

**NACIONALNI CENTAR ZA VANJSKO  
VREDNOVANJE OBRAZOVANJA**

**Trg Marka Marulića 18, 10 000 Zagreb**

**ELABORAT O ODREĐIVANJU  
PRAGA PROLAZNOSTI NA  
ISPITIMA IZ PROBNE DRŽAVNE MATURE**

**dr. sc. Jasminka Buljan Culej**

**Zagreb, ožujak, 2009.**



**NACIONALNI CENTAR ZA VANJSKO  
VREDNOVANJE OBRAZOVANJA**

|  |    |
|--|----|
| <b>1. UVOD</b>   | 3  |
| <b>2. Metode temeljene na odlukama, standardima i prosudbama</b>                                   | 3  |
| 2.1 Odluke   | 3  |
| 2.2. Standardi   | 3  |
| 2.3. Prosudbe  | 4  |
| 2.4. Dva modela pogrešnih odluka   | 4  |
| <b>3. Metode temeljene na procjeni težine ispitnih zadataka</b>                                    | 5  |
| 3.1. Metoda Nedelsky   | 6  |
| 3.2. Angoffova metoda  | 7  |
| 3.3. Ebelova metoda  | 7  |
| 3.4. Odabir praga prolaznosti  | 8  |
| <b>4. Metode temeljene na procjeni individualnih postignuća na ispitu</b>                          | 8  |
| 4.1. Metoda granične grupe   | 10 |
| 4.2. Metoda kontrastne grupe   | 10 |
| 4.3. "Zaglađivanje" podataka   | 11 |
| 4.4. Odabir praga prolaznosti  | 12 |
| <b>5. Metode temeljene na procjeni skupnog postignuća na ispitu</b>                                | 13 |
| <b>6. Odabir jedinstvenih odrednica za postavljanje standarda praga prolaznosti</b>                | 14 |
| <b>7. Prijedlozi za određivanje praga na ispitima iz probne državne mature u Hrvatskim školama</b> | 15 |
| <b>8. Literatura</b>   | 18 |

## 1. UVOD

Određivanje praga prolaznosti na ispitima visokih uloga, kao što je primjerice državna matura, je važno pitanje kako za učenike tako i za cjelokupnu javnost. Transparentnost postupka i potkrijepljenost teorijama osiguravaju uvid cjelokupne zajednice u odabranu metodu određivanja praga prolaznosti. Iako se radi o važnoj temi, do danas još uvijek nije postignut dogovor oko toga koja metoda je najprimjerenija. Prilikom odabira praga prolaznosti moguće je očekivati nezadovoljstvo određenih skupina; oni koji su ispod praga prolaznosti smatrat će da je prag postavljen nepravedno visoko, dok oni iznad praga prolaznosti mogu smatrati da je postavljena razina apsurdno niska. Kako ne postoji potpuno objektivni način za postavljanje standarda prolaznosti, on će uvijek ovisiti o nekoj vrsti subjektivne prosudbe. Iako se standardi ne mogu objektivno odrediti oni se svakako mogu i moraju objektivno primjenjivati.

Da bi se objasnile razne teorije o postavljanju standarda za određivanje praga prolaznosti detaljnije će biti razrađene sljedeće teme:

1. Metode temeljene na odlukama, standardima i prosudbama
2. Metode temeljene na procjeni težine ispitnih zadataka
  - a. Metoda Nedelsky
  - b. Angoffova metoda
  - c. Ebelova metoda
3. Metode temeljene na procjeni individualnoga postignuća na ispitu
4. Metode temeljene na procjeni grupnog postignuća na ispitu
5. Odabir jedinstvenih odrednica za postavljanje standarda praga prolaznosti
6. Prijedlozi za određivanje praga na ispitu državne mature u Hrvatskim školama

## 2. Metode temeljene na odlukama, standardima i prosudbama

### 2.1. Odluke

Jedna od metoda određivanja praga prolaznosti je razmatranje rezultata pristupnika na npr. državnoj maturi i usporedba s ostalim informacijama o toj osobi (npr. ocjene iz srednje škole). Temeljem ovih činjenica moguće je donijeti prosudbu i odluku. Prednosti ove metode su višestruke; kriteriji se ne moraju unaprijed odrediti, u razmatranje se mogu uzeti sve relevantne informacije o pristupniku i postoji mogućnost individualnog razmatranja svakoga pristupnika. Metoda svakako ima i svojih nedostataka, a to su; subjektivnost iz nemogućnosti objektivnog odabira kriterija, pristupnici osjećaju nesigurnost jer se odlučivanje provodi „od slučaja do slučaja“, a postoji i mogućnost optužbi zbog arbitrarnog određivanja praga prolaznosti koje može biti utemeljeno na favoriziranju ili predrasudama.

Često se prilikom metode odlučivanja koristi pravilo klasificiranja pristupnika u dvije grupe: grupu s višim rezultatima i grupu s nižim rezultatima. Pravila o odlučivanju ove vrste koriste se u brojnim ispitnim situacijama, npr. grupa s višim rezultatima dobit će diplomu ili svjedodžbu, dok grupa s nižim rezultatima neće.

### 2.2. Standardi

Standard je odgovor na pitanje „Koliko dobro je dovoljno dobro?“. Kad se odabere prag prolaznosti tada je postavljen i standard za rezultate na ispitu.

Odabir praga prolaznosti ne bi bio problem kad bi rezultati pristupnika uvijek jasno pripadali dvjema grupama: u grupu sa savršenim rezultatima i u grupu s rezultatima koji su na ili blizu razine vjerojatnoga prolaska. Na žalost u stvarnom svijetu ispita rijetko dobivamo tako jasne rezultate, stoga se moramo suočiti s teškim zadatkom odlučivanja koliko je dobro dovoljno dobro. Standardi mogu biti apsolutni ili relativni. Relativni standard ovisi o usporedbama između pojedinaca dok apsolutni ne ovisi.

Rezultat pojedinog pristupnika uspoređivat će se prema standardu koji ovisi o rezultatima drugih pristupnika. Što su bolji rezultati drugih pristupnika, viši će biti standard. Nasuprot tome, apsolutni standard je onaj koji ne ovisi o rezultatima pristupnika koji će se prema tome ocjenjivati. Za osobu koja polaže ispit kod kojeg će se koristiti apsolutni

standard, nije važno kako će proći drugi pristupnici, jer njihovi rezultati ne utječu na standard.

Kad se kaže „Prag prolaznosti je 60 od mogućih 100“, to ne govori puno, osim ako se ne zna što „60“ znači. Ako to znači „60% odgovora je odgovoreno točno“, prag prolaznosti predstavlja apsolutni standard. Ako to znači „bolje od 60% pristupnika“ ili „dvije standardne devijacije ispod prosječnog rezultata pristupnika“, prag prolaznosti predstavlja relativni standard.

Temeljem prethodno navedenoga možemo zaključiti da prag prolaznosti predstavlja relativni standard – izabere se rezultat kojeg prelazi željeni broj ili postotak pristupnika i taj rezultat predstavlja standard.

### 2.3. *Prosudbe*

Odabir praga prolaznosti uključuje prosudbe u nekom trenutku postupka.

Važno je da te prosudbe:

donose osobe koje su kvalificirane za njihovo donošenje;

imaju smisla za osobe koje ih donose; i

napravljene su tako da uzimaju u obzir svrhu ispita.

Ova tri zahtjeva međusobno su povezana. Različite metode odabira praga prolaznosti traže različite vrste prosudbi te stoga donekle drugačije kvalifikacije ocjenjivača.

### 2.4. *Dva modela pogrešnih odluka*

Kad god se koriste ispiti za podjelu pristupnika u dvije grupe, mogu se dogoditi dvije vrste pogrešnih odluka:

Pristupnik koji zapravo pripada u nižu grupu može dobiti rezultat iznad praga prolaznosti.

Pristupnik koji zapravo pripada u višu grupu može dobiti rezultat ispod praga prolaznosti.

Ove pogrešne odluke događaju se jer ispiti skoro nikad nisu savršena mjera znanja i vještina koje se njima mjere. Vještine pristupnika mogu se razlikovati od dana do dana i čak od sata do sata. Pristupnik može pogoditi odgovor na neko od pitanja i nema načina da se razdvoji slučajno pogađanje od odgovora koji je pristupnik zaista znao. Za većinu ispita, zadaci ili problemi ne uključuju svaki element znanja i svaku moguću primjenu vještina koje ispit namjerava mjeriti. Zadaci ili problemi samo su uzorak svih onih koji su mogli biti uključeni i mogu dati pogrešnu sliku o vještinama nekih pristupnika.

Zbog svih ovih razloga, na većini ispita nemoguće je izabrati prag prolaznosti koji bi u potpunosti eliminirao pogrešne odluke. Može se smanjiti vjerojatnost za prolaz pristupniku koji bi inače pao kad bi se koristio viši prag prolaznosti. Međutim ako se odluči za ovu metodu, povećat će se vjerojatnost za pad na ispitu pristupniku koji bi inače prošao. Isto tako može se smanjiti mogućnost za padanje pristupnika koji je trebao proći da se koristio niži prag prolaznosti, ali će se time povećati mogućnost pristupniku koji je trebao pasti. Unaprjeđenjem strukture ispita smanjit će se broj pogrešnih odluka, ali neće ih se ukloniti u cijelosti.

### **3. Metode temeljene na procjeni težine ispitnih zadataka**

Ove metode temelje se na ideji da prag prolaznosti treba biti rezultat koji se očekuje od osobe čije vještine su na granici (Nedelsky, 1954). Granični pristupnik je onaj pristupnik čije znanje i vještine se nalaze na granici između gornje i donje grupe pristupnika. Metode su relativno za primjenu bilo prije ili poslije provođenja ispita. Osim toga, u postupku donošenja prosudbi o ispitnim zadacima pažnja ocjenjivača fokusira se usko na sadržaj ispita. Najvažnije je da se potrebni podaci - prosudbe o ispitnim zadacima mogu gotovo uvijek dobiti. Međutim, tip prosudbe koje ove metode traže nije jednostavno vrjednovanje nečijih rezultata koje ocjenjivač provodi. Ocjenjivači moraju donijeti odluku kako bi granični pristupnik odgovorio na svako od ispitnih zadataka. Zbog hipotetske prirode ovih prosudbi, smatramo da ove metode trebaju „realnu provjeru“. Kod primjene ove metode potrebna je se dopuna s nekom vrstom informacija o stvarnim rezultatima pristupnika. Ako ova informacija jasno pokaže da rezultati metode ne opisuju rezultate graničnog pristupnika, treba biti spreman priznati da metoda možda nije najbolja za određeni ispit i odabrati prag prolaznosti na neki drugi način. Prvi korak u bilo kojoj od navedenih metoda je odabir ocjenjivača. Kvalificirani ocjenjivači moraju odlučiti koja je razina znanja ili vještina koje se ispitom mjere potrebna za prag prolaznosti.

Na primjer, ako se test čitanja koristi na završnom ispitu u srednjoj školi, ocjenjivači moraju biti kvalificirani odlučiti što se u školskoj diplomi treba navesti o sposobnosti osobe za čitanje. U nekim slučajevima ima svega nekoliko ljudi koji mogu biti ocjenjivači; u drugim slučajevima mnogi mogu biti kvalificirani. Ako samo nekoliko ljudi ima kvalifikacije treba ih, ako je moguće, sve uključiti kao ocjenjivače. Ako to nije moguće, treba osigurati da su ocjenjivači koji sudjeluju tipični za sve osobe koje su kvalificirane biti ocjenjivači. Sve važne teme trebaju se predstaviti na sastanku ocjenjivača.

Koliko ocjenjivača treba izabrati? Ako ih ima premalo, na postupak može biti bitno utjecati jedna ili dvije osobe koja ima neuobičajeno visoke ili niske standarde. U tom smislu, što je više ocjenjivača, to bolje.

Kako bi se odredilo „granično“ znanje i vještine, potrebno je provjeriti da li ocjenjivači razumiju što ispit mjeri i kako će rezultati ispita biti korišteni. Nakon toga ocjenjivači bi trebali biti sposobni vlastitim riječima opisati osobu čije bi znanje i vještine bile granične između prihvatljive i neprihvatljive razine znanja i vještina koje se ispitom mjere. Ocjenjivači mogu smatrati prikladnim opise rezultata određenih ljudi s kojima su radili, a koje oni klasificiraju kao „granične“. Postavljanjem odgovarajućih pitanja može im se pomoći u postupku. Ocjenjivačima treba osigurati dovoljno vremena da se slože oko definicije graničnog znanja i vještina. Ako postoji bitna razlika u mišljenjima koja se ne mogu riješiti kompromisom, potrebno je nastaviti bez definicije oko koje bi se grupa ocjenjivača mogla složiti. U svakom slučaju potreban je sporazum, ako je moguće. Kad su se ocjenjivači dogovorili oko definicije, potrebno je ih je pismeno dokumentirati zajedno s primjerima, tako da postoji pisani dokument o standardu kojeg bi prag prolaznosti trebao predstavljati.

Tri metode koje će biti objašnjene nazvane su po osobama koje su ih prve predložile u svojim knjigama i člancima o edukacijskim mjerenjima. Ove metode poznate su kao metoda Nedelsky, Angoffova metoda i Ebelova metoda. Svaka od te tri metode zahtijeva drugi tip prosudbe.

### *3.1. Metoda Nedelsky*

Ovu metodu Leo Nedelsky predložio je 1954. godine i može se koristiti samo kod zadataka s više ponuđenih odgovora, obzirom da zahtijeva prosudbu o svakom netočnom odgovoru. Zadatak ocjenjivača je pogledati pitanje i identificirati pogrešan odgovor koji bi granični pristupnik smatrao krivim, to jest ne najboljim od ponuđenih odgovora.

Ocjenjivači bi trebali davati svoje mišljenje pojedinačno ali i pokušati postići kompromis. Kada broj ocjenjivača nije prevelik, tada davanje pojedinačnih mišljenja, isto kao i postizanje kompromisa daje dobre rezultate. Ocjenjivači mogu donijeti bolje prosudbe ako dijele informacije i mišljenja jedni s drugima. Jedno ograničenje kod ovog postupka je to što traži od svih ocjenjivača da svoje mišljenje daju istovremeno i na istom mjestu. Drugo ograničenje je da je postupak prilično spor (iako ni izbliza tako spor kao kad se pokušava postići grupni konsenzus o svakom pitanju).

Metoda Nedelsky temelji se na ideji da granični pristupnik reagira na zadatke s više ponuđenih odgovora tako da prvo eliminira odgovore koje

on ili ona prepoznaje kao netočne i onda pogađa naslijepo između preostalih odgovora. Ako se ispit ocjenjuje bez korekcije za pogađanje, relativno je lako pronaći rezultat koji se očekuje od pristupnika, kroz primjenu sljedećih pravila.

Kod metode Nedelsky, očekivani rezultat pristupnika za bilo koje pitanje je 1 podijeljeno s brojem odgovora na koje pristupnik mora pogađati.

Da bi se saznao očekivani rezultat pristupnika za cijeli ispit, dodajte očekivani rezultat pristupnika za pojedinačna pitanja.

Na primjer, ako je granični pristupnik eliminirao sve osim tri moguća odgovora, on ili ona ima 1:3 šanse za pogađanje točnog odgovora. Stoga je njegov ili njen očekivan rezultat za to pitanje 1 podijeljeno s 3 ili 0,33.

### 3.2. *Angoffova metoda*

Metoda koju je predložio William H. Angoff 1971. godine slična je metodi Nedelsky, s razlikom da se može koristiti na ispitima kod kojih nema zadataka s više ponuđenih odgovora. Kod Angoffove metode prag prolaznosti izračunava se iz očekivanih rezultata za individualna pitanja, kao i kod metode Nedelsky. Međutim Angoffova metoda ne traži od ocjenjivača razmatranje svakog mogućeg netočnog odgovora pojedinačno. Umjesto toga, ocjenjivač razmatra svako pitanje kao cjelinu i donosi prosudbu o vjerojatnosti da će granični pristupnik odgovoriti na pitanje točno. Ovaj zadatak može biti težak za neke ocjenjivače. Ako zadaci imaju više ponuđenih odgovora, vjerojatnost bi obično trebala biti veća od mogućnosti da se točan odgovor pogodi pukom srećom (to jest, 1,00 podijeljeno na broj izbora).

Pronalaženje očekivanog ispitnog rezultata za graničnog pristupnika radi se zapravo na isti način kao kod metode Nedelsky. Ako je rezultat bez korekcije za pogađanje, vjerojatnost točnog odgovora je očekivani rezultat pristupnika za to pitanje.

### 3.3. *Ebelova metoda*

Za razliku od prethodne dvije metode, Ebelova metoda predstavlja postupak u dvije faze. Svaki ocjenjivač prvo klasificira pitanja u grupe i onda radi pojedinačnu brojčanu prosudbu za svaku grupu pitanja. Podjela pitanja u grupe temelji se na dvije vrste prosudbi o svakom pitanju: prosudba njegove težine i prosudba relevantnosti (ili važnosti). Ebel je predložio tri razine težine, nazvavši ih „lagano“, „srednje“ i „teško“, te četiri kategorije relevantnosti, nazvavši ih „ključno“, „važno“,



„prihvatljivo“ i „upitno“. Prvi zadatak ocjenjivača je klasificirati sva pitanja iz ispita. Drugi zadatak ocjenjivača je donijeti prosudbu o rezultatima graničnog pristupnika. Ocjenjivač mora donijeti takvu prosudbu za svako od 12 polja u klasifikacijskoj tablici (osim za ona koja su prazna). To znači da ocjenjivač mora donijeti jednu prosudbu za pitanja klasificirana „ključno, lagano“, drugu za „ključno, srednje“ i tako dalje sve do „upitno, teško“. Prosudba se sastoji od odgovora na pitanje: „Ako granični pristupnik mora odgovoriti na veliki broj pitanja kao što su ova, koji postotak bi on ili ona odgovorio/la točno?“

Grupni postupak prikupljanja prosudbi koji je preporučen za metodu Nedelsky i Angoffovu metodu može se prilagoditi i za Ebelovu metodu. Međutim to bi bilo suviše složeno jer ocjenjivači moraju donijeti dvije odluke o svakom ispitnom pitanju: o težini i relevantnosti, te moraju donijeti prosudbu o rezultatu graničnog pristupnika u svakoj od 12 grupa pitanja. Ako se koristi ovaj postupak kod Ebelove metode, uputno ga je primijeniti odvojeno na svaku od dvije faze Ebelove metode.

### *3.4. Odabir praga prolaznosti*

Da bi se dobio očekivani rezultat ispita za graničnog pristupnika, potrebno je koristiti sljedeći postupak:

- Pomnožiti prosuđeni postotak točnosti za prvu kategoriju („ključno, lagano“) prema broju pitanja u toj kategoriji da bi se dobio očekivani rezultat graničnog pristupnika za prvu kategoriju.
- Ponoviti prvi korak za svaku od preostalih 11 kategorija.
- Dodati očekivani rezultat za 12 kategorija kako bi se dobio očekivani rezultat za cijeli ispit.

Mogu se kombinirati rezultati koje su izračunati za pojedinačne ocjenjivače na isti način kao kod metode Nedelsky, izračunavanjem aritmetičke sredine, medijana ili trimmed aritmetičke sredine.

## **4. Metode temeljene na procjeni individualnih postignuća na ispitu**

Ove metode predstavljene temelje se na informacijama o pojedinim pristupnicima. One zahtijevaju dvije vrste informacija o svakom pristupniku:

- Rezultat ispita osobe.
- Prosudbu o adekvatnosti znanja i vještina pristupnika.

Metode uključuju metodu „granične grupe“ i metodu „kontrastne grupe“. Glavna prednost ovih metoda je da su ljudi u našem društvu navikli ocjenjivati vještine drugih ljudi kao adekvatne ili neadekvatne iz nekog razloga – posebice u obrazovnom i profesionalnom okruženju. Prilikom primjene ovih metoda potrebno je zadovoljiti sljedeće zahtjeve:

- Prosudbe trebaju obavljati osobe koje su za njih kvalificirane.
- Prosudbe se trebaju baviti znanjem i vještinama koje test namjerava mjeriti.
- Prosudbe moraju odražavati vještine pristupnika u vrijeme ispita.
- Prosudbe trebaju odražavati stvarno mišljenje ocjenjivača.

Prvi zahtjev odnosi se na bilo koju metodu odabira praga prolaznosti: prosudbe moraju obavljati kvalificirane osobe. Kod metoda koje se temelje na prosudbi pojedinačnih pristupnika, potrebne su dvije vrste kvalifikacija:

- Ocjenjivači moraju biti sposobni odrediti znanje i vještine svakog pristupnika.
- Ocjenjivači moraju znati koju razinu znanja i vještina koju pristupnik treba imati da bi prošao na ispitu.

Važno je da ocjenjivači posjeduju obje navedene kvalifikacije.

Drugi zahtjev je da se prosudbe moraju temeljiti na vještinama i znanju koje se namjeravaju mjeriti ispitom. Problem je što prosudbe individualnih vještina mogu biti podložne djelovanju čimbenika koji su nebitni za svrhu ispita. Upute ocjenjivačima mogu pomoći da se smanji utjecaj ovih nebitnih čimbenika. Ocjenjivači moraju jasno shvatiti koje karakteristike pristupnika oni trebaju ocjenjivati, a koje trebaju zanemariti.

Treći zahtjev je da prosudbe moraju odražavati vještine pristupnika u vrijeme ispita. Ako se prosudbe temelje na ocjenjivačevom poznavanju znanja i vještina pristupnika, prosudba bi trebala biti obavljena što je bliže moguće vremenu ispita. Ako se prosudbe temelje na posebnom promatranju, izvedba koju ocjenjivači promatraju treba biti obavljena što je bliže moguće vremenu ispita.

Četvrti zahtjev je da prosudbe trebaju održavati pravo mišljenje ocjenjivača. Važno je provjeriti da ocjenjivači nemaju osobne razloge zbog kojih bi bili posebno strogi ili blagi kod prosuđivanja vještina pristupnika.

Svakako se preporuča da se ocjenjivači ne upoznaju s rezultatima ispita za pristupnike sve dok postupak prosudbe nije završen. Čak i ako se prosudbe temelje na izvedbi koja je i sama dio ispita, one trebaju biti prosudbe izvedbe, a ne ispitnih rezultata. Opasnost je u tome da ocjenjivač koji je upoznat s rezultatom ispita pristupnika može koristiti

rezultate prvih nekoliko pristupnika za uspostavljanje standarda i onda ocjenjivati ostale pristupnike uspoređivanjem njihovih rezultata na ispitu s onih prvih nekoliko. Ako prvih nekoliko pristupnika nisu tipični, to će imati utjecaja na sve preostale prosudbe. Međutim ako ocjenjivači nemaju pristup rezultatima ispita, oni će morati prosuđivati svakog pristupnika pojedinačno i postupak postavljanja standarda djelovat će kako bi i trebao.

#### *4.1. Metoda granične grupe*

Ova metoda temelji se na ideji da bi prag prolaznosti trebao biti rezultat koji bi se očekivao od pristupnika čije vještine granične. Metoda ima sličnosti s ostalim metodama koje se temelje na prosudbi ispitnih pitanja. Međutim umjesto traženja od ocjenjivača da naprave utemeljena predviđanja o načinu kako bi granični pristupnik trebao odgovarati, ova metoda poziva ocjenjivače da identificiraju stvarne pristupnike kao granične u znanju i vještinama koje ispit mjeri. Ocjenjivači ne moraju ocjenjivati sve pristupnike, niti njihov reprezentativni uzorak. Oni trebaju samo identificirati one pristupnike koji prema njihovoj prosudbi najbolje odgovaraju definiciji graničnog pristupnika. Zatim se prag prolaznosti postavlja na rezultat medijana (50%) ove granične grupe. Jednostavnost je glavna prednost ove metode, jer je jednostavna za korištenje i za objašnjavanje. Glavni nedostatak ove metode je da granični pristupnici obično čine mali postotak svih pristupnika. Ocjenjivači bi mogli imati problema u identifikaciji koji su pristupnici zbilja granični.

Razlozi korištenja medijana umjesto aritmetičke sredine (uobičajeni prosjek) je da je medijan puno manje podložan utjecaju nekoliko vrlo visokih ili vrlo niskih rezultata. Ova osobina medijana je posebno važna kod metode granične grupe, jer je pristupnik s vrlo visokom ili vrlo niskim rezultatom vjerojatno netko tko zaista ne spada u graničnu grupu. Ako je većina ispitnih rezultata granične grupe vrlo blizu jedan drugog, onda metoda dobro funkcionira. Međutim ako su rezultati granične grupe raspoređeni u širem rasponu mogućih rezultata, metoda ne funkcionira dobro.

#### *4.2. Metoda kontrastne grupe*

Metoda kontrastne grupe temelji se na ideji da se pristupnici mogu podijeliti u dvije kontrastne grupe – kvalificirana i nekvalificirana grupa – na temelju prosudbe njihovog znanja i vještina.

Većina pristupnika s vrlo visokim rezultatom bit će u kvalificiranoj grupi. Kako se ide prema dolje na ljestvici rezultata, broj pristupnika koji su kvalificirani smanjivat će se. Na razini najnižih rezultata nekvalificirani pristupnici brojčano će nadmašiti kvalificirane pristupnike. Jedan očigledan odabir praga prolaznosti bio bi rezultat pri kojem je isti broj kvalificiranih i nekvalificiranih pristupnika.

Pristupnici se mogu odabrati tako da su njihovi rezultati raspoređeni ravnomjerno kroz dijelove raspona rezultata gdje bi mogao biti prag prolaznosti. Na primjer, na ispitu koji ima 100 pitanja, može se odabrati 10 pristupnika iz svakog intervala od 5 bodova (31-35, 36-40 itd.). Važno načelo koje treba imati na umu je da uzorak pristupnika koji se odabire na svakoj razini rezultata reprezentativan za sve pristupnike na njihovoj razini rezultata.

#### 4.3. „Zaglađivanje“ podataka

Kod izračuna postotka kvalificiranih pristupnika na svakoj razini rezultata, može se dogoditi da postotak ne raste jednako od jedne do druge razine. Umjesto toga možda se kreće po cik-cak liniji. Ovakva vrsta rezultata je vjerojatna pogotovo kad je broj pristupnika kod svakog rezultata malen. Razumno je pretpostaviti da kad bi mogli dobiti prosudbe za sve moguće pristupnike, postotak kvalificiranih bi se znatno povećao od jedne razine rezultata do druge (moguće izravnane najviših i najnižih razina). U tom slučaju potrebno je uskladiti postotke i približiti ih onome što bi se dobilo u slučaju da postoje rezultati ispita za sve pristupnike.

Općeniti naziv za usklađivanje ove vrste je zaglađivanje. Postoji nekoliko tehnika zaglađivanja promatranih postotaka. Neke tehnike zaglađivanja uključuju komplicirane izračune, ali ima i drugih koje su vrlo jednostavne. Sve metode zaglađivanja temelje se na ideji da prosudbe pristupnika na svakoj razini rezultata govore nešto o znanju i vještinama pristupnika na najbližoj razini rezultata. Jedna metoda zaglađivanja koja je jednostavna za primjenu je nacrtati grafikon koji pokazuje postotke kao bodove. Zatim pokušajte nacrtati glatku krivulju koja se približava bodovima najbliže moguće. Ako broj pristupnika varira od jedne razine od druge, potrebno je približiti krivulju najbližim mogućim bodovima koji predstavljaju veći broj pristupnika. Ova tehnika zove se grafičko zaglađivanje i donekle je subjektivna, odnosno različite osobe u primjeni ove metode mogu doći do donekle različitih rezultata. Bez obzira na sve, metoda dobro funkcionira, odnosno proizvodi rezultate koji su vrlo slični rezultatima puno objektivnijih metoda zaglađivanja.

Druga jednostavna metoda zaglađivanja je zamjena promatranog postotka na svakoj razini rezultata s prosjekom postotaka za tu razinu rezultata i dvije najbliže razine rezultata. Na primjer, zaglađeni postotak – kvalificirani za ispit – razina rezultata 86 bio bi prosjek razina rezultata ispita od 85, 86 i 87. Poboljšanje kod ove metode je usporedba svakog postotka s brojem pristupnika na svakoj razini rezultata. Ovaj postupak ima učinak kombiniranja pristupnika na tri razine rezultata i izračunavanja postotka kvalificiranih za ovu povećanu grupu. Pomaknuti prosjek ne može se računati na najnižoj i najvišoj razini ispitnih rezultata, ali ovo ograničenje ne bi trebalo predstavljati ozbiljan problem u postavljanju graničnih vrijednosti.

#### 4.4. *Odabir praga prolaznosti*

Posljednji korak u primjeni metode kontrastne grupe je izbor praga prolaznosti. Jedan od logičnih odabira bio bi rezultat ispita kod kojeg zaglađeni postotak kvalificiranih iznosi točno 50%. Na bilo kojoj nižoj razini ispitnih rezultata, pristupnik će se prije smatrati nekvalificirani nego kvalificiranim, dok obrnuto važi na bilo kojoj višoj razini ispitnih rezultata.

Razlog postavljanja praga prolaznosti na rezultatu ispita koji odgovara 50%-noj vjerojatnost za prosudbu o kvalificiranosti temelji se na pretpostavci da su obje vrste pogrešnih odluka o pristupniku jednako ozbiljne. Ali što ako nisu? Na primjer, što ako je dva puta gore pustiti nekvalificiranog pristupnika nego što je srušiti kvalificiranog pristupnika? U ovom slučaju prag prolaznosti trebao bi biti viši, ali koliko viši? Teorija statističkog odlučivanja (koja u najjednostavnijem smislu zapravo predstavlja zdrav razum pretočen u matematički jezik) daje odgovor na ovo pitanje.

Odgovor se zasniva na ideji da bi odabir praga prolaznosti trebao ovisiti o ukupnoj šteti od svih pogrešnih odluka koje se očekuju. Ako je dva puta gore pustiti nekvalificiranog pristupnika nego što je srušiti kvalificiranog pristupnika, onda bi prolaz nekvalificiranog pristupnika bio isti kao da dva kvalificirana pristupnika ne prođu. Najbolji odabir praga prolaznosti bio bi rezultat ispita pri kojem postoje točno dva kvalificirana pristupnika za svakog nekvalificiranog pristupnika. To bi bio rezultat ispita koji odgovara 67% kvalificiranih. Na isti način, ako bi bilo tri puta gore pustiti nekvalificiranog pristupnika nego što je srušiti kvalificiranog pristupnika, prag prolaznosti bio bi rezultat ispita pri kojem kvalificirani pristupnici brojčano nadmašuju nekvalificirane u odnosu tri naprama jedan, tj. prag prolaznosti bio bi rezultat ispita koji odgovara 75% kvalificiranih.

Koji je broj pristupnika potreban prilikom provedbe metode?

Kod velikog uzorka bolja je zastupljenost populacije pristupnika pa je i veća preciznost u određivanju praga prolaznosti. Stupanj preciznosti koji se može dobiti s danim brojem pristupnika ovisi o nekoliko čimbenika:

- mjeri u kojoj i rezultati ispita i prosudbe odražavaju iste sposobnosti pristupnika;
- mjeri u kojoj su rezultati ispita i prosudbe nepodložni drugim utjecajima;
- konzistentnosti rezultata pristupnika;
- mjeri u kojoj ocjenjivači imaju iste standarde - ako različiti ocjenjivači prosuđuju različite pristupnike;
- konzistentnosti s kojom ocjenjivači primjenjuju svoje standarde u prosudbi pristupnika.

Stupanj preciznosti ovisi o broju osoba na koje će odabir praga prolaznosti utjecati i o posljedicama prolaza ili padanja ispita. Također će ovisiti o tome koliko je jasna razlika koja se namjerava napraviti. Izbor između pragova prolaznosti od 3 i 4 boda na ispitu koji ima 5 bodova puno je lakši nego odabrati između pragova prolaznosti 73 i 74 kod ispita koji ima 100 bodova.

Troškovi pribavljanja prosudbi individualnih pristupnika, preciznost koju osigurava broj pristupnika i potreba za preciznošću u postavljanju praga prolaznosti variraju od jedne ispitne situacije do druge. Općenito, može se savjetovati da se:

- Uključi što više moguće pristupnika testova;
- Konzultiraju statističari za savjet koji se odnosi na određenu ispitnu situaciju.

## **5. Metode temeljene na procjeni skupnog postignuća na ispitu**

Ove metode temelje se na prosudbama o grupi pristupnika – preferira se velika grupa. Ova grupa se često zove referentna grupa. Najjednostavnija od ovih metoda, i jedna s najočitijim opravdanjem, je odabir praga prolaznosti koji bi prošao određeni broj (ili određeni postotak) pristupnika u referentnoj grupi. Na primjer, ako postoje razlozi koji ukazuju da je 85% prošlogodišnjih pristupnika bilo kvalificirano, može se saznati rezultat koji bi omogućio da 85% prošlogodišnjih pristupnika prođe i koristiti ovaj rezultat kao prag prolaznosti za ovogodišnje pristupnike. Ako se ispit mijenja od godine do godine, morat će se pronaći rezultat na ovogodišnjem ispitu kojeg bi prošlo 85%

prošlogodišnjih pristupnika korištenjem statističke tehnike koja se zove poravnanje (Angoff WH, 1971). Prosudba o postotku pristupnika u referentnoj grupi koji su bili kvalificirani vodi izravno do odabira praga prolaznosti.

Da li prag prolaznosti odabran ovom metodom predstavlja apsolutni ili relativni standard? Odgovor na ovo pitanje ovisi o referentnoj grupi. Ako je referentna grupa upravo grupa pristupnika na koje će se primijeniti prag prolaznosti, onda je standard relativan. U tom slučaju pristupnikov relativni položaj u grupi određuje da li ili ne on/ona prolazi ispit. Međutim ako je referentna grupa prijašnja grupa pristupnika, to ima učinak postavljanja apsolutnog standarda. Iz kuta gledanja pristupnika, prag prolaznosti već je određen. Bilo koji pristupnik čiji će rezultat biti viši od tog rezultata prolazi ispit, čak i ako drugi pristupnici i sami imaju visok rezultat. Bilo koji pristupnik čiji rezultat će biti niži od tog rezultata pada ispit, bez obzira koliko su loši drugi bili.

## **6. Odabir jedinstvenih odrednica za postavljanje standarda praga prolaznosti**

Nakon uvida u navedene metode koje se koriste pri određivanju standarda za prag prolaznosti nameće se sljedeće pitanje: „Koja je metoda najbolja?“. Nema metode koja je najbolja za sve ispitne situacije. Odabir metode treba ovisiti o tome koju se vrsta prosudbe može dobiti s obzirom na svrhu primjene ispita i broj polaznika. Smatra se da je su najkorisniji podaci na temelju kojih se može izabrati metoda – rezultati ispita stvarnih pristupnika. Metoda kontrastne grupe najčešće se koristi kod ispita esejskoga tipa. Za ispite sa zadacima koji imaju više ponuđenih odgovora, preporuča se metoda kontrastne grupe samo kada postoji dokazana sigurnost da će ocjenjivači temeljiti svoje prosudbe na istim kvalitetama pristupnika, istom znanju i vještinama koje test mjeri. Metoda kontrastne grupe ima najjaču teoretsku podlogu od svih ostalih metoda – onu teorije statističke odluke. To je jedina metoda postavljanja standarda koja omogućava procjenu učestalosti dva tipa pogrešaka. Glavni nedostatak metode kontrastne grupe je poteškoća u pribavljanju potrebnih prosudbi.

Ako se ne mogu dobiti valjane prosudbe odgovarajućih uzoraka pristupnika, ali svaki ocjenjivač može sa sigurnošću identificirati individualne pristupnike kao dobre primjere za ljude s graničnim kvalifikacijama, preporuča se metoda granične grupe. Ako ocjenjivači mogu najbolje iskazati svoje standarde u smislu rezultata određene grupe pristupnika (na primjer, „dobar barem kao prosječni učenik s ocjenom 3“), preporuča se definiranje standarda u tom smislu.

Ako se nijedan od ovih uvjeta ne može zadovoljiti, predlaže se korištenje jedne od metoda koje se temelje na prosudbi ispitnih pitanja (Nedelsky, Angoff ili Ebel), ali također uz preporuku da se rezultati te metode usporede sa stvarnim ispitnim rezultatima. Pri ovoj metodi potrebna je spremnost na kompromis ako usporedba pokaže da su standardi ocjenjivača bili nerealistični.

Metode kao što su Nedelsky, Angoffova i Ebelova metoda posebno su korisne kad je važno da prag prolaznosti predstavlja standard za veliku i raznoliku grupu ljudi.

Angoffova metoda zahtijeva od ocjenjivača da razmišljaju o vjerojatnosti – što je teško za većinu ljudi – ili da zamisle grupu graničnih pristupnika, što može biti izvan uobičajenog iskustva ocjenjivača. Bez obzira na sve, Angoffova metoda je najlakša za objasniti i najbrža za korištenje od tri metode.

Ebelova metoda omogućuje ocjenjivačima da uzmu u obzir težinu i važnost svakog ispitnog pitanja. Ova karakteristika ima posebnu vrijednost kad se ispitna pitanja razlikuju značajno po svojoj važnosti. Njeni nedostaci su sporost i neprikladnost za kratke ispite.

Metoda Nedelsky uzima u obzir činjenicu da težina zadataka s više ponuđenih odgovora ovisi o tome koliko su odgovori netočni. Međutim metoda Nedelsky može biti teška za korištenje kod pitanja koja su negativno intonirana ili sadrže neke druge komplikacije.

## **7. Prijedlozi za određivanje praga na ispitima iz probne državne mature u Hrvatskim školama**

*Nacionalni centar za vanjsko vrednovanje obrazovanja* je školske godine 2007./2008. proveo nacionalne ispite na kojima su postignuti bodovi na svakom od provedenih ispita bili izraženi na relativnoj skali od 0 do 100, odnosno u postocima od maksimalnoga broja bodova. Granične rezultate tom su prilikom određivali članovi stručnih radnih skupina i Centar na osnovu distribucije rezultata, težine ispita, a sve u skladu s distribucijom ocjena koje su učenici postigli u školi. Ovakva metoda određivanja standarda može se usporediti s metodom odabira praga prolaznosti temeljene na odluci i definiranju standarda. Rezultat pojedinog pristupnika uspoređivao se prema standardu koji je ovisio o rezultatima drugih ispitanika na ispitu ali i na školskom uspjehu, tj. ovdje se radi o relativnom standardu.

Prosudbu o načinu određivanja standarda donijele su kompetentne osobe visoko kvalificirane za sadržaje ispita na koje se standard odnosio. Ocjenjivači su bili članovi stručnih radnih skupina koje su konstruirale ispite, kao i djelatnici Centra odgovorni za ovo područje rada.



Temeljem dostupnih podataka o važećim metodama određivanja praga prolaznosti i imajući u vidu probnu državnu maturu koja će se održati školske godine 2008./2009. predlažem primjenu metode temeljene na procjeni grupnog postignuća na ispitu. U tu svrhu potrebno je odrediti sljedeće:

1. Identificirati referentnu grupu ispitanika (slučajni uzorak)
2. Normalizirati distribuciju bruto rezultata za svaki ispit posebno
3. Pretvoriti bruto rezultate u skalu z-vrijednosti
4. Odrediti 5 razreda za svaki ispit posebno
5. Odabrati skupinu ocjenjivača za grupe predmeta (npr. jezična grupa, društvena, prirodoslovna, itd)
6. Ocjenjivači bi svakako trebali biti i članovi stručnih radnih skupina koji su dobro upoznati s predmetom, sadržajem i svrhom ispita
7. Ocjenjivače upoznati s metodom određivanja praga prolaznosti
8. Definirati adekvatne i neadekvatne razine znanja i vještina koje se ispituju
9. Prikupiti prosudbe o postotku polaznika koji u referentnoj grupi imaju određenu razinu znanja i vještina
10. Usporediti međusobno prosudbe ocjenjivača iz jedne skupine kako bi se odabrao prag prolaznosti za svaki ispit pojedinačno

Koraci 5. i 6. međusobno su ovisni, a Centar može jamčiti da će osigurati kompetentne ocjenjivače, članove stručnih radnih skupina, za svaki ispit. Referentna grupa treba biti prilično velika tako da prosudbe o postotku pristupnika koji su kvalificirani ne ovise puno o jednom ili dva pristupnika. Nije potrebno znati ispitne rezultate individualnih pristupnika, ali potrebno je znati koliko pristupnika u svakoj grupi ima isti ispitni rezultat.

Ocjenjivači moraju biti sposobni ocijeniti koliko je – ili koji postotak – pristupnika u referentnoj grupi kvalificirano u znanju i vještinama koje ispit mjeri. Stoga oni moraju znati koje su ispitne mjere i razina ovih vještina potrebni. Nije potrebno identificirati određene pojedince kao kvalificirane ili nekvalificirane, ali treba procijeniti koliko je otprilike pristupnika kvalificiranih, tj. trebaju se moći odrediti oni ispitanici koji su zadovoljili na ispitu. Ocjenjivači mogu donositi svoje prosudbe individualno ili kao grupa. Predlažem sljedeće:

- Neka svaki ocjenjivač napravi preliminarnu prosudbu,

- Ispisati prosudbe i dati ih svima istovremeno na uvid (projekcija na plohu),
- Raspraviti s ocjenjivačima koji su odabrali visoku prosudbu i tražiti da se to argumentira,
- Raspraviti s ocjenjivačima koji su odabrali nisku prosudbu i tražiti da se to argumentira.
- Dati ocjenjivačima priliku za promjenu prosudbe, ako to žele, a zatim prikupiti revidirane prosudbe.

Sve gore navedeno može se provesti na podacima s ispita održanih školske godine 2007./2008. prilikom uvježbavanja ocjenjivača.

Očekivane poteškoće prilikom primjene metode:

Očekuje se veliko raspršenje oko aritmetičke sredine jer će sudjelovati skupine učenika koje se značajno razlikuju po postignućima na dosad održanim nacionalnim ispitima. Naime, radi se o dvama velikim skupinama učenika; iz gimnazijskih i iz strukovnih programa. Također, to može utjecati na određivanje standarda prolaznosti na ispitu koji bi se nekim skupinama mogao činiti nepravedno niskim. Sve ovo može se provjeriti na simulaciji određivanja standarda praga prolaznosti za obavezne predmete na obje razine temeljene na rezultatima nacionalnih ispita školske godine 2007./2008.

Kod ispita kao što je državna matura važan je kontinuitet standarda. Polaganje državne mature preduvjet je za svjedodžbu, pa bi se značenje svjedodžbe moglo promijeniti ako se standardi izmijene. Međutim, ispiti se mijenjaju iz godine u godinu, i jedne godine može biti lakši, a druge teži. Jedan od načina za održavanje konstantnog standarda je prilagodba praga prolaznosti kako bi se uzela u obzir težina ispita. Takvo prilagođavanje se ipak može shvatiti kao promjena standarda, iako mu je cilj izbjeći promjenu standarda. Stoga prilagođavanje može uzrokovati razne probleme. Ipak, standard se može održati i prilagođavanjem ispitnih rezultata kako bi se kompenzirala težina ispita, a prag prolaznosti ostao nepromijenjen. Ovaj oblik prilagođavanja naziva se poravnanje. To je prihvatljiva i često korištena tehnika u edukacijskom testiranju, ali zahtijeva određene vrste informacija koje povezuju dva oblika ispita. Na primjer, dva oblika ispita mogu biti tako konstruirani da imaju nekoliko zajedničkih pitanja.

## 8. Literatura

1. Andrew, BJ i Hecht, JT "A Preliminary Investigation of Two Procedures for Setting Examination Standards." *Educational and Psychological Measurement*, 1976, v. 36, no. 1, pp. 45-50. (Report of a small-scale experiment comparing Nedelsky's method and Ebel's method.)
2. Angoff, WH. Scales, Norms, and Equivalent Scores. In R. L. Thorndike (ed.), *Educational Measurement*. Washington, D.C., American Council on Education, 1971, pp. 514-515. (Source document for Angoff's method.)
3. Berk, RA. "Determination of Optimal Cutting Scores in Criterion-Referenced Measurement." *Journal of Experimental Education*, 1976, v. 15, no. 4, pp. 4-9. (A method based on the comparison between instructed and uninstructed students.)
4. Brennan, RL i Lockwood, RE. "A Comparison of the Nedelsky and Angoff Cutting Score Procedures Using Generalizability Theory." *Applied Psychological Measurement*, 1980, v. 4, no. 2, pp. 219-240.
5. Ebel, RL. *Essentials of Educational Measurement*. Englewood Cliffs, NJ.: Prentice-Hall, 1972, pp. 492-494. (Source document for Ebel's method.) Hambleton, RK. "Test Score Validity and Standard-Setting Methods." In RA. Berk (ed.), *Criterion-Referenced Measurement: The State of the Art*. Baltimore: Johns Hopkins University Press, 1980, pp. 80-123.
6. Glass GV, Burton NW, Scriven M, Hambleton RK, Block JH, Popham WJ, Linn RL, Levin HM. *Journal of Educational Measurement*, 1978, v. 15, no. 4. Special issue on standard-setting. Contains articles by.
7. Hambleton RK. Setting standadrs on performance Assessments: Promising new methods and technival issues. Annual meeting of the APA, NY, 1995)
8. Huynh, H. "Statistical Consideration of Mastery Scores." *Psychometrika*, 1976, v. 41, no. 1, pp. 65-78. (Mathematical theory for the contrasting-groups method.)
9. Livingston, SA. "Choosing Minimum Passing Scores by Stochastic Approximation Techniques." *Educational and Psychological Measurement*, 1980, v. 40, no. 4, pp. 859-873. (Includes a detailed presentation of the up-and-down method.)