

UDK:371.3:004.738.5
Predhodno priopćenje
Primljeno: 20.05.2005.
Prihvaceno: 15.06.2005.

Vrednovanje ucinkovitosti procesa ucenja i poucavanja u sustavima za e-ucenje

Slavomir Stankov, Ani Grubišić, Branko Žitko, Divna Krpan
Fakultet prirodoslovno-matematičkih znanosti i odgojnih području Sveučilišta u Splitu
Teslina 12, Split
e-mail: slavomir.stankov{ani.grubisic,branko.zitko, divna.krpan}@pmfst.hr

Sažetak - *U radu smo se orijentirali na jedan aspekt e-ucenja, na ucenje i poucavanje uz pomoc Web-orijentiranih inteligentnih tutorskih sustava. B. Bloom je u istraživanju provedenom 1984. pokazao da se individualno i tradicionalno poucavanje u razredu razlikuju za dvije standardne devijacije, tj. da između njih postoji "2-sigma razlika" u korist individualnog poucavanja. Također, poucavanje u razredu se ne može natjecati s individualnim poucavanjem u interaktivnosti i prilagodljivosti. Uključivanjem racunala u obrazovanje i pojavom inteligentnih tutorskih sustava otvaraju se nove mogućnosti u ucenju i poucavanju jer su za razliku od ljudskih tutora, racunala ipak jeftinija. U trenutku kada inteligentni tutorski sustavi postaju Web-orijentirani, prelaze se prostorna i vremenska ograničenja. Naravno, potrebno je ispitati ucinkovitost tih sustava. U ovom radu je predstavljeno istraživanje ucinkovitosti Web-orijentiranog inteligentnog tutorskog sustava DTE_x-Sys, uz opis analize dobivenih rezultata i načina ispitivanja metrickih karakteristika testova.*

Ključne riječi: *sustavi za e-ucenje, inteligentni tutorski sustavi, Web-orijentirani inteligentni tutorski sustavi, ucenje i poucavanje, vrednovanje ucinkovitosti procesa ucenja i poucavanja.*

I. UVODNO RAZMATRANJE

Individualno ili tutorsko poučavanje je poučavanje tipa "jedan učitelj – jedan učenik", za razliku od tradicionalnog poučavanja u razredu gdje na jednog učitelja dolazi oko 20-30 učenika, iz čega se vidi razlika u vremenu koje učitelj može posvetiti svakom učeniku. Kod tradicionalnog predavanja u razredu svaki učenik se mora izboriti za svojih "5 minuta", pod uvjetom da to želi. Često se u velikim razredima događa da učenik pasivno prati nastavu, a učitelj to ne primjećuje. Naravno, nije dovoljno da učenik samo bude prisutan, te da uči tek kada se najavi ispitivanje; učenik mora aktivno sudjelovati kako bi znanje i vještine koje usvoji bili što kvalitetniji. U situaciji kada jedan učitelj poučava samo jednog učenika, on mora biti aktivan, ne može se "sakriti", može dobiti dodatno objašnjenje ako želi, a ne mora se dugo zadržavati na onome što je dobro usvojio.

Razmišljanje o razlikama između individualnog i tradicionalnog poučavanja navodi na zaključak da individualno poučavanje ima prednosti u odnosu na tradicionalno poučavanje. No, isticanje prednosti ili nedostataka nije dovoljno. Kako bismo se uvjerali u veću ili manju učinkovitost nekog načina poučavanja, treba provesti istraživanje o učinkovitosti toga načina poučavanja. Takvo istraživanje o razlikama između individualnog i tradicionalnog načina poučavanja je provodio Benjamin Bloom (Bloom, 1984.). Dobivena razlika od dvije standardne devijacije ili "2-sigma razlika", išla je u korist individualnog poučavanja (što se i očekivalo). To je približno ekvivalentno povećanju od 50-og do 98-og centila ili razine postignuća. Ako se uspjeh nekog učenika, primjerice, nalazi između 85-og i 86-og centila, to znači da je 85% učenika ostvarilo lošiji, a 14% bolji rezultat od njega. Ako neki učenik na testu ostvari uspjeh koji se nalazi iznad 99-og centila, to ne znači da je riješio 99% testa, nego je njegov uspjeh među 1% najboljih. Razlika od dvije standardne devijacije stoga znači da je svaki učenik koji je bio individualno poučavan, ostvario uspjeh koji pripada među 2% najboljih.

Zašto onda svi učenici ne uče individualnim načinom? Zato što je ekonomski nemoguće imati tutora za svakog učenika. Međutim, uključivanjem računala u obrazovanje i pojavom *inteligentnih tutorskih sustava* (eng. Intelligent Tutoring Systems, kratica ITS) otvaraju se nove mogućnosti u poučavanju. Posljednjih desetljeća su se

brzina racunala te njihovi kapaciteti za spremanje podataka udvostrucavali svakih 18-24 mjeseca (Phipps Merisotis, 1999.), a u isto se vrijeme smanjivala njihova cijena. U usporedbi s ljudskim tutorima racunala su jeftinija, pa bi se uz njihovu pomoc možda mogao popuniti dio razlike koja postoji izmedu individualnog i tradicionalnog poucavanja (Fletcher, 2003.).

E-ucenje (eng. e-learning) nova je paradigma ucenja uz pomoc razlicitih uredaja koji se temelje na elektronicnoj tehnologiji, a odnosi se na dostavljanje sadržaja ucenja ucenicima preko svih vrsta elektronicnih medija, ukljucujuci racunala, Internet, satelitsko emitiranje, audio- i videovrpce, interaktivnu televiziju, CD-ROM-ove ili DVD-ove i sl. (definicija preuzeta sa <http://www.learnframe.com/aboutelearning>). U ovom cemo se radu orijentirati samo na jedan aspekt e-ucenja, na Web-orijentirano ucenje i poucavanje uz pomoc Web-orijentiranih inteligentnih tutorskih sustava. Web-orijentirani inteligentni tutorski sustavi prelaze ogranicenja prostora i vremena, pa tako postaju sve zanimljiviji. U ovom radu predstavljamo istraživanje o ucinkovitosti Web-orijentiranog ITS-a DTEEx-Sys (Distributed Tutor Expert System) (Rosic, 2000.).

U drugom poglavlju je ukratko opisana struktura sustava DTEEx-Sys na kojem je provedeno istraživanje. U trecem poglavlju je opisano naše istraživanje ukljucujuci planiranje i realizaciju istraživanja, te analizu dobivenih rezultata i ispitivanje mjernih instrumenata.

II. STRUKTURA SUSTAVA DTEEX-SYS

Sustav DTEEx-Sys ne ovisi o platformi, vremenu i mjestu pristupa, ima jedinstven nacin spremanja informacija za razlicita podrucna znanja, mogucnost ucenja u hipermedijskoj okolini, mogucnost pracenja napretka ucenika, te mogucnost testiranja uz preporuke za nastavak. Sustav je zasnovan na oponašanju ljudskih ucitelja i omogucuje svakom uceniku ucenje u bilo kojem trenutku, onoliko koliko mu je potrebno kako bi stekao željenu razinu znanja.

Znanje u sustavu DTEEx-Sys je predstavljeno semantickim mrežama s okvirima ciji su osnovni elementi cvorovi i veze. Cvorovi se koriste za predstavljanje objekata podrucnog znanja, a veze prikazuju odnose medu njima. U sustavu DTEEx-Sys se baze podrucnog

znanja, najjednostavnije receno, sastoje od cvorova i veza. Cvorovi u bazi podrucnog znanja pored naziva i veza prema ostalim cvorovima mogu imati i strukturne atribute: tekstualni opis, sliku i animaciju, te URL adresu.

Korisnici sustava mogu odabrati sljedece usluge: pristup bazama podrucnog znanja, testiranje znanja ili zatražiti rezultate testova. Specificnost modula testiranja je u tome što se nakon svakog ciklusa od dva pitanja generiraju nova pitanja cija težina ovisi o rezultatima ucenika, te se na kraju uz ocjenu daju i preporuke za dalji rad.

DTEEx-Sys je implementiran kao 3-slojna klijent-server arhitektura (Rosic i sur., 2001.) koja se sastoji:

- 1) od *korisnickog sucelja* – sloj korisnickog sucelja sustava DTEEx-Sys pruža sucelje za pristup sustavu, bazama znanja, rješavanju testa i pregledu ocjena koje ucenici postižu;
- 2) od *logike aplikacije* – srednji sloj generira dokument koji se distribuira korisniku prema njegovu zahtjevu, ovisno što korisnik želi raditi (npr. uciti, rješavati test);
- 3) od *baze podataka* – sloj baza podataka se sastoji od baza podrucnog znanja i baza podataka koje sadrže podatke o korisnicima sustava. Baze podataka o korisnicima sustava sadrže elementarne podatke o korisnicima, njihovu ponašanje pri korištenju sustava i rezultatima njihova testiranja.

III. PRISTUP VLASTITOM ISTRAŽIVANJU

U ovom radu se predstavlja istraživanje o djelovanju sustava DTEEx-Sys na proces ucenja i poučavanja, kako bi se pokušalo odgovoriti na pitanja: Koliko DTEEx-Sys utjece na ucenje i poučavanje i kolika je ucinkovitost DTEEx-Sys-a? Odabirom metode eksperimentalnog istraživanja (Iqbal i sur., 1999.) pokušali smo ustvrditi hoće li biti poboljšanja u postignucima studenata u odnosu na tradicionalno predavanje. Eksperimentalno istraživanje je uobicajena metoda vrednovanja postignuca u psihologiji i obrazovanju, a prema Marku i Greeru (1993.) može se primijeniti i na inteligentne tutorske sustave.

A. Planiranje i realizacija eksperimenta

Cilj istraživanja je ispitati je li DTE_x-Sys utjecao na poučavanje studenata u određenom područnom znanju i koliki je taj utjecaj, izraženo u standardnoj mjeri *velicine ucinka* (eng. effect size). Velicina ucinka pokazuje je li *poučavanje na sustavu DTE_x-Sys* učinkovitije od *tradicionalnog poučavanja* u procesu *ucenja* i *poučavanja*.

Za uzorak na kojem će se izvoditi eksperiment je odabrana skupina od ukupno 31 studenta druge godine, Fakulteta prirodoslovno-matematičkih znanosti i odgojnih područja Sveučilišta u Splitu, triju studijskih grupa: matematike i informatike, matematike, te informatike i tehničke kulture, a svi skupa slušaju kolegij *Racunalni praktikum I*. Područno znanje na kojem su studenti testirani je objektno-orijentirana paradigma programiranja. Studenti su podijeljeni na eksperimentalnu (20 studenata) i kontrolnu skupinu (11 studenata) slučajnim odabirom. (Ispitana je situacija s dvije eksperimentalne skupine, ali su dobiveni isti rezultati), a zatim su svi zajedno slušali predavanje o temeljnim pojmovima objektno-orijentirane paradigme u trajanju od dva školska sata. Studentima je omogućen i pristup sadržaju predavanja u tekstualnom obliku. Svi studenti su pisali pismeni **inicijalni test** koji se sastojao od 20 pitanja iz područja objektno-orijentiranog programiranja, koji je ocjenjivan bodovima od 0 do 100, da bi se vidjelo ima li kakvih razlika u predznanju između kontrolne i eksperimentalne skupine.

Eksperimentalnoj skupini je slijedio jedan tjedan ucenja i poučavanja o objektno-orijentiranom programiranju na sustavu DTE_x-Sys. Baza područnog znanja iz objektno-orijentiranog programiranja znanja se sastoji od 202 cvora i 189 veza, a gotovo svi cvorovi imaju objašnjenje u obliku hiperteksta. Neki od cvorova imaju i dodatna objašnjenja u obliku slika ili prezentacija (ukupno 45 prezentacija i 45 slika). Testiranje u tom razdoblju nije bilo moguće. Kontrolna skupina nije imala pristup, te je ona nastavila učiti iz ranije dobivenih tekstualnih materijala.

Na kraju toga tjedna studenti eksperimentalne skupine su se testirali na sustavu DTE_x-Sys, a zatim su nakon toga zajedno sa studentima kontrolne skupine pisali **završni test** koji je imao 16 pitanja (13 teorijskih pitanja i 3 zadatka), kako bi se vidjelo ima li značajne razlike između eksperimentalne i kontrolne skupine nakon korištenja sustava. Test je također ocjenjivan bodovima od 0 do 100.

B. Analiza rezultata

Analiza rezultata zapocinje s utvrđivanjem da li postoji statisticki znacajna razlika izmedu kontrolne i eksperimentalne skupine u pocetnim uvjetima, tj. u rezultatima iz inicijalnog ispitivanja, a zatim i u rezultatima završnog ispitivanja. Termin *statisticka znacajnost* (eng. statistical significance) važan je i cest termin u psihologijskim istraživanjima. U vezi s tim, prvi korak je postavljanje nul-hipoteza za rezultate iz inicijalnog ispitivanja H1: *Nema znacajnih razlika izmedu eksperimentalne i kontrolne skupine u rezultatima inicijalnog ispitivanja*, te za završno ispitivanje H2: *Nema znacajnih razlika izmedu eksperimentalne i kontrolne skupine u rezultatima završnog ispitivanja*.

Kako bi se odabrao odgovarajuci test, potrebno je razmisliti o kakvu se tipu podataka radi, tj. jesu li podaci iz nekog intervala (eng. interval data) ili se odnose na poredak (eng. ordinal data) (Becker, 1999.). U našem istraživanju se radi o podacima iz intervala 0-100 koji se odnose na dvije nezavisne skupine, kontrolnu i eksperimentalnu, pa je odabran i nezavisni t-test, koji cemo dalje u tekstu zvati t-test.

T-test se može izracunati primjenom nekog statistickog racunalnog programa. Ovdje je korišten program StatSoft Inc. STATISTICA 6 (StatSoft, Inc, 2004.). Kao što se vidi iz tablice 1, vrijednost $p=0.74$ je veca od 0.05 pa se hipoteza H1 prihvaca i zakljucuje da nema znacajnih razlika izmedu eksperimentalne i kontrolne skupine u pocetnim uvjetima.

Tablica 1: Rezultati t-testa za inicijalno ispitivanje

	Vrijednosti t-testa	Znacajna razlika
Eksperimentalna vs. Kontrolna	$t = 0.34$ $p = 0.74$	Ne

Rezultati u tablici 2 pokazuju da je $p>0.05$ pa se prihvaca nul-hipoteza H2 i zakljucuje da nema statisticki znacajnih razlika izmedu kontrolne i eksperimentalne skupine u rezultatima završnog ispitivanja.

Tablica 2: Rezultati t-testa za završno ispitivanje i postignuce

	t-test (završno ispitivanje)	t-test (postignuce)	Znacajna razlika
Eksperimentalna vs. Kontrolna	t = 0.58 p = 0.56	t = 0.29 p = 0.77	Ne

Dakle, prema rezultatima t-testa možemo zaključiti da tretman eksperimentalne skupine, tj. učenje na sustavu DTEEx-Sys nije imalo značajnog utjecaja na njihov uspjeh. Bilo kakva razlika između eksperimentalne i kontrolne skupine je mogla nastati kao posljedica greške.

Međutim, test kojim se utvrđuje značajnost razlika između dviju skupina nije potpun bez provjere veličine učinka jer ne možemo vidjeti snagu veze između tretmana i postignuca u populaciji iz koje smo odabrali uzorak ispitanika. Iz toga se razloga koristi veličina učinka.

Velicina učinka je općenita mjera veličine utjecaja nekog novog načina poučavanja u eksperimentalnoj skupini, u odnosu na kontrolnu koja je koristila tradicionalni način poučavanja. Postoji više načina za računanje veličine učinka, a najčešće se koristi formula (1) (Kulik & Kulik, 1989.; prema Yaakub, 1998.):

$$ES = \frac{AS_E - AS_K}{SD_K} \quad (1)$$

AS_E = Aritmetička sredina postignuca eksperimentalne skupine

AS_K = Aritmetička sredina postignuca kontrolne skupine

SD_K = Standardna devijacija kontrolne skupine

Postignuce (eng. gain) izračunano je na način da se od rezultata koje su studenti postigli na završnom ispitivanju oduzmu rezultati iz inicijalnog ispitivanja.

Standardna devijacija se računa formulom:

$$s = SD = \sqrt{\sum_i \frac{(X_i - M)^2}{N - 1}} \quad (2)$$

X_i = izmjerena vrijednost ili rezultat

M = aritmeticka sredina rezultata

N = broj rezultata

U tablici 3 su prikazane vrijednosti aritmetickih sredina i standardnih devijacija za testove i postignuce. Vrijednosti aritmetickih sredina i standardnih devijacija se koriste za izracunavanje velicine ucinka.

Tablica 3: Aritmeticke sredine i standardne devijacije

	Eksperimentalna	Kontrolna
Inicijalno ispitivanje	AS= 59.35 SD= 26.00	AS= 56.18 SD= 22.61
Završno ispitivanje	AS= 46.05 SD= 23.89	AS= 41.00 SD= 21.61
Postignuce (eng. gain)	AS= -13.3 SD= 17.93	AS= -15.18 SD= 15.12

Izracunali smo velicinu ucinka (ES) po formuli (1) za naše istraživanje:

$$ES = \frac{-13.3 - (-15.18)}{15.12} = \mathbf{0.12} \quad (3)$$

Dakle, velicina ucinka sustava DTEEx-Sys, koja je dobivena istraživanjem je 0.12. U konacnici Pedhazur i Schmelkin (Petz, 2004.) velicinu ucinka manju od 0.2 smatraju *malom velicinom*, velicinu ucinka jednaku 0.5 *srednjom*, a velicina ucinka vecu ili jednaku 0.8 *velikom velicinom* ucinka. Naše je istraživanje rezultiralo malom velicinom ucinka, a kako smo ocekivali znacajnije rezultate, bilo je potrebno ispitati što je moglo utjecati na istraživanje. U tu svrhu proveli smo provjeru mjernih instrumenata, tj. testova koji su se koristili za ispitivanje znanja studenata.

C. Metricke karakteristike testova

Kod ucenja i poučavanja najčešće se ispituju stavovi i postignuca ucenika (Phipps i Merisotis, 1999.). Obično se za mjerenje *stavova* (eng. attitudes) koriste upitnici, a za mjerenje postignuca testovi. U našem istraživanju je bilo potrebno izmjeriti postignuce ucenika, pa smo odabrali testove kao instrumente za mjerenje. Testovi bi trebali točno mjeriti postignuce ucenika, kako bi istraživanje imalo smisla. Dakle, potrebno je ispitati metricke karakteristike testova. Mužic (Mužic, 1979.) definira sljedece metricke karakteristike testova: (i) valjanost, (ii) pouzdanost, (iii) objektivnost, (iv) osjetljivost, (v) diskriminativna vrijednost zadataka.

Kod našeg istraživanja nije ispitana objektivnost testova iz praktičnih razloga. Ovi su se testovi, naime, trebali upotrijebiti samo jedanput te se ne bi isplatilo okupiti veći broj ocjenjivaca koji bi morali poznavati specifično područno znanje i biti upoznati s istraživanjem. Ta je karakteristika primjerenija ako se testovi planiraju koristiti i za neku buduću namjenu. Također, nije ispitana ni osjetljivost jer osjetljivost nema smisla ispitivati ako su pouzdanost i objektivnost testa niske ili nisu ispitivane (Mužic, 1979.).

1) Valjanost

Valjanost je karakteristika testa koja nam pokazuje da li i koliko test mjeri ono što njime želimo mjeriti. Ne postoji "opća" valjanost testa, tj. valjanost se može odrediti samo u vezi s konkretnom namjenom za koju se test koristi. Jedan od kriterija za ispitivanje valjanosti je nastavni program. To znači da bi se sadržaj i ciljevi nastave trebali slagati sa sadržajem testa. Ako se testom točno ispituje ono što bi ucenik trebao naučiti, onda je test valjan. Uobičajen postupak je određivanje ciljeva nastave koji se zatim uspoređuju sa sadržajem testa. Većinom se testom ne ispituje cijeli nastavni sadržaj, nego samo jedan dio koji na što bolji način treba predstavljati cjelinu.

Prilikom provedbe našeg istraživanja nastojalo se da se sadržaji testova podudaraju sa sadržajem predavanja o objektno-orijentiranoj paradigmi koje je trajalo dva školska sata. Prema tome, sadržaji testova i predavanja se podudaraju.

2) Pouzdanost

Ova karakteristika određuje možemo li se i u kojoj mjeri osloniti na rezultat koji se dobije testom, tj. da li test točno mjeri, bez obzira što se mjeri. U slučaju negativnih rezultata, tj. kad bi se pokazalo da testovi na promatranom uzorku nisu pouzdani, onda se ne bi uopće mogli osloniti na rezultate testova. Za pouzdanost je važna jednoznacnost pitanja, jer učenik koji zna odgovor, može odgovoriti netočno ako je pogrešno razumio pitanje. Ako postoje zadaci s dvoclanim izborom (npr. točno/netočno ili istina/laž), onda postoji i veća mogućnost da učenik slučajno pogodi odgovor što nema veze s njegovim znanjem, a smanjuje pouzdanost testa. Faktor koji se povezuje s pouzdanosti je duljina testa, tj. broj zadataka ili pitanja koje test sadrži. Povećanjem broja zadataka povećava se i pouzdanost testa, pod uvjetom da su novi zadaci iste kvalitete kao postojeći. Pouzdanost se može odrediti na nekoliko načina, a spomenut ćemo dva kojima je ispitana pouzdanost testova iz našeg istraživanja.

Prvi način određivanja pouzdanosti je *metodom izračunavanja koeficijenta homogenosti*. U našem istraživanju za svakog studenta posebno smo zbrojili bodove ostvarene na parnim zadacima (prva varijabla), te bodove ostvarene na neparnim zadacima (druga varijabla). Izračunavanjem korelacije između parnih i neparnih zadataka dobije se pokazatelj pouzdanosti. Koristeći program STATISTICA 6 dobiveni su koeficijenti korelacije pola testa za inicijalno ispitivanje 0.747 i za završno ispitivanje 0.753. Kako se radi o pola testa, potrebno je primijeniti Spearman - Brownovu formulu (Siegle, 2004.):

$$r_{cijeli} = \frac{2 \cdot r_{pola}}{1 + r_{pola}} \quad (4)$$

r_{cijeli} = koeficijent korelacije cijelog testa

r_{pola} = koeficijent korelacije pola testa

Dakle, iz formule (4) slijedi da je pouzdanost za inicijalno ispitivanje i završno ispitivanje približno jednaka i iznosi 0.85, što je, prema Mužicu (Mužic, 1979.), dovoljno visok koeficijent.

Drugi način određivanja pouzdanosti je pomoću *Cronbachova α koeficijenta* (eng. Cronbach's Coefficient Alpha) (Hempel, 2003.). Taj se način često koristi za procjenu pouzdanosti jer se njime može ocjenjivati pouzdanost testova čiji se zadaci ocjenjuju samo s 0 ili 1 bod (odnosno točno/netočno), te testovi čiji se zadaci ocjenjuju s više od jednog boda (Siegle, 2003.). Koeficijent α daje najnižu procjenu pouzdanosti koja se može očekivati. Ako je koeficijent visok, znači da je pouzdanost visoka. Međutim, ako je koeficijent nizak, ne može se ništa zaključiti o pouzdanosti, te je treba ispitati na neki drugi način. Cronbachov koeficijent se računa kao aritmetička sredina svih mogućih podjela testa.

Test je pouzdan ako je $\alpha > 0.70$ (Wisher i Olson, 2003.), ali preveliko približavanje 1 može značiti redundanciju, tj. nepotrebno ponavljanje istoga kroz različite zadatke. Pri izračunavanju Cronbachova α koeficijenta za naše testove koristili smo program STATISTICA 6 i dobili $\alpha > 0.70$.

Dakle, primjenom dva načina ispitivanja pouzdanosti dobili smo pozitivne rezultate, tj. koeficijent homogenosti > 0.80 i $\alpha > 0.70$, što znači da se možemo osloniti na rezultate testova.

3) *Diskriminativna vrijednost zadataka*

Pojedini zadaci u testu su više valjani što se više slaže uspjeh učenika na tom zadatku s uspjehom na citavom testu. Prema tome, valjani zadaci omogućuju razlikovanje uspješnih i neuspješnih učenika. Ako uspješni učenici dobro rješavaju neki zadatak, a neuspješni slabo, onda će stupanj slaganja između uspjeha u tom zadatku i ukupnog uspjeha u testu biti visok, a diskriminativna vrijednost zadatka visoka i pozitivna. Može se dogoditi i da učenici sa slabijim uspjehom neki zadatak rješavaju, a uspješniji učenici ne. U tom slučaju je diskriminativna vrijednost zadatka negativna, a takav zadatak ne bi smio biti dio testa.

Diskriminativna vrijednost zadataka se određuje računanjem "point-biserijalnog" koeficijenta korelacije koji se obično označava s r_b (Petz, 2004.). Obično se provodi sljedeći postupak. Prvo se "numerira" dihotomna varijabla, tj. jedna karakteristika se

oznaci s jednim brojem, a druga s drugim, npr. 0 i 1, te se tako dobije varijabla koja sadrži samo dvije različite vrijednosti.

Sljedeći korak je računanje Pearsonova koeficijenta korelacije. Taj koeficijent korelacije je izračunan na cijelom uzorku. Dobiveni su niski koeficijenti korelacije za neke od zadataka (zadatak 3. ($r_{pb}=0.19$) i zadatak 10. ($r_{pb}=0.01$) u inicijalnom ispitivanju i zadatak 2. ($r_{pb}=0.12$) u završnom ispitivanju). Analizom rezultata bez tih zadataka dobije se velicina ucinka od 0.18 standardnih devijacija, što je neznatno povećanje u odnosu na rizik gubitka pouzdanosti i osjetljivosti ako bi se ti zadaci izbacili iz testa.

IV. ZAKLJUCAK

Uz Web-orijentirane inteligentne tutorske sustave učenici se više ne moraju okupljati u isto vrijeme na određenom mjestu radi učenja i poučavanja. Ti sustavi obećavaju dostupnost kvalitetnog obrazovanja svim osobama kojima je ono potrebno, pod uvjetom da imaju odgovarajuću tehnologiju za pristup. Ta tehnologija postaje sve bolja i jeftinija, a na taj način i lakše dostupna.

Iako se čini da Web-orijentirani ITS-ovi nude mnoge prednosti, potrebno je provesti istraživanje kojim će se ispitati njihova učinkovitost. Svako provedeno istraživanje pridonosi općenitoj ocjeni učinkovitosti Web-orijentiranih ITS-ova.

Istraživanje je potrebno planirati. Kada se odredi što će se konkretno istraživati, potrebno je odabrati uzorak ispitanika. Važan korak je priprema testova i provjera njihovih metrickih karakteristika, zato što o njima ovisi ispravnost rezultata istraživanja. Dobiveni rezultati istraživanja se analiziraju odabranim statističkim metodama.

Prilikom provedbe našeg istraživanja sustava DTEEx-Sys nismo dobili visoku veličinu ucinka, ali smo stekli dragocjeno iskustvo koje nam može pomoći pri budućim istraživanjima. Smatramo da je nemotiviranost studenata koji su sudjelovali u istraživanju mogla imati negativnog utjecaja na provedbu i rezultate istraživanja. Potrebno je prilikom analize istraživanja uzeti u obzir i koliko vremena su studenti učili on-line na sustavu

DTEx-Sys, te ispitati imaju li svi odgovarajuću tehnologiju za pristup (pristup internetu, računalo i odgovarajuću programsku podršku).

LITERATURA

1. Bloom B. S. (1984). The Two-Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*. 13,4-16.
2. Phipps, R. & Merisotis, J. (1999). What's the Difference? - A Review of Contemporary Research on the Effectiveness of Distance Learning in Higher Education, The Institute for Higher Education Policy, Washington, (URL <http://www.ihep.com/Pubs/PDF/Difference.pdf>).
3. Fletcher, J. D. (2003). Evidence for Learning From Technology-Assisted Instruction, in H. F. O'Neal, R. S. Perez, *Technology applications in education: a learning view*, Mahwah, NJ: Lawrence Erlbaum Associates, str. 79-99.
4. Rosic, M. (2000). Zasnivanje sustava obrazovanja na daljinu unutar informacijske infrastrukture, magistarski rad, Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu, Zagreb.
5. Rosic, M., Stankov, S., Glavinic, V. (2001). DTEx-Sys – A Web Oriented Intelligent Tutoring System, in *Proceedings of Intelligent Conference On Trends in Communication – EUROCON 2001*, Vol 2/2, str. 255-258.
6. Iqbal, A., Oppermann, R., Patel, A. and Kinshuk (1999). A Classification of Evaluation Methods for Intelligent Tutoring Systems", *Software Ergonomie '99 – Design von Informationswelten* (Eds. U. Arend, E. Eberleh & K.Pitschke), B. G. Teubner Stuttgart, Leipzig, str.169-181.
7. Mark, M. A. & Greer, J. (1993). Evaluation Methodologies for Intelligent Tutoring Systems, *Journal of Artificial Intelligence and education*, 4 (2/3), str. 129-153.
8. Becker, L. A. (1999). Testing for Differences Between Two Groups: t test, (URL <http://web.uccs.edu/lbecker/SPSS/ttest.htm>).
9. StatSoft, Inc. (2004). Electronic Statistics Textbook, (URL <http://www.statsoft.com/textbook/stathome.html>).

10. Yaakub, M. N. (1998). Meta-Analysis of the Effectiveness of Computer-Assisted Instruction in Technical Education and Training, doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.
11. Petz, B. (2004). Osnovne statističke metode za nematematičare, Naklada Slap, Jastrebarsko.
12. Mužic, V. (1979). Metodologija pedagoškog istraživanja, Svjetlost, Sarajevo.
13. Siegle, D. (2003). Reliability, Neag School of Education - University of Connecticut, (*URL <http://www.gifted.uconn.edu/siegle/research/InstrumentReliabilityandValidity/Reliability.htm>*)
14. Hempel, S. (2003). Reliability, University of Derby, (*URL <http://ibs.derby.ac.uk/~susanne/PTT/lectures/PTTrelability2003.pdf>*).
15. Wisner, R.A. & Olson, T. M. (2003). The Effectiveness of Web-Based Instruction, U.S. Army Research Institute for the Behavioral and Social Sciences, Virginia, Research Report 1802.

Evaluating Influence of Learning and Teaching Process in e-learning systems

Abstract – In this paper we have focused on one aspect of e-learning, on learning and teaching using Web-based intelligent tutoring systems. B. Bloom in his research in 1984. had shown that individualized and traditional instruction in classroom differ in two standard deviation, that is, there is "2-sigma difference" in advantage of individualized instruction. Also, traditional instruction in classroom cannot compete individualized instruction in interactivity and adoptability. Involving computers in education and development of intelligent tutoring systems give new potential to learning and teaching process because computers are much cheaper than human tutors. In a moment when intelligent tutoring systems become Web-oriented, time and space boundaries are completely crossed. Evidently, educational influence of those systems should be evaluated. In this paper we present a research that was conducted to evaluate educational influence of Web-based intelligent tutoring system DTE_x-Sys, along with description of result analyses and metric characteristics of used tests.

Keywords: E-learning systems, intelligent tutoring systems, Web-based intelligent tutoring systems, learning and teaching, evaluating influence of learning and teaching process.