

## AUTENTIČNO OCJENJIVANJE, ŠTO JE TO?

### 1. UVOD

Osnovni čimbenik razvoja ljudske vrste je velika raznolikost ljudskog ponašanja. Ljudi su posebno osposobljeni za učenje i imaju pretpostavke za stjecanje velike količine znanja. Različita znanja nužna su za čovjekovu prilagodbu okolini, za selekciju iz okoline i mijenjanje okoline, kaže Sternberg (1997). Znanja su omogućila da ljudi opstaju u vrlo različitim uvjetima, na polovima i na ekvatoru, u podmornicama i svemirskim stanicama. Znanje će omogućiti, po svemu sudeći u ne tako dalekoj budućnosti, koloniziranje drugih planeta.

Zdravi ljudi uče cijeli život. Uče u roditeljskom domu, školuju se u institucijama, uče promatrajući tuđa ponašanja, uče iz priručnika, uče prateći medije. Posebno važna su znanja koja stječemo školovanjem. Ta bi znanja, za razliku od stečenih na druge načine, trebala biti **iskoristiva**. Beskorisna znanja, tipa znanja telefonskih brojeva koje nikad nećemo nazvati, često učimo slučajno, na primjer, putem agresivnih reklama. No, beskorisna znanja nisu dopuštena u školskom gradivu.

Mnoga znanja koja stječemo odmah su iskoristiva i njihov utilitet je lako razumljiv. Djeci koja su naučila plivati dopustit će se vožnja čamcem po jezeru, trk rivom ili igra uz rijeku. Naučili ste voziti bicikl i odsad možete brže prelaziti put nego pješice, bilo zbog potrebe ili zabave.

Mnoga znanja su nam nužna da bi na njih mogli nadograđivati druga. Prvo ćemo naučiti bacati i hvatati loptu, a tek onda igrati graničara, ili kasnije igrati rukomet. Moramo naučiti trčati, a tek onda ćemo naučiti baciti koplje, igrati nogomet, preskočiti letvicu.

Znanje nam omogućuje da razumijemo svijet u kojem živimo i da budemo aktivni dio toga svijeta. Znanje penjanja omogućilo je čovjeku iskustvo najvišeg vrha svijeta, znanje ronjenja omogućilo je pogled u dubine mora.

Znanje nam je potrebno da bi se aktualizirali kao osobe. Naše sposobnosti i druge osobine iskazat će se kroz naučene oblike ponašanja. Motorički sposobni pojedinci iskazat će se u motoričkim vještinama i pokazati čovječanstvu koliko se može brzo plivati ili klizati, daleko ili visoko skočiti, vješto prebaciti lopticu preko mreže, precizno ubaciti u koš slobodnim bacanjem ili vješto nogom izvesti jedanaesterac.

Znanje nam je potrebno da razumijemo i da uživamo u dometima darovitih pojedinaca. Sport čija pravila ne poznamo može nam izgledati kao besmisleno komešanje. Bez znanja i razumijevanja vrhunski dometi gimnastike, skokova u vodu, umjetničkog klizanja doživjet će se kao jednostavna zabava. Tek znanje o potrebnoj razini sposobnosti i količini i kvaliteti treninga potrebnog za određenu razinu izvedbe omogućuje pravo razumijevanje i uvažavanje motoričkog ponašanja.

Neka će znanja rijetko biti korištena, no mogu biti ključna za čovjekovu dobrobit. Znanje kako spretno pasti stečeno u djetinjstvu može dugo stajati neiskorišteno u memoriji. Aktivirat će se u rijetkim trenucima gubitka ravnoteže na skliskoj podlozi ili ljestvama i tada može biti presudno za očuvanje fizičkog integriteta.

Osnovni cilj školovanja je predaja znanja koja će biti iskoristiva za dobrobit pojedinaca. Tijekom školovanja potrebno je provjeravati ne samo postoji li određeno znanje u dugoročnoj memoriji, već i to jeli iskoristivo u realnim uvjetima života. Otuda potreba za razvojem novih sustava ocjenjivanja.

### **Definicije autentičnog ocjenjivanja**

Autentično ocjenjivanje je isprva zamišljeno kao poboljšanje procjene svladavanja intelektualnih zadataka pa Wiggins (1989, 1990) daje sljedeću definiciju:

«Ocjenjivanje je autentično u situacijama kada direktno ispituje učenikovu izvedbu pažnje vrijednih intelektualnih zadataka. Tradicionalno ocjenjivanje, nasuprot tome, bazirano je na indirektnim procjenama zadataka koji su pojednostavljene zamjene za pravu izvedbu.»

Definicija se proširila i danas uključuje sva znanja, ne samo simbolička nego i vještine: To je način ocjenjivanja u kojem se od učenika traži da izvode zadatke iz realnog života u kojima će pokazati smislenu primjenu osnovnih znanja i vještina (Mueller, 2003).

U ovu široku definiciju mogli bi svrstati ocjenjivanja koja su poznata pod drugim nazivima - direktno ocjenjivanje, ocjenjivanje izvedbe (performance assessment), ocjenjivanje uz pomoć dosijea .

Findak (2003) naglašava da je najveća vrijednost autentičnog ocjenjivanja u tome što omogućava i olakšava samopraćenje i samoocjenjivanje učenika. Kako navode Vizek-Vidović i suradnice (2003), mogućnost nadgledanja vlastitog procesa stjecanja znanja važan je faktor uspješnog učenja. Niz istraživanja koja to potvrđuju ilustriraju citatom Konfucija: "Reći da znaš kada znaš i reći da ne znaš kada ne znaš - to je znanje".

Findakova definicija glasi:

Autentično ocjenjivanje uključuje praćenje i vrednovanje kompletnog uratka učenika, što znači svih komponenti ocjenjivanja, koje se mogu i moraju kvantificirati, s ciljem da se ne samo nastavnici nego i učenici mogu na temelju definiranih elemenata praćenja i vrednovanja suditi o svom radu i napretku, kako u tekućoj školskoj godini, tako i tijekom čitavog školovanja.

### **Provedba autentičnog ocjenjivanja**

Koncept autentičnog ocjenjivanja višestruko je operacionaliziran i primijenjuje se za procjenu različitih znanja i vještina. Sve šira primjena može se očekivati i stoga što se pojavio komercijalni interes za izradu pomagala i edukaciju nastavnika za autentično ocjenjivanje. Na primjer, izdavači iz Pearson grupe razvijaju (i prodaju) udžbenike i video materijal za edukaciju nastavnika - ocjenjivača, te dosijee za autentično ocjenjivanje u različitim područjima.

Profesor Jonathan Mueller s North Central College-a (Naperville, Illinois) razvio je sustav koji bi mogao biti putokaz za razvoj autentičnog ocjenjivanja u svakoj novoj, konkretnoj situaciji. Mueller (2003) navodi da su nužna četiri koraka: (1) identifikacija standarda, (2) izbor autentičnog zadatka, (3) određivanja kriterija za zadatak i (4) kreiranje rubrika.

### **Identifikacija standarda**

Pod standardima Mueller podrazumijeva listu znanja i ponašanja koja se očekuju nakon uspješnog učenja. Standarde ćemo utvrditi odgovorom na pitanje što učenik treba znati ili što se mora učiniti nakon određene faze učenja. Mueller razlikuje tri tipa standarda: sadržajne, procesne i vrijednosne.

Sadržajni standardi su katalogi znanja i vještina koje učenik treba usvojiti. Na primjer, od studenata kineziologije se očekuje da će naučiti kako funkcionira lokomotorni sustav, tko je Franjo Bučar, što je aritmetička sredina, i očekuje se da zna preskočiti preponu, izvesti uspješni servis, gušiti protivnika, plivati leđno.

Procesni standardi su vještine koje omogućavaju efikasnije i kvalitetnije učenje. Posebno su vrijedni stoga što se odnose na vještine koje vrijede općenito i koje će olakšati svako novo učenje onom koji ih je usvojio. Za studente kineziologije vrlo važna takva vještina je postavljanje realističnih ciljeva u izvođenju motoričkih zadataka, a s obzirom na vlastite motoričke i morfološke karakteristike. Procesni standard koji je važan za sve učenike i studente je, na primjer, vještina pronalaženja relevantnih informacija.

Vrijednosni standardi se odnose na stavove koje želimo razviti kod učenika. U najmanju ruku očekujemo da će školovanjem pojedinci razviti pozitivan stav prema znanju. U nastavi tjelesne i zdravstvene kulture očekujemo da će učenici razviti pozitivan stav prema sportu i vježbanju te da će prihvatiti uvjerenje da je redovito vježbanje korisno i potrebno. Od studenata kineziologije očekujemo da će nakon četiri godine studija imati pozitivan stav i cijeniti sve sportove i sve oblike vježbanja. Očekujemo da će odbaciti stereotipe, na primjer spolne, i da pri kraju studija neće iskazivati uvjerenje da je aerobika za žene a nogomet za muškarce.

### **Izbor autentičnog zadatka**

Za određeni standard potrebno je razviti zadatke u kojima će studenti moći pokazati da su usvojili potrebna znanja ili vještine. Za razliku od uobičajenih testova znanja, autentični zadaci su oni gdje se (1) od učenika očekuje da sam uobliči svoj odgovor a ne da samo bira među ponuđenim, i (2) u kojima situacija nije artifičijelna i pojednostavljena već je zadatak blizak realnosti.

Mueller navodi učenje golfa kao primjer prirodnog izbora autentičnog zadatka. Pri provjeri znanja igranja golfa svom učeniku neće ponuditi test višestrukog izbora. Odvest će ga na golf teren i tražiti da igra golf. U terminima autentičnog ocjenjivanja, dovest će ga u situaciju realnog života da sam uobliči svoj odgovor. Motorička znanja su sama po sebi takva da se njihova provjera gotovo i ne može provesti izvan autentičnog zadatka. Teško je zamisliti kako indirektno mjeriti nečije znanje bacanja kugle ili plivanja prsnim stilom. Kao izuzetak

može se navesti padobranstvo, u kojem iz razumljivih razloga nije dozvoljena provjera znanja u realnim uvjetima.

Mueller (2003) navodi sljedeće razlike između tradicionalnih testova znanja i autentičnih zadataka:

<b>Karakteristike tradicionalnih testova znanja</b>	<b>Karakteristike autentičnih zadataka</b>
Izbor među ponuđenim odgovorima	Izvođenje zadatka
Artificijelni zadaci	Zadaci slični realnom životu
Procesi dosjećanja i prepoznavanja	Procesi konstrukcije i primjene
Strukturira ih nastavnik	Strukturira ih učenik
Indirektna procjena znanja	Direktna procjena znanja

### **Određivanje kriterija za zadatak**

Pokazatelje dobre izvedbe zadatka Mueller naziva kriterijima za zadatke. U ovom koraku potrebno je definirati indikatore koji će nam reći jeli učenik usvojio standard ili ne. Kriteriji će, naravno, ovisiti o zadatku. U igranju košarke jedan od kriterija će biti proporcija pogodaka u koš i promašaja, u umjetničkom klizanju proporcija ispravnih doskoka, u skoku s motkom visina letvice koju je sportaš preskočio bez rušenja.

Dobro određeni kriteriji imaju neke zajedničke karakteristike, bez obzira na vrstu zadatka. Dobar kriterij treba biti nedvosmisleno definiran. Treba sadržavati opis ponašanja, i to takav da je ponašanje moguće opažati. Treba biti opisan sažeto i jednostavnim jezikom razumljivim ne samo nastavniku nego i učeniku.

Broj kriterija treba prilagoditi zadatku. Kompleksni, opsežni i važni zadaci opisivat će se s više kriterija, dok će za jednostavnije i manje važne zadatke najčešće biti dostatan manji broj kriterija. Mueller navodi da broj kriterija ne bi trebao biti prevelik, bez obzira na vrstu zadatka, jer u protivnom ocjenjivanje postaje pretežak zadatak za ocjenjivača. Tvrdi da je za gotovo sve zadatke dovoljno do deset kriterija, a za mnoge da je dostatno i tri. Mueller preporučuje da u prvom pokušaju utvrđivanja kriterija pobrojimo više njih i da iz tog većeg skupa izdvojimo najvažnije, a odbacimo manje važne i one koji se sadržajem prekrivaju. Preporučuje da u određivanju kriterija konzultiramo kolege i učenike.

### **Kreiranje rubrika**

Rubrike su za Muellera skale na kojima će studentova izvedba zadatka biti konačno evaluirana. Rubrike se formiraju od najmanje dva kriterija. Za svaki kriterij treba odrediti

dvije ili više razina kvalitete izvedbe zadatka putem kojih će se moći razlikovati uspješni od manje uspješnih studenata. Kombinacijom kriterija i nivoa izvedbe dobivamo rubriku za pojedini zadatak.

Evaluiranje kvalitete izvedbe na svakom kriteriju Mueller naziva analitičkom rubrikom. Određivanje opće ocjene preko svih kriterija Mueller naziva holističkom rubrikom. Prilikom praćenja procesa učenja ocjenjivači će se najčešće poslužiti analitičkom rubrikom da bi mogli imati pod kontrolom sve relevantne kriterije uspješnog učenja. U slučaju kada treba rangirati učenike, za uvrštenje u školsku reprezentaciju, za dodjeljivanje priznanja ili za izdavanje preporuke za posao, prikladna će biti holistička rubrika.

Skale na kojima će se evaluirati razinu izvedbe zadatka trebaju biti pažljivo promišljene. Treba pomno odrediti što će se smatrati nezadovoljavajućim, što dobrim a što odličnim. Tamo gdje nije moguće objektivno odrediti vrijednost izvedbe, skalu treba odrediti tako da ocjenjivač čini što je moguće manju pogrešku prosudbe. Za kvalitetniju ocjenu, ponekad je potrebno angažirati i više ocjenjivača.

Mueller preporučuje da se rubrike konstruiraju iterativno, do verzije koja će zadovoljiti. Naglašava da je teško od prve izraditi dovoljno dobre rubrike. Preporučuje probnu primjenu prve varijante i korekcije na temelju tog iskustva, pa ponovnu probnu primjenu itd. do trenutka kad mijenjanje rubrika ne bi značilo bitno poboljšanje.

Najveća poteškoća nastavnicima u pokušajima uvođenja autentičnog ocjenjivanja često je baš izrada rubrika. Nastavnici su stručnjaci u svom području i imaju formirane sudove o tome što učenici trebaju znati, na kojim zadacima to mogu iskazati i kako prepoznati da li su dosegli standarde znanja. Konstrukcija skala za evaluaciju je psihometrijski problem za koji nastavnici u pravilu nisu educirani, pa im je pri izradi rubrika dobro došla pomoć.

### **Jeli autentično ocjenjivanje zaista novost?**

Odgovor je **ne**. Pojam **autentično ocjenjivanje** pojavio se krajem osamdesetih godina i postao popularan polovicom devedesetih godina prošlog stoljeća. No, postupci koji su opisani u autentičnom ocjenjivanju su stariji i primjenju se i bez korištenja dotičnog termina.

Pogledajmo primjer ocjenjivanja uspjeha u predmetu Tjelesna i zdravstvena kultura. Findak (1999., str. 243) navodi da treba ocijeniti je li svaki učenik ostvario utvrđeni cilj i zadaće programa ili nije, te jeli svaki učenik uspio svladati osnovni program ili ne. Prevedeno u pojmove autentičnog ocjenjivanja, Findak navodi da treba odrediti standarde i zatim utvrditi jesu li učenici zadovoljili standarde ili ne. U opisu komponenti koje trebaju činiti ocjenu Findak (1992., str. 179) navodi ne samo postignuti uspjeh u motoričkim znanjima već i rezultate u motoričkim dostignućima, razvoju sposobnosti, odgojnim efektima rada. U skladu s definicijama Muellera, Findak preporučuje sadržajne, procesne i vrijednosne standarde.

U načelima ocjenjivanja i sadržaju ocjenjivanja Findak (1992.) osuđuje donošenje ocjene temeljem izvođenja nekolicine pojedinačnih elemenata ili uspjeha u pokojoj disciplini (str. 172). Smatra potrebnim ocjenjivati sadržaje koji mogu pružiti informaciju o ukupnom dostignuću, a posebno davati prednost sadržajima koji imaju veću utilitarnu vrijednost,

drugim riječima, koji su primjenjivi u realnom životu (str. 175). Dakle, nema dvojbe, Findak se zalaže za autentične zadatke.

Opisujući metode provjeravanja Findak (1999) na primjeru vođenja lopte rukom daje ilustraciju kako opisati dobro ili loše obavljen zadatak. U autentičnom ocjenjivanju, pokazuje kako sastaviti kriterije (str. 212). Findak navodi sljedeće kriterije:

- pozicija lopte u odnosu na tijelo
- kontrola nad loptom
- potiskivanje lopte nasuprot udaranja (pljuskanja)
- meki nasuprot krutim pokretima ruke i šake
- smjer pogleda.

Findak (1999) uz pomoć navedenih kriterija izvodi pet-stupanjsku skalu:

	<b>pozicija lopte u odnosu na tijelo</b>	<b>kontrola nad loptom</b>	<b>potiskivanje lopte/ udaranja</b>	<b>meki / kruti pokreti ruke i šake</b>	<b>smjer pogleda</b>
<b>odličan</b>	pored i malo ispred tijela	potpuna	potiskuje loptu	meki pokreti	ispred sebe
<b>vrlo dobar</b>	pored i malo ispred tijela	gotovo potpuna	potiskuje loptu	uglavnom meki pokreti	prema lopti
<b>dobar</b>	pored i malo ispred tijela	uglavnom ima kontrolu	ponekad udara	kruti pokreti	prema lopti
<b>dovoljan</b>	ispred tijela	povremena kontrola	najčešće udara	kruti pokreti	prema lopti
<b>nedovoljan</b>	ispred tijela	nema kontrole	stalno udara	kruti pokreti	prema lopti

Treba svakako dodati da se Findak (1999) zalaže za redovito vođenje evidencija o radu te školske dokumentacije koja uključuje i osobni karton učenika (str. 275).

Autentično ocjenjivanje pretpostavlja informiranje učenika o standardima učenja i sudjelovanje učenika u oblikovanju autentičnih zadataka, kriterija i rubrika. Govoreći o metodama provjeravanja, Findak (1999) kaže: "Uopće nije i ne bi smjelo biti sporno aktivno sudjelovanje učenika u provedbi metode promatranja" (str. 213). Opisujući metodu mjerenja Findak (1999) preporučuje: "...u provedbu tog provjeravanja treba, kada god je to moguće, uključiti i učenike" (str. 214).

Rezimirajući sve navedeno, može se zaključiti da naša sredina ne samo da je spremna za autentično ocjenjivanje, već da i prakticira jedan oblik tog ocjenjivanja.

Dakle, autentično ocjenjivanje nije novo. Novo je to da veći broj stručnjaka pokušava razraditi sustav putem kojeg bi se nastavnike educiralo i putem kojeg bi se nastavnicima olakšalo ocjenjivanje, s ciljem objektivne, valjane i pouzdane ocjene.

### **Jeli autentično ocjenjivanje u konfrontaciji s tradicionalnim ocjenjivanjem?**

Odgovor je ponovo **ne**.

Znanje provjeravamo zbog više razloga (Vizek-Vidović i sur., 2003; Findak, 1999):

- zbog povratne informacije nužne učeniku i nastavniku za dalje oblikovanje procesa učenja,
- zbog izdavanja certifikata o educiranosti za određeni posao,
- zbog utvrđivanja stanja u populaciji za planiranje politike obrazovanja u zajednici (na primjer, provjera pismenosti populacije, provjera znanja plivanja u populaciji) itd.

Znanje će se mjeriti na način koji je najprikladniji za cilj mjerenja, pa će svoje mjesto naći različite metode ocjenjivanja.

### **Zaključak**

Kako bismo dobili objektivniju, valjaniju i pouzdaniju sliku znanja pojedinca i populacije trebalo bi koristiti sve načine provjere znanja. Kako bi se unaprijedilo ocjenjivanje trebalo bi poticati razvoj i upotrebu takozvanih alternativnih načina, među koja spada i autentično ocjenjivanje.

No, alternativni načini ocjenjivanja su zahtjevniji od tradicionalnih. Autentično ocjenjivanje traži veći napor i više vremena od nastavnika, a ponekad je potrebna ocjena od strane više ocjenjivača. Wiggins (1990) navodi da tradicionalno ocjenjivanje predstavlja minimalni, gotovo nikakav materijalni izdatak. Izračunao je da je autentično ocjenjivanje dvadeset puta skuplje. U zemljama u kojima se štedi na obrazovanju to bi mogli biti osnovni razlozi zbog kojih se sustavi ocjenjivanja ne unapređuju.

### **Literatura**

1. Findak, V. (1992). Metodika tjelesne i zdravstvene kulture. Zagreb: Školska knjiga.
2. Findak, V. (1999). Metodika tjelesne i zdravstvene kulture - priručnik za nastavnike tjelesne i zdravstvene kulture. Zagreb: Školska knjiga.
3. Findak, V. (2003). Konceptualna osnova praćenja i provjeravanja u tjelesnoj i zdravstvenoj kulturi. U zborniku "Assessment in Physical Education in Alpe Adriatic Countries. The National Education Institute Slovenija, Ljubljana, str. 25-32.
4. Mueller, J. (2003). Authentic Assessment Toolbox. <http://jonathan.mueller.faculty.noctrl.edu/toolbox/>
5. Sternberg, R.J. (1997). The Triarchic Theory of Intelligence. In: Flanagan, D.P., Genshaft, J.L. & Harrison, P.L. (Eds.). Contemporary Intellectual Assessment: *Theories, Tests and Issues*. New York: The Guilford Press.

6. Vizek-Vidović, V., Rijavec, M., Vlahović-Štetić, V., Miljković, D. (2003). Psihologija obrazovanja. Zagreb: IEP-VERN.
7. Wiggins, G. (1989). "A true Test: Toward More Authentic and Equitable Assessment", Phi Delta Kappan, 70, 9.
8. Wiggins, G. (1990). The Case for Authentic Assessment. ERIC Digest. ERIC Document Reproduction Service. <http://ericae.net/edo/ED328611.htm>