

Rekrutacja pod Nową Maturą

Wnioski z doświadczeń osobistych i nie tylko

Wstęp

Wprowadzenie od 2005 roku Nowej Matury zrewolucjonowało proces rekrutacji na wyższe uczelnie. Rezygnację z własnych egzaminów wstępnych wymusił prawodawca, a także praktyka Ministerstwa Edukacji, ograniczającego od 2006 roku możliwość przeprowadzania własnych „testów kompetencji” do rzeczywiście uzasadnionych przypadków. Choć entuzjazm wielu decydentów uniwersyteckich studziło ograniczone zaufanie do nowej matury – uzasadnione skądinąd jej rzeczywistą „nowością”, a więc, z jednej strony, ryzykiem związanym z brakiem doświadczenia jej twórców i organizatorów, z drugiej zaś – brakiem pozytywnych (bo i jakichkolwiek) doświadczeń odbiorców jej wyników, to wszyscy ludzie wyższych uczelni z ulgą przyjęli przejście do historii koszmaru egzaminów wstępnych, z ogromem pracy niedosypiających komisji rekrutacyjnych czy dezorganizującego życie tygodnia „wszystkich rąk na pokładzie”, gdy wszyscy młodszy pracownicy naukowcy zaganiani byli do twórczej pracy pilnowania kandydatów. Nie bez znaczenia były jednak także nie zawsze najlepsze doświadczenia z jakością egzaminów wstępnych, a także świadomość, że częstokroć nawet przy najlepszej woli komisji czy twórców testów egzaminacyjnych, rekrutacja na najbardziej obłążone kierunki niebezpiecznie zaczynała przypominać loterię – w której zapewne słabi kandydaci nie mieli szans na wygraną, ale także ci rzeczywiście dobrzy mogli przegrać z losowych powodów.

Nowa Matura zdjęła z uczelni znaczną część obciążeń organizacyjnych i odpowiedzialności za rzetelność egzaminacyjnego pomiaru. Jednocześnie jednak pojawiły się – lub zintensyfikowały – inne problemy, przede wszystkim dotyczące porównywalności i współmierności różnych egzaminów. Kolejnym problemem – nie nowym wprawdzie, lecz

zintensyfikowanym przez likwidację większości egzaminów wstępnych – jest rozszerzające się zjawisko współkandydowania na liczne kierunki, co w momencie, gdy dobry kandydat zostanie zaakceptowany na kilku wydziałach jednocześnie, prowadzi do organizacyjnego chaosu.

W naszej prezentacji przedstawimy problemy związane z rekrutacją pod Nową Maturą w świetle doświadczeń, jakie zdobyliśmy pełniąc funkcję sekretarzy kolejnych komisji rekrutacyjnych w Instytucie Socjologii UW w latach 2004-2006, a także współtworząc i obserwując funkcjonowanie systemów rekrutacyjnych w tym, oraz wcześniejszym i późniejszym okresie. Wykorzystamy również wyniki analiz danych z Internetowej Rejestracji Kandydatów dotyczących wszystkich osób zdających na Uniwersytet Warszawski. Pisząc o możliwych rozwiązaniach tych problemów przedstawimy te przetestowane – z lepszym lub gorszym skutkiem – w praktyce IS UW, inne, rozważane przez nas rozwiązania możliwe do wprowadzenia przez same wydziałowe komisje rekrutacyjne, ale także propozycje wymagające zmian systemowych na poziomie bądź uczelni, bądź w ogóle systemu egzaminacyjnego. Niektóre wydają się proste do wprowadzenia, niektóre zapewne mniej, mamy jednak nadzieję na przynajmniej zapoczątkowanie dyskusji nad nimi.

Problemy formuły rekrutacyjnej

Jak wiadomo, istnieją dwa typy egzaminów: te, których celem jest jedynie zweryfikowanie wiedzy egzaminowanego oraz egzaminy konkursowe. Z punktu widzenia uniwersytetu najlepsza byłaby sytuacja, gdyby o przyjęciu kandydata decydowało jedynie to, czy jest on wystarczająco przygotowany i sprawny intelektualnie do podjęcia studiów uniwersyteckich. W praktyce, niestety, o przyjęciu na studia decydują możliwości uczelni, która nie zawsze może przyjąć wszystkich zdających kandydatów, zaś obecnie – co wydaje się jeszcze bardziej niepokojące – także potrzeby finansowe: gdy za studentem idą pieniądze (jego lub państwowe), konieczne jest zapełnienie limitów przyjęć na studia dzienne i (zwłaszcza) wieczorowe, nawet kosztem obniżenia wymagań względem kandydatów. W tej sytuacji procedura rekrutacyjna musi przede wszystkim umożliwić stworzenie rankingu kandydatów – od najbardziej, do najmniej zdających.

Rozważmy możliwe problemy analizując konsekwencje trzech różnych sytuacji:

- Wszyscy kandydaci zdają ten sam, jeden egzamin.
- Wszyscy kandydaci zdają po dwa lub więcej, tych samych egzaminów.
- Kandydaci mogą wybrać jeden spośród kilku egzaminów.

Możliwe są również inne sytuacje. Ponieważ są one kombinacjami powyższych sytuacji – np. kandydaci zdają jeden wspólny egzamin oraz jeden egzamin do wyboru – i nie wnoszą one nic szczególnie istotnego do dalszych rozważań, odwoływać się będziemy do nich jedynie w razie potrzeby, koncentrując się na trzech podstawowych sytuacjach.

Wszyscy zdają jeden egzamin

Sytuacja, w której wszyscy zdają jeden egzamin wydaje się najprostsza (słusznie) i najlepsza (co już niekoniecznie). Zalety rozwiązania „jeden wspólny egzamin” są tak oczywiste, że nie wymagają omówienia. Skoncentrujemy się zatem na jego wadach

- Zwłaszcza w przypadku kierunków studiów, nie mających bezpośrednich odpowiedników wśród przedmiotów szkolnych, uzależnianie przyjęcia na studia od osiągnięć w tylko jednej dziedzinie nie wydaje się racjonalne. Taki sposób rekrutacji nie pozwala na zweryfikowanie kompetencji potrzebnych do studiowania, a co więcej nie jest sprawiedliwy.
- Ograniczenie kryteriów rekrutacji do jednego tylko przedmiotu szkolnego skutkowałoby wysłaniem do potencjalnych kandydatów błędnego sygnału o rzeczywistych wymaganiach na studiach, z reguły nie ograniczających się do jednej, wąskiej umiejętności.
- W interesie uczelni jest przyjęcie kandydatów o różnorodnych zdolnościach i umiejętnościach, gdyż różne specjalizacje w ramach nawet tych samych studiów mają różne wymagania.
- Choć przy jednym wspólnym egzaminie ustalony ranking kandydatów nie budzi wątpliwości, to liczba różnych pozycji w tym rankingu ograniczona jest do liczby możliwych różnych wyników egzaminu. Gdyby na egzaminie można było dostać od 0 do 50 punktów, zaś na dany kierunek startowałoby 1000 kandydatów, należałoby się spodziewać, że na „granicy przyjęcia” wspólną pozycję zajmowałoby kilkadziesiąt osób. Co to oznacza dla kierunku, na którym jest, powiedzmy, 100 miejsc, łatwo sobie wyobrazić.

Niezależnie od powyższych wad, system „jeden wspólny egzamin” w warunkach Nowej Matury jest właściwie niemożliwy w czystej postaci. Co prawda jest wiele kierunków studiów, które wymagają wyłącznie jednego przedmiotu – na Uniwersytecie Warszawskim są to przede wszystkim filologie: angielska (wymagany tylko j. angielski), polonistyka (j. polski), germanistyka (j. niemiecki), czy filologia ukraińska (j. polski). Pamiętajmy jednak, że różne egzaminy to nie tylko na przykład egzaminy z polskiego i matematyki. Różnymi egzaminami jest także Nowa Matura z matematyki z roku 2005 i 2006 (nie mówiąc już o sesjach poprawkowych), a tym bardziej Nowa Matura i Centralny Egzamin Wstępny (CEW)

na UW dla kandydatów ze starą maturą. I choć egzaminy maturalne są konstruowane w taki sposób aby zachowywać porównywalność pomiędzy latami, to jednak tworzenie egzaminów zachowujących zawsze ten sam poziom trudności jest właściwie niemożliwe. Zakładanie, że różne egzaminy z tego samego przedmiotu będą porównywalne jest niesłuszne, czego przykładem mogą być masowe protesty kandydatów zdających egzaminy CEW w rekrutacji na UW w roku 2006.

Wszyscy zdają te same egzaminy

Wymóg, by kandydaci zdawali kilka różnych egzaminów, łatwo jest uzasadnić. Studia wyższe wymagają różnorodnych umiejętności – np. studiowanie biologii wymaga nie tylko znajomości na poziomie szkolnym biologii, ale także chemii, po trochu też fizyki i matematyki oraz obcego języka. Dobry kandydat musi mieć wiedzę wszechstronną, zaś uczeń marzący o studiowaniu danego kierunku powinien już w szkole nastawić się na zdobywanie różnych umiejętności – a zakres egzaminów wstępnych jest także czytelnym komunikatem dla przyszłych studentów. Odwołanie się do różnych egzaminów przy tworzeniu jednego, wspólnego rankingu kandydatów wywołuje jednak trudności nawet wtedy, gdy egzaminy te zdają wszyscy kandydaci. Jaki bowiem ma być wpływ każdego z egzaminów na miejsce w rankingu?

Załóżmy, że z egzaminu A można było uzyskać od 0 do 20 punktów, zaś z egzaminu B – od 0 do 100 punktów, zaś o miejscu w rankingu decyduje suma uzyskanych punktów. Jest oczywiste, że na przyjęcia kandydatów wielokrotnie większy wpływ będą miały wyniki egzaminu B – osoby bardzo dobre z przedmiotu A nie będą miały szans z osobami beznadziejnymi z A, ale nieco lepszymi z B. Jeżeli nie jest to efekt zamierzony (egzamin A jest „pomocniczy”, zaś B „główny”), to trzeba by uznać, że procedura rekrutacyjna była niesprawiedliwa.

Czy wystarczy jednak, by z obu egzaminów przyznawano po 100 punktów? Niestety, niekoniecznie. Wyobraźmy sobie, że egzamin A okazał się bardzo łatwy i większość zdających uzyskała wyniki w granicach 80-100 punktów, zaś egzamin B był średnio trudny i zdający uzyskiwali z niego wyniki od 0 do 100 punktów. Efekt będzie taki sam jak w poprzednim przypadku, tyle, że tym razem będzie to efekt nieprzewidywalny przed egzaminem.

Proste sumowanie wyników różnych egzaminów zawsze wiąże się z ryzykiem niesprawiedliwego potraktowania kandydatów (co oznacza także nieoptymalny dla uczelni wybór studentów).

W warunkach Nowej Matury (przynajmniej 2004-2005) sytuacja dwóch jednocześnie zdawanych egzaminów dotyczyła także matury podstawowej i rozszerzonej.

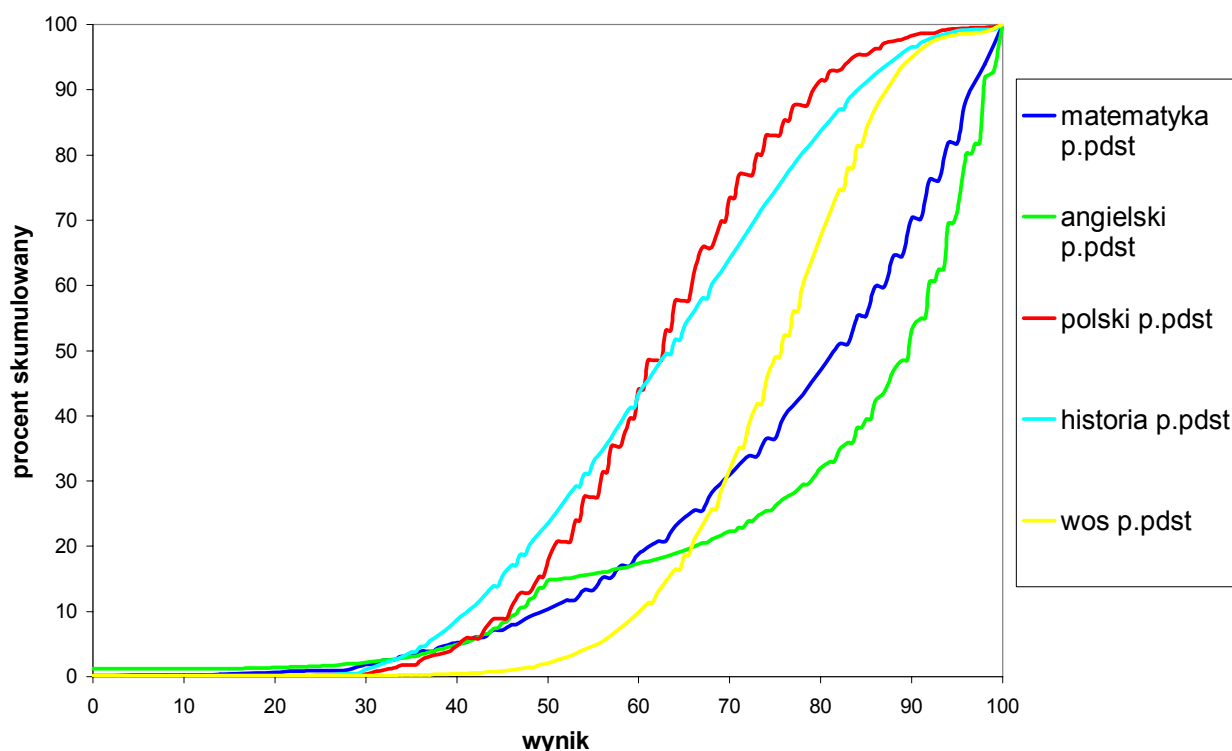
Kandydaci zdają różne egzaminy

Przy rekrutacji pod Nową Maturą jest to sytuacja powszechna. Kandydaci mogą zdawać maturę tegoroczną, zeszłoroczną lub w sesji poprawkowej. Kandydaci mogli również zdawać starą maturę oraz zastępujący Nową Maturę Centralny Egzamin Wstępny. Jednocześnie, na większości wydziałów dopuszcza się możliwość wyboru przez kandydatów przedmiotów egzaminacyjnych. Dotyczy to najczęściej egzaminów z języków obcych gdzie wymagany jest jeden z kilku języków. Takie rozwiązanie ułatwia nabór kandydatów o różnorodnych umiejętnościach, poszerza bazę potencjalnych studentów, a także stanowi łagodniejszą formę komunikatu o przyszłych wymaganiach. Przykładowo, kandydat na studia w IS UW, widząc, że jednym z dopuszczalnych przedmiotów egzaminacyjnych jest matematyka, powinien zauważyć, że najwyraźniej znajomość matematyki do czegoś na tych studiach jest przydatna.

Problem z wykorzystaniem wyników dwóch różnych egzaminów jest oczywisty. Jeśli z jednego egzaminu łatwiej jest uzyskać więcej punktów niż z drugiego – nie można mówić o równym traktowaniu egzaminowanych. Osoby zdające trudniejszy egzamin mają mniejsze szanse dostania się na studia. Jednak tego typu problem nie jest wcale rzadki. W 2005 roku dotyczył chociażby osób zdających na SGH, a w 2006 osób zdających CEW na UW.

Rzeczywiste zróżnicowanie wyników

Analizy wyników matur osób, które zdawały na Uniwersytet Warszawski w 2006 roku pokazują znaczące różnice pomiędzy przedmiotami. Poniższy wykres przedstawia skumulowany rozkład wyników matur z różnych przedmiotów na poziomie podstawowym. Poszczególne wartości pokazują jaki procent osób miało dany lub gorszy wynik. Rozkłady punktów z matury dla poszczególnych przedmiotów są bardzo różne. Uzyskanie dużej liczby punktów z matury z angielskiego było stosunkowo proste – prawie połowa zdających uzyskała 90 i więcej punktów. Dużo trudniejszy była matura z języka polskiego, czy historii gdzie tylko niewielka część zdających otrzymywała dużą liczbę punktów.



W konsekwencji tych różnic proste porównywanie punktów uzyskanych na maturze z różnych przedmiotów nie ma sensu. Faworyzuje ono osoby, które zdawały przedmioty łatwiejsze (w sytuacji gdy kandydaci zdają różne egzaminy) lub osoby lepsze z łatwiejszych przedmiotów kosztem lepszych z przedmiotów trudniejszych (gdy kandydaci zdają te same przedmioty).

Porównywalność i współmierność

Każda procedura porównywania wyników różnych egzaminów oraz wyznaczania wyników rekrutacji polega na pewnym przekształceniu wyników egzaminów, czyli w gruncie rzeczy sprowadza się do jakichś operacjach na liczbach. Warto się więc zastanowić jakie reguły powinny rządzić tymi operacjami. Z liczbami możemy zrobić wszystko – ale nie wszystko ma sens. Dwie liczby możemy do siebie dodać, bądź też powiedzieć, która jest większa – ale jeśli są one wynikami pomiarów, ma to sens tylko wtedy, gdy wyrażone są one w tych samych jednostkach. Choć pięć jest większe od dwóch, to pięć cali nie jest większe od dwóch metrów; nie ma też sensu stwierdzenie, że pięć watów to więcej niż dwa wolty, ani tym bardziej, że pięć watów plus dwa wolty to w sumie siedem.

Gdy chcemy porównać ze sobą poziom wiedzy matematycznej zmierzonej za pomocą różnych egzaminów, musimy się liczyć z tym, że jeden pomiar jest „w metrach”, zaś drugi – „w calach” – chyba, że potrafimy wykazać, że rzeczywiście dokonane zostały w tych samych jednostkach, bądź też potrafimy przeliczyć „cale” na „metry”. Gdy porównujemy poziom

wiedzy matematycznej i historycznej, musimy liczyć się z tym, że porównujemy „waty” z „woltami” – a jeśli chcemy twierdzić, że sprowadziliśmy je do wspólnej jednostki, powinniśmy umieć to uzasadnić.

Czy jest to w ogóle możliwe?

Obiektywnie rzecz biorąc, każda próba rozwiązania tego problemu będzie dalece niedoskonała, nie znaczy to jednak, że nie należy ich podejmować. Egzaminy wstępne trzeba *jakoś* przeprowadzić – i skoro nie ma co liczyć, że procedury będą w pełni racjonalne i obiektywnie sprawiedliwe to należy starać się, by były możliwie najracjonalniejsze i najsprawiedliwsze.

Generalnie, porównując (bądź też sumując) wyniki dwóch egzaminów możemy przyjąć jedno z dwóch założeń:

1. albo równoważne są egzaminy, czyli poziom ich trudności jest identyczny (a ściślej – takiemu samemu poziomowi kompetencji odpowiada taka sama liczba uzyskanych punktów)
2. albo równoważne (równie dobre) są populacje zdających oba egzaminy / w populacji rozkład (parametry rozkładu) poziomu wiedzy z przedmiotów egzaminacyjnych jest taki sam.

W pierwszym przypadku porównywać/sumować będziemy bezpośrednio wyniki egzaminów, zaś jednostka pomiaru będzie zależna od narzędzi testowych. W drugim, wyniki egzaminów poddawać będziemy wprawdzie jakiejś formie standaryzacji, zaś jednostka pomiaru będzie od rozkładu wyników w populacji zdających.

Słabość pierwszego rozwiązania ujawniła się z całą mocą w czasie rekrutacji na wyższe uczelnie w 2005 roku, gdy w szeregu przypadków mechaniczne porównywanie wyników różnych matur (np. z historii i geografii) doprowadziło do jawnie niesprawiedliwych decyzji i przynajmniej kilku ogólnopolskich skandali. Jednym z niewielu miejsc, gdzie rekrutacja przebiegła bez tego rodzaju zakłóceń był Instytut Socjologii UW, gdzie przyjęto drugie rozwiązanie. Zastosowanie pewnej specyficznej formy standaryzacji wyników egzaminów było efektem wieloletnich doświadczeń i ewolucji stosowanego w IS UW systemu rekrutacji. Ewolucji, która pomimo sukcesu z 2005 roku nie uległa zahamowaniu – na bazie tak pozytywnych, jak i negatywnych doświadczeń przyjęto wtedy zmodyfikowane rozwiązania na rok 2006, zaś doświadczenia z rekrutacji 2006 skłaniają do ich kolejnych modyfikacji.

Różne rozwiązania „standaryzacyjne”, ich wady i zalety w ewolucji i doświadczeniach systemu rekrutacyjnego IS UW

W dawnych, dobrych czasach rekrutacja do IS UW opierała się na egzaminach ustnych, w czasie których wypowiedzi zdających oceniane były przez komisję na skali ocen od 2 do 5.. Gdy jednak liczba kandydatów zaczęła się wyrażać nie w dziesiątkach, lecz w setkach i tysiącach, system taki odejść musiał do lamusa. Decydowała o tym z jednej strony techniczna niemożliwość ustnego przeegzaminowania takiej masy kandydatów, z drugiej zaś świadomość, że oceny wystawiane przez różne, pracujące równolegle komisje są nieporównywalne – choć zapewne osoba, która w jednej komisji dostała dwóję, w żadnej nie dostałaby piątki, to jednak za to, że czwórkowicz z jednej komisji jest rzeczywiście gorszy od piątkowicza z innej, żaden członek komisji rekrutacyjnej głowy by nie dał – a właśnie to, kto dostał czwórkę, a kto piątkę decydowało o przyjęciu na studia. W tej sytuacji konieczne stało się przejście na testowe egzaminy pisemne.

Pod koniec lat 90-tych kandydaci na socjologię zdawali egzamin z WOS, oraz, do wyboru, historię, polski lub matematykę. O pozycji na liście rankingowej decydowała suma punktów uzyskanych z obu zdawanych egzaminów. Efekt łatwo sobie przewidzieć. Dodajmy, że w tym wypadku okazało się, że to egzamin z historii przygotowany dla socjologii był wyjątkowo łatwy... W tej sytuacji, w kolejnym latach zdecydowano się na zmianę systemu. Poza rezygnacją z egzaminu z polskiego, podstawą reformą było zastosowanie punktów centylowych.

Punkty centylowe

Idea punktów centylowych jest zasadniczo prosta: kandydat z danego egzaminu otrzymuje liczbę punktów centylowych równą odsetkowi zdających ten egzamin, którzy uzyskali wynik nie lepszy od niego. Ten, kto zdał egzamin najlepiej, uzyska 100 punktów, kto zdał go nie gorzej, niż 80% konkurentów, uzyska 80 punktów itd. Tym samym, poziom trudności egzaminu nie ma znaczenia dla wyniku uzyskanego przez kandydata – jeśli a wie więcej niż b , to powinien uzyskać więcej punktów zarówno na łatwym, jak i trudnym egzaminie, zaś wynik kandydata zależy jedynie od jego pozycji w uporządkowaniu, a nie od liczby „surowych” punktów. Oznacza to także, że dopuszczalne staje się porównywanie wyników także między przedmiotami – kogoś, kto jest lepszy od 70% współzdających historię uznajemy za lepszego od osoby lepszej jedynie od 50% współzdających matematykę. Oczywiście – ma to sens, jeśli poziom wiedzy matematycznej i historycznej w populacji (populacjach) zdających przyjmiemy za równy (a więc przyjmujemy drugie założenie).

Ostateczny wynik kandydata określany był przez sumę punktów centylowych uzyskanych na egzaminie z WOS i egzaminie do wyboru.

System punktów centylowych ma swoje istotne ograniczenia, do których wrócimy potem. Nie one były raczej przyczyną, dla której system ten przetrwał w swojej pierwotnej postaci dość krótko, po czym został zastąpiony „systemem dwóch konkursów”.

System dwóch (lub wielu) konkursów

Jeśli kandydaci zdający matematykę i kandydaci zdający historię są – ogólnie rzecz biorąc – równie dobrzy, to ostatecznie wśród przyjętych liczby zdających historię i matematykę powinny być ze sobą w takiej samej proporcji, jak wśród kandydatów. Zasadę tę można odwrócić – skoro *przyjmujemy*, że kandydaci zdający matematykę i kandydaci zdający historię są – ogólnie rzecz biorąc – równie dobrzy, to *procedura rekrutacji powinna gwarantować*, by ostatecznie wśród przyjętych liczby zdających historię i matematykę *były* ze sobą w takiej samej proporcji, jak wśród kandydatów.

Aby to osiągnąć, procedura przeprowadzana jest w ramach dwóch¹, praktycznie osobnych konkursów. Kandydaci zdawali (razem) egzamin z WOS oraz (osobno) egzaminy z historii lub matematyki. Wyniki egzaminów - także egzaminu z WOS – ogłaszano jednak osobno dla konkursu „historycznego” i „matematycznego”. Wynik kandydata był sumą pozycji w rankingu z obu zdawanych egzaminach, ogłaszano jednak dwie osobne listy rankingowe, zaś limit przyjęć rozdzielano między obie listy proporcjonalnie do liczby wpisanych na nią osób².

System dwóch konkursów utrzymał się do roku 2004. Potrzeba jego zmiany związana była ze zjawiskiem masowego kandydowania na wiele kierunków i, w konsekwencji, masowych rezygnacji zakwalifikowanych kandydatów (którzy dostali się także na wyżej preferowane kierunki) oraz znaczne rozciągnięcie w czasie procesu rekrutacji.

Bezpośrednie przesłanki decyzji odejścia od systemu „dwóch konkursów” były dwie.

Pierwszą był fakt, że zakwalifikowani na studia kandydaci zdający matematykę blisko dwukrotnie rzadziej niż kandydaci zdający historię rezygnowali z podjęcia studiów. Założeniem systemu „dwóch konkursów” było zachowanie tych samych proporcji obu kategorii kandydatów tak wśród przyjętych, jak i ogółu kandydatów. Co to jest jednak „ogół kandydatów”? Czy należy zaliczać do kandydatów osoby, które w istocie nie są zainteresowane podjęciem studiów? Jeśli spośród kandydatów zdających historię, tylko

¹ Liczba konkursów zależy od liczby różnych egzaminów, gdyż dla każdego egzaminu prowadzony jest oddzielny konkurs. Dla uproszczenia w tym miejscu piszemy o dwóch egzaminach i konkursach.

² W sytuacji, gdy na jednej liście znajdowały się osoby zdające te same egzaminy, sumowanie pozycji w rankingu z tych egzaminów byłoby równoważne sumowaniu punktów centylowych; generalnie, punkty centylowe można by zastąpić „surową” pozycją w rankingu w sytuacji, gdyby porównywane rankingi uwzględniały dokładnie taką samą liczbę osób.

połowa zakwalifikowanych podejmuje studia, zaś wśród zdających matematykę – 75%, to czy proporcja przyjętych nie powinna uwzględniać tej różnicy?

Przesłanką drugą było wprowadzenie od 2005 Nowej Matury. Mieliśmy pełną świadomość (choć, jak życie pokazało, miał ją mało kto poza nami), że bezpośrednie uznanie porównywalności różnych matur, a także matur i Centralnych Egzaminów Wstępnych, byłoby niedopuszczalne. W tej sytuacji, „system dwóch konkursów” należałoby zastąpić „systemem czterech konkursów” (dla zdających Nową Maturę z historii, Nową Maturę z matematyki, CEW z historii, CEW z matematyki), a po egzaminach ogłosić cztery osobne listy rankingowe, wymagające osobnego administrowania w dalszym procesie rekrutacji. Ponieważ już administrowanie dwoma listami rankingowymi okazało się uciążliwe, zespół sekretarzy Instytutowej Komisji Rekrutacyjnej IS UW w składzie dr Joanna Konieczna, dr Mikołaj Jasiński i dr Jacek Haman, zaproponował „nowy” system – „system unormowanej pozycji”, pozwalający na umieszczenie kandydatów zdających różne egzaminy na jednej liście rankingowej.

System unormowanej pozycji – nowe doświadczenia z punktami centylowymi

„Nowość” systemu unormowanej pozycji była względna – w istocie, był to nieznacznie zmodyfikowany system punktów centylowych, aczkolwiek ujęty w postaci zupełnie innego algorytmu. Aby nie komplikować wyводу, zamiast podawać zmodyfikowany algorytm, wskażemy tylko na efektywną różnicę – system unormowanej pozycji był równoważny systemowi, w którym kandydat za dany egzamin uzyskuje liczbę punktów będącą średnią odsetka osób, które uzyskały wynik gorszy od niego, oraz odsetka osób, które uzyskały wynik nie lepszy (tj. gorszy lub taki sam) jak on³.

Ponieważ kandydaci zdawali po dwa egzaminy (do wyboru historia lub matematyka – w postaci Nowej Matury lub CEW – oraz wspólny dla wszystkich test kompetencji), ostateczna pozycja osoby na liście rankingowej zależała od sumy punktów centylowych uzyskanych na obu zdawanych egzaminach, przy czym punkty centylowe za test kompetencji wyznaczone były ze względu na pozycję wśród wszystkich zdających, natomiast za egzamin do wyboru – ze względu na pozycję wśród osób zdających dokładnie ten sam egzamin (tj. osobno wśród osób zdających CEW z matematyki, osobno Nową Maturę z matematyki itd.).

³ Aby nie komplikować opisu wprowadzaniem innego, ale równoważnego metodzie punktów centylowych algorytmu, w dalszym ciągu tekstu procedura z roku 2005 przedstawiana będzie tak, jakby stosowane były w niej punkty centylowe.

Efekt działania takiej procedury byłby identyczny do działania systemu „dwóch [wielu] konkursów”, gdyby populacje kandydatów zdających poszczególne egzaminy były identyczne ze względu na dwie cechy:

- poziom kompetencji mierzonych „testem kompetencji socjologicznych”;
- skłonność do podjęcia studiów w przypadku zakwalifikowania na nie.

Gdyby warunki te były spełnione, proporcje przyjętych zdających różne egzaminy do wyboru byłyby identyczne, jak analogiczne proporcje wśród ogółu kandydatów. Przeciwnie – gdyby, np., kandydaci zdający CEW z matematyki osiągnęli lepsze niż inni wyniki z testu kompetencji socjologicznych (co skądinąd miało miejsce), ich udział wśród zakwalifikowanych na studia byłby wyższy, niż wynikało to jedynie z ich udziału w ogóle kandydatów. Podobnie, gdyby kandydaci zdający CEW z matematyki w przypadku zakwalifikowania rzadziej niż inni rezygnowali z podjęcia studiów, ich udział także by rósł. Obie te własności systemu należy traktować jako pożądane i stanowiące o wyższości systemu punktów centylowych/unormowanej pozycji nad „systemem wielu konkursów”.

Zastosowanie systemu unormowanej pozycji w rekrutacji do IS UW w roku 2005 okazało się zdecydowanym sukcesem. Podczas gdy w wielu renomowanych uczelniach, gdzie bezpodstawnie przyjęto założenie o porównywalności „surowych” wyników różnych egzaminów Nowej Matury lub wewnątrzuczelnianych, na tle rekrutacji dochodziło do tarć czy wręcz skandali, odbijających się szerokim echem w prasie, rekrutacja w IS przebiegła bez zakłóceń, a procedury nie były kwestionowane.

Klasyczna standaryzacja

Mimo tego sukcesu, działający w tym samym składzie zespół Sekretarzy IKR IS UW zaproponował na rok 2006 jedną, istotną modyfikację: zastąpienie „unormowanych pozycji” (punktów centylowych), standaryzowanymi wynikami egzaminów.

Standaryzacja (tym razem w wąskim znaczeniu) polega na takim przekształceniu liniowym wartości zmiennej, by uzyskała ona średnią 0 i odchylenie standardowe 1. Wynik standaryzowany uzyskujemy poprzez podzielenie różnicy pomiędzy wynikiem danej osoby i średniej wyników wszystkich osób przez odchylenie standardowe wśród wszystkich osób. Ze względów praktycznych, zaproponowaliśmy zastosowanie formuły standaryzacji prowadzącej do średniej standaryzowanego wyniku równej 50 przy odchyleniu standardowym 10 (chodziło o uniknięcie sytuacji, gdy część kandydatów zobaczy, że uzyskali z egzaminów „ujemne” wyniki).

Standaryzacja, podobnie jak punkty centylowe, pozwala na „ucieczkę” od problemu jednostki, na której dokonano pomiaru, oraz na porównywanie wyników różnych pomiarów

różnych cech. Inaczej niż punkty centylowe, wykorzystuje ona nie tylko informacje o uporządkowaniu kandydatów ze względu na wyniki uzyskane przez nich na egzaminie, ale także informacje o wielkościach różnic między poszczególnymi kandydatami⁴.

Nie miało by to praktycznego znaczenia, gdyby problemem, przed którym stoją rekrutujący, było jedynie porównanie różnych kandydatów zdających różne egzaminy. Jednakże, w procedurach rekrutacyjnych do IS UW (jak i w wielu innych), obecny jest także inny element – generowanie wypadkowego wyniku z wyników różnych egzaminów zdawanych przez tę samą osobę. Wyznaczanie tej wartości jako sumy pozycji w rankingach lub punktów centylowych uzyskanych z różnych egzaminów można kwestionować ze względów teoretycznych, ale również można na prostych przykładach wskazać jego potencjalną niesprawiedliwość, związaną z niekorzystnym traktowaniem kandydatów, którzy z jednego z egzaminów uzyskali wynik wybitny, przy słabszym – ale wciąż niezłym – wyniku drugiego egzaminu.

Załóżmy – a wydaje się to założenie rozsądne – że „poziom kompetencji w dziedzinie A (*resp.* B)”, mierzony za pomocą egzaminu A (*resp.* B) jest cechą mającą w populacji rozkład normalny. Niech osoba x uzyska z egzaminu A standaryzowany wynik 2,1, zaś z egzaminu B 0,1. Niech osoba y uzyska z egzaminu A wynik 1,7, a z egzaminu B 0,2, a więc osoba x jest bardzo dobra (i wyraźnie lepsza niż y) w dziedzinie A, natomiast w dziedzinie B obaj kandydaci uzyskali bardzo zbliżony wynik, lokujący ich nieco powyżej średniej. Wydaje się jednoznaczne, że lepszym kandydatem jest x – ale gdy przyznamy im punkty centylowe, okaże się, że x uzyska za egzamin A 98 punktów, zaś za B – 54 punkty i w sumie 152 punkty, podczas gdy y uzyska odpowiednio 96, 58 i 154 punkty. Wynika to z faktu, że niewielkie różnice uzyskane z danego egzaminu przez kandydatów uzyskujących przeciętne, bliskie średniej, wyniki mają większe przełożenie na różnicę w punktach centylowych niż znacznie większe różnice między kandydatami z dobrymi i bardzo dobrymi wynikami: gęstość prawdopodobieństwa rozkładu normalnego jest największa w rejonie wyników bliskich średniej, a niewielka w rejonie wyników wysokich.

W powyższym przykładzie standaryzacja doprowadziłaby, oczywiście, do wyższej pozycji kandydata x niż y . Zauważmy przy tym, że w systemie rekrutacyjnym, w którym kandydaci zdają tylko po jednym (choć niekoniecznie tym samym) egzaminie, zastosowanie

⁴ Zastosowanie punktów centylowych wymaga jedynie, by pomiar kompetencji za pomocą testu egzaminacyjnego dokonywany był na skali porządkowej (rangowej), ignorując przy tym informacje uzyskiwane w przypadku pomiaru na mocniejszej skali przedziałowej (interwałowej). Standaryzacja wymaga, by „surowy” wynik egzaminu miał charakter zmiennej przedziałowej (i wykorzystuje walory mocniejszej skali). W przypadku dobrze przygotowanych egzaminów założenie, że ich wyniki określone są na skali przedziałowej wydaje się dopuszczalne.

standaryzacji i punktów centylowych da dokładnie ten sam efekt pod warunkiem, że rozkłady wyników obu egzaminów mają taki sam kształt, np. są rozkładami normalnymi (ale niekoniecznie o takiej samej średniej i wariancji).

Efekty zastosowania standaryzacji wyników egzaminów w rekrutacji w roku 2006 okazały się jednak nie w pełni zgodne z naszymi oczekiwaniami. Wynikało to z faktu, że rozkłady wyników różnych egzaminów okazały się znacznie mniej regularne, niż można by się tego spodziewać np. oglądając ogólnopolskie rozkłady wyników Nowej Matury. Choć w skali całego kraju rozkłady te są zbliżone do normalnego (co najwyżej lekko skośne), to w specyficznej populacji, jaką stanowili kandydaci na studia w IS UW część z tych rozkładów od normalności silnie odbiegała. W efekcie różnic w kształtach poszczególnych rozkładów, standaryzowane wyniki najlepszych kandydatów zdających różne egzaminy znacząco się od siebie różniły, a proporcje zakwalifikowanych na studia wśród osób zdających różne egzaminy – zwłaszcza w początkowym okresie rekrutacji, gdy kwalifikowane była zaledwie kilka procent kandydatów (omawiane efekty działały głównie na krańcach rozkładów) – znacznie odbiegały od proporcji wśród ogółu kandydatów.

Punkty centylowe czy klasyczna standaryzacja – a może coś pośredniego?

Problemy ze standaryzacją, obserwowane w 2006 roku, skłaniają do powrotu do koncepcji punktów centylowych. Możliwe jest jednak połączenie zalet standaryzacji z odpornością punktów centylowych na „dziwne” rozkłady wyników egzaminów. Rozwiązanie takie można nazwać „wymuszoną normalizacją”: Załóżmy, że dana osoba jest lepsza od 90% zdających egzamin (uzyskuje 90 punktów centylowych). Gdyby rozkład surowych wyników egzaminu był normalny standaryzowany, osoba taka uzyskałaby z niego wynik 1,28 (gdyż dystrybuanta rozkładu normalnego standaryzowanego wynosi 0,9 w punkcie 1,28). „Wymuszona normalizacja” przekształca punkty centylowe w wynik, jaki takiej wartości punktów centylowych odpowiadałby przy normalnym rozkładzie wyników egzaminu. Innymi słowy, kandydatowi lepszemu od 90% współzdających przypisujemy wynik 1,28 *niezależnie od tego*, czy rzeczywisty rozkład wyników egzaminu był normalny, czy też nie.

Dalsze problemy i propozycje na przyszłość

Założenie o „równoważności populacji” i jego ograniczenia. Postulowane zmiany w systemie maturalnym

Każda standaryzacja (tak standaryzacja *sensu stricte*, jak i rozwiązania w rodzaju punktów centylowych), wymaga przyjęcia, że przeciętne wyniki z poszczególnych egzaminów

uzyskane w populacjach zdających te egzaminy określają „równie dobry” poziom kompetencji badanych przez te egzaminy – a więc, kandydaci na socjologię zdający matematykę są przeciętnie równie dobrzy z matematyki, jak kandydaci zdający historię są dobrzy z historii. Twierdzenie takie można przyjąć jako aksjomat (skoro nie ma sensownego sposobu porównania wiedzy historycznej z matematyczną, to przyjęcie takiego założenia można uznać za zasadne), można też starać się weryfikować je pośrednio. Na przykład monitorując funkcjonowanie systemu rekrutacji na socjologię, kontrolowaliśmy średnie wyniki z testów kompetencji, kandydatów zdających matematykę i historię, a także zdających Nową Maturę i CEW. Fakt, że przeciętne wyniki ze zdanego przez wszystkich egzaminu w poszczególnych grupach nie odbiegały znacząco od siebie traktowaliśmy jako argument za utrzymaniem takiego założenia.

O ile jednak w przypadku porównywania zupełnie różnych egzaminów (jak matematyka – historia) założenie to można w ostateczności przyjąć arbitralnie, to w przypadku porównywania różnych egzaminów z tego samego przedmiotu, na dłuższą metę może być nie do utrzymania. Dodatkowo, o ile liczba typów egzaminów ze względu na przedmioty uwzględniane w rekrutacji będzie zawsze ograniczona (w przypadku IS UW wynosi ona dwa – matematyka i historia), to liczba egzaminów w ramach tego samego przedmiotu będzie wciąż rosła. Zauważmy, że w ciągu jednego roku są obecnie trzy sesje egzaminów maturalnych, zaś wśród ubiegających się o przyjęcie na studia znaleźć można osoby, które zdawała maturę w bieżącym, poprzednim, ale także jeszcze w kilku poprzednich latach. O ile założenie, że osoby zdające maturę w tym i poprzednim roku reprezentują podobny poziom wiedzy być może da się przyjąć na stałe, to czy można być tego pewnym w przypadku osób które zdawały egzamin w terminie poprawkowym, albo charakteryzują się tak nietypową ścieżką kariery edukacyjnej, że starają się o przyjęcie na studia w wieku 30 lat? Nie jest tu istotne, czy takie osoby miałyby być „lepsze” czy „gorsze” – ale trzeba się jednak liczyć z tym, że są „inne” niż „typowi” kandydaci. Ponadto, poza dwoma najpopularniejszymi kategoriami (tegorocznymi i zeszłorocznymi maturzystami), pozostałe grupy byłyby na tyle nieliczne, że z czysto losowych względów założenie, że przeciętnie reprezentują taki sam poziom kompetencji mogłoby być obciążone nieakceptowalnym błędem.

Pewnym rozwiązaniem tej sytuacji mogłoby się stać standaryzowanie wyników Nowej Matury nie względem populacji kandydatów na dany kierunek, ale względem ogółu zdających dany egzamin. Aby było to wykonalne, Centralna Komisja Egzaminacyjna musiałaby jednak udostępniać zbiorcze dane o wynikach matur nie później niż po ogłoszeniu wyników indywidualnych (co zapewne byłoby wykonalne, ale póki co nie jest czynione), bądź też na świadectwach maturalnych oprócz wyników „surowych” powinny być także podawane wyniki „standaryzowane” (np. w postaci punktów centylowych). Standaryzacja do populacji ogółu zdających nie jest jednak rozwiązaniem w pełni satysfakcjonującym, można

bowiem wskazać na przyczyny, dla których populacje zdających różne edycje egzaminów miałyby się od siebie różnić. Na przykład, założenie, że zdający egzamin w sesji poprawkowej nie są słabsi od zdających egzamin w sesji zasadniczej utrzymać byłoby dość trudno. Podobnie, po wprowadzeniu obowiązkowej matury z matematyki można się spodziewać, że przeciętny poziom kompetencji zdających matematykę będzie niższy, niż wtedy, gdy była przedmiotem wybieranym tylko przez chętnych. Z drugiej strony, można się spodziewać, że jeśli jakiś przedmiot w danym roku uznany zostanie za „łatwy” (mały odsetek abiturientów „polegnie”), w roku kolejnym stanie się popularniejszy wśród zwłaszcza słabszych uczniów i, w efekcie, w kolejnej edycji stanie się „trudny”, nawet, jeśli obie edycje egzaminu obiektywnie charakteryzować się będą tą samą trudnością.

Czy istnieje jeszcze inne rozwiązanie? Alternatywą dla założenia o „równoważności populacji” jest założenie o „równoważności egzaminów”. Analiza wyników różnych „edycji” Nowej Matury być może w przyszłości dostarczy argumentów za zaakceptowaniem Nowych Matur zdawanych w różnych terminach jako równoważnych – pamiętać jednak trzeba, że nawet jeśli będziemy dysponowali danymi potwierdzającymi względną równoważność egzaminów już przeprowadzonych, to nie dadzą one gwarancji w odniesieniu do egzaminów przeprowadzanych w przyszłości.

Czy można wobec tego uciec od obydwóch powyższych założeń? Przy aktualnie obowiązujących regułach prawnych raczej nie. Optymalne rozwiązanie wymaga zmiany systemu.

Zmiana ta polegałaby na częściowym rozdzieleniu funkcji Nowej Matury jako egzaminu dojrzałości i egzaminu wstępnego. O ile sam egzamin dojrzałości jest ważny przez całe życie abiturienta, nie ma konieczności, by jego ważność jako egzaminu wstępnego trwała dłużej niż rok. Już w tej chwili liczni kandydaci poprawiają wyniki egzaminu maturalnego, by zwiększyć swoje szanse przyjęcia na studia. Gdyby wszystkie osoby ubiegające się w danym roku o przyjęcie na studia zdawały odpowiednie egzaminy maturalne (wyłącznie jako egzaminy wstępne i wyłącznie te, które są uwzględniane w procesie rekrutacji) znikłby problem porównywania różnych edycji matury (uwzględniana byłaby wyłącznie „aktualna”). Czy jednak państwowy system egzaminacyjny przetworzy rocznie dodatkowe kilkaset tysięcy egzaminów? Można chyba przyjąć, że gdyby zapewnione były środki na pokrycie dodatkowych kosztów – w tym w szczególności na godziwe wynagrodzenie egzaminatorów, a być może także na wynajem dodatkowych pomieszczeń na egzaminy, to byłoby to wykonalne. A zapewnienie tych środków byłoby możliwe pod warunkiem dodatkowej zmiany systemowej – obciążenia kosztami zainteresowanych. O ile z konstytucyjnego prawa do bezpłatnej edukacji można wywieść prawo do darmowego egzaminu maturalnego, to nie ma podstaw do twierdzenia, że każdy ma prawo na koszt podatnika poprawiać w

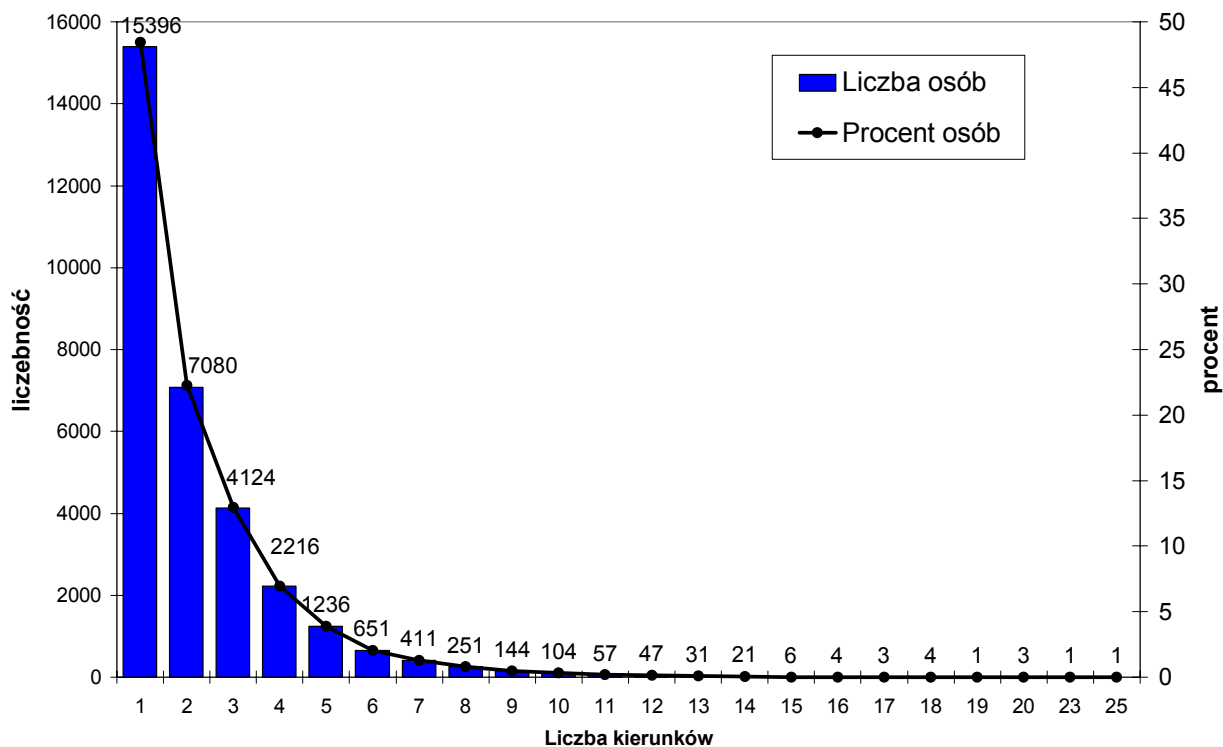
nieskończoność maturę! Dodatkowy egzamin maturalny – już po zakończeniu nauki w szkole średniej – powinien być „usługą” płatną, podobnie, jak „od zawsze” płatne były egzaminy wstępne na wyższe uczelnie (tzw. opłata administracyjna). Nie wydaje się przy tym, aby wysokość takich opłat, skalkulowanych tak, by pokryły koszty przeprowadzenia egzaminu, miała stanowić istotne ograniczenie w dostępie do wyższego wykształcenia.

Reforma taka, choć radykalna, rozwiązałaby definitywnie najpoważniejsze problemy związane z rekrutacją na studia na podstawie Nowej Matury. Nie wydaje się przy tym, by jej wprowadzenie w jakikolwiek sposób sprzeciwiało się ideom, które za Nową Maturą stoją.

Problem współkandydowania i postulowane zmiany w systemach rekrutacyjnych wyższych uczelni

Źródłem poważnych problemów organizacyjnych w czasie rekrutacji jest to, że większość kandydatów ubiega się o przyjęcie na kilka kierunków studiów. Nowa Matura przyczynia się do pogłębienia tego zjawiska pośrednio, ale zasadniczo – ponieważ te same egzaminy stają się podstawą do ubiegania się o miejsce na kilku wydziałach, spadają koszty kandydowania tak finansowe (obniżenie opłat administracyjnych), jak i związane z wysiłkiem (nie trzeba przygotowywać się do większej liczby egzaminów). A ponieważ na studia przyjmowani są – w miarę możliwości – najlepsi, ci, którzy zakwalifikowali się na jeden kierunek, mają znaczną szansę zakwalifikować się także na inne. W efekcie, znaczna część zakwalifikowanych kandydatów nie jest zainteresowana podjęciem studiów na danym kierunku, gdyż dostała się na inne, wyżej preferowane. Z kolej część kandydatów, którzy złożyli już dokumenty, wycofuje je, gdy miejsce na subiektywnie lepszym kierunku studiów uzyskają na dalszym etapie rekrutacji.

Podczas rekrutacji na Uniwersytet Warszawski w roku 2006 współkandydowanie okazało się poważnym problemem. Poniższy wykres przedstawia rozkład liczby kierunków na którą zdawali kandydaci na UW. Niespełna połowa (48,4proc.) kandydowała tylko na jeden kierunek studiów. Rekordzista kandydował na 25 kierunków studiów. Średnia liczba kierunków, na które zapisywała się jedna osoba to 2,21.



W warunkach masowego współkandydowania i konkurencji instytucji edukacyjnych o jak najlepszych dziennych, i jak najlepszych, a w dalszej kolejności o jakichkolwiek płatnych studentów wieczorowych studentów, rekrutacja staje się polem bitwy między poszczególnymi wydziałami i uczelniami, w której bronią są m. in. terminy ogłaszania wyników (najlepiej być pierwszym⁵), obowiązujące terminy składania dokumentów (tutaj nie jest do końca pewne, jaka strategia jest najlepsza), a czasem także – choć to środek nielegalny – utrudnianie kandydatom wycofywania dokumentów.

Bałagan i przedłużanie się okresu rekrutacji to nie tylko koszt administracji, a zwłaszcza członków komisji rekrutacyjnych. To także wielki koszt – głównie psychiczny – dla kandydatów, ale przede wszystkim czynnik, ograniczający efektywność rekrutacji – nie ma bowiem gwarancji, że rzeczywiście na poszczególne kierunki przyjęte zostaną najlepsze osoby chętne na studia w danym miejscu. Na przykład możliwe jest, że na wydział A dostanie się osoba x , na wydział B osoba y , pomimo, że osoby te byłyby skłonne zamienić się miejscami, a także oba wydziały byłyby zainteresowane taką zamianą.

Rozwiązaniem tego problemu byłoby stworzenie systemu wykorzystującego zgłoszone przez kandydatów (przy rejestracji) preferencje uporządkowujące kierunki, przyjęciem na które jest

⁵ Najprawdopodobniej jednak największe uczelnie Warszawskie uzgodnią ogłaszanie wyników rekrutacji tego samego dnia, począwszy od rekrutacji w roku 2007.

zainteresowany. System ten powinien pozwalać na centralne ustalenie podziału kandydatów pomiędzy kierunki, tak aby preferencje kandydatów były uwzględnione, a jednocześnie by żaden kandydat nie był jednocześnie zakwalifikowany na więcej niż jeden kierunek studiów. Szczegółowe zasady na jakich powinien opierać się taki system nie będą tu omawiane. Warto jednak zaznaczyć, że zasady te nie są wcale oczywiste. Nie rozwiązuje to także problemów związanych z kandydowaniem na różne uczelnie.

Zakończenie

Nowa Matura i związane z nią zmiany w rekrutacji na wyższe uczelnie stwarzają szansę na zafunkcjonowanie bardziej sprawiedliwych, bardziej efektywnych i wymagających mniej wysiłku od wszystkich uczestników procedur rekrutacyjnych. Jak w przypadku każdej nowości, generują one także ryzyko, że – jeśli wyniki nowych matur nie zostaną sensownie użyte – procedury te staną się mniej sprawiedliwe i efektywne oraz bardziej pracochłonne niż dawniej.

Założenie, że wyniki różnych egzaminów można dowolnie porównywać czy sumować jest grzechem pierworodnym wielu systemów rekrutacyjnych na wyższe uczelnie. Drugim, częstym źródłem problemów jest generalne niedocenywanie znaczenia procedur rekrutacyjnych. O ile jeszcze sam zestaw egzaminów wstępnych, bądź ich zakres czy poziom trudności, budzą niejednokrotnie emocje, to, co z ich wynikami będzie się działo później, nie jest traktowane jako kwestia istotna. Tymczasem jest to kwestia o znaczeniu zasadniczym. Planując i przeprowadzając proces rekrutacji należy korzystać z kumulujących się doświadczeń, tak własnych, jak i cudzych, ale także z wiedzy o właściwościach i problemach pomiaru dydaktycznego.

W prezentowanym tekście staraliśmy się przedstawić „skumulowane doświadczenia” rekrutacji w Instytucie Socjologii – tak jej sukcesy, jak i porażki, w nadziei, że pozwolą one innym uczyć się na naszych błędach i korzystać z naszych osiągnięć. Przedstawiliśmy także nasze propozycje innych rozwiązań, zarówno dostępnych samym jednostkom przeprowadzającym rekrutację (wydziałom, instytutom), jak i dotyczące zmian systemowych. Propozycje te są wynikiem przemyśleń wynikających głównie z doświadczeń praktycznych, ale także naszej wiedzy i doświadczeń związanych z teoretyczną statystyką, problemami pomiaru dydaktycznego czy sprawiedliwego podziału dóbr. Problem rekrutacji na wyższe uczelnie jest *praktycznym* problemem instytucji naukowych i pracujących w nich naukowców. Najwyższa pora, by zaczął być rozwiązywany z wykorzystaniem tych narzędzi, jakie daje nauka.