fundacja Rozwoju UG, Gdańsk, 2005; także: www.ptde.org.pl.
- [25] Niewiadomski A., Kryger P., Szczepaniak P.S., *Fuzzyfication of Indiscernibility Relation for Structurizing List of Synonyms and Stop-List for Search Engines*. [w:] L. Rutkowski i inni (red.), *ICAISC 2004, LNAI 3070*, Springer-Verlag, Berlin Heidelberg, 2004, 504–509.
- [26] Nwana H.S., *Intelligent tutoring systems: an overview*. Artificial Intelligence Review, 4, 1990, 251–277.
- [27] Pawlak Z., *Rough Set Elements*. [w:] L. Polkowski, A. Skowron (red.), *Rough Sets in Knowledge Discovery*. Springer Verlag, 1998.

- [28] Pokropek A., *Trafność metody edukacyjnej wartości dodanej*. [w:] Edukacyjna wartość dodana, cz. 2, Biuletyn Badawczy CKE, nr 14, Warszawa, 2007, 100–139.
- [29] Przybyszewski K., *A new method for validation of students' final exams*. 4th PD Forum-Conference on Computer Science, Smardzewice-Łódź, May 19–21, 2008.
- [30] Przybyszewski K., *Zastosowanie zbiorów rozmytych do ewaluacji różnych aspektów systemów kształcenia*. Automatyka (półrocznik AGH), 3 (12), 2008, 1033–1045.
- [31] Przybyszewski K., *A new evaluation method for e-learning systems*. [w:] L. Rutkowski *et al.* (red.). ICAISC 2006, LNAI 4029, Springer-Verlag, Berlin Heidelberg, 2006a, 1209–1216
- [32] Przybyszewski K., Filutowicz Z., Cader A., *Standaryzacja wyników ewaluacji studenta/ucznia dla potrzeb prognozowania efektywności klasycznych i e-learningowych systemów kształcenia*. XVII Krajowa Konferencja EDI-EC, Łódź-Rochna, 8–9 czerwiec 2009 (przyjęty referat).
- [33] Przybyszewski K., Cader A., Filutowicz Z., *Automatyzacja i obiektywizacja ewaluacji w procesach kształcenia*. [w:] A. Cader *et al.* (red.), Wybrane zagadnienia inżynierii wiedzy. Wydawnictwo SWSPiZ, Łódź, 2008, 36–76.
- [34] Przybyszewski K., Cader A., Filutowicz Z., *Zarządzanie informacją w interaktywnych systemach nauczania*. Zeszyty Naukowe WSHE 4 (9), 2000, 90–102.
- [35] Różewski P., Kusztnina E., Zaikin O., *Modele i metody zarządzania procesem otwartego nauczania zdalnego*. IBS PAN, Warszawa–Szczecin, 2008.
- [36] Rutkowski L., *Metody i techniki sztucznej inteligencji*. PWN, Warszawa, 2005.
- [37] Strumiłło G., Cader A., Przybyszewski K., Tonn K., *Experts' Knowledge Acquisition for Design of Expert System Supporting Students' Evaluation in Long-Time Learning/Teaching Process*. [w:] D. Rutkowska i inni (red.), Some New Ideas and Research Results in Computer Science. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2007.
- [38] Tadeusiewicz R., *Cybernetyczny model nauczania wspomagane komputerowo*. Automatyka (półrocznik AGH), 3 (8), 2004, 643–664.
- [39] Zaczyński W.P., *Statystyka w pracy badawczej nauczyciela*. Wydawnictwo Żak, Warszawa, 1997.