

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Ani Grubišić

**VREDNOVANJE UČINKA INTELIGENTNIH
SUSTAVA E-UČENJA**

MAGISTARSKI RAD

Zagreb, 2007.

UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Ani Grubišić

**EVALUATING EDUCATIONAL INFLUENCE OF
INTELLIGENT E-LEARNING SYSTEMS**

MASTER THESIS

Zagreb, 2007.

Magistarski rad je izrađen na Zavodu za elektroniku, mikroelektroniku, računalne i inteligentne sustave Fakulteta elektrotehnike i računarstva Sveučilišta u Zagrebu.

Dio istraživanja je obavljen na Fakultetu prirodoslovno-matematičkih znanosti i kineziologije Sveučilišta u Splitu, te na Kemijsko tehnološkom fakultetu Sveučilišta u Splitu.

Mentor: prof.dr.sc. Slavomir Stankov sa Fakulteta prirodoslovno-matematičkih znanosti i kineziologije Sveučilišta u Splitu

Magistarski rad sadrži 168 stranica teksta bez priloga, a prilozi imaju 26 stranica.

Rad br.:

Povjerenstvo za ocjenu magistarskog rada:

1. prof.dr.sc. Vlado Glavinić, Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu – predsjednik
2. prof.dr.sc. Slavomir Stankov – Fakultet prirodoslovno-matematičkih znanosti i kineziologije Sveučilišta u Splitu – mentor
3. prof.dr.sc. Bojana Dalbelo-Bašić, Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu

Povjerenstvo za obranu magistarskog rada:

4. prof.dr.sc. Vlado Glavinić, Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu – predsjednik
5. prof.dr.sc. Slavomir Stankov – Fakultet prirodoslovno-matematičkih znanosti i kineziologije Sveučilišta u Splitu – mentor
6. prof.dr.sc. Bojana Dalbelo-Bašić, Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu

Datum obrane magistarskog rada: 01.02.2007.

ZAHVALA

Zahvaljujem svom mentoru prof.dr.sc. Slavomiru Stankovu koji mi je strpljivo posvetio svoj trud i vrijeme kako bi mi pomogao vrijednim činjenicama i mišljenjima te brojnim stručnim savjetima i korisnim raspravama usmjeravao tijek pisanja rada.

Zahvaljujem svom kolegi mr.sc. Branku Žitku na pomoći pruženoj prilikom realizacije implementacije programskog sustava.

Hvala svim dragim i bliskim ljudima koji su mi pružali podršku. Najveću zahvalu upućujem svojim najdražima – Fabjanu i Mislavu – koji su moja najveća inspiracija i kojima dugujem zahvalnost za ljubav, vjeru i podršku.

SAŽETAK

S obzirom na veliku prisutnost e-učenja u svijetu obrazovanja, potrebno je osigurati valjano i učinkovito okruženje u kojem bi se e-učenje odvijalo. Upravo zbog činjenice da sustavi e-učenja zbog svoje prisutnosti u procesu učenja i poučavanja, utječu na taj isti proces, kao i na postignuća učenika, smatramo da se u nastavi mogu koristiti samo oni sustavi čija je učinkovitost provjerena.

U ovom radu smo inteligentne tutorske sustave prikazali kao posebnu vrstu inteligentnih sustava e-učenja, a samim time i sustava e-učenja općenito. Na temelju spoznaja iz literature dali smo pregled i klasifikaciju metoda vrednovanja, objasnili postupke i instrumente prikupljanja podataka. Pošto je proces vrednovanja učinka dugotrajan, a izbor metoda za vrednovanje velik, odlučili smo pristupiti definiranju vlastite metodologije koja predstavlja sintetički pristup vrednovanju učinka kojim se povezuju eksperiment i ideja o provjeravanju učinka sustava u što većem broju međustanja. Ova metodologija je implementirana u okviru programskog alata za automatsko vrednovanje učinka sustava e-učenja koji omogućava jednostavno vrednovanje učinka.

Nastavak istraživanja je orijentiran prema primjeni definirane metodologije, kao i programskog alata za automatsko vrednovanje učinka sustava e-učenja, za vrednovanje učinka što većeg broja sustava e-učenja različitih kategorija, što bi omogućilo provedbu meta-analize i procjenu sveukupne veličine učinka različitih podkategorija sustava e-učenja.

Ključne riječi:

e-učenje, sustavi e-učenja, inteligentni sustavi e-učenja, inteligentni tutorski sustavi, vrednovanje učinka, veličina učinka, 2-sigma problem

SUMMARY

Concerning great presence of the e-learning in education, it is necessary to enable valid and efficient environment for its realization. Due to the fact that e-learning systems influence learning and teaching process, as well as students' achievements, we believe that only those systems, whose efficiency has been evaluated, can be used in educational process.

In this thesis we have shown that intelligent tutoring systems are one special class of intelligent e-learning systems, and e-learning systems in general. Based on literature findings, we have presented classification of evaluation methods, as well as explained methods and instruments for gathering data. As efficiency evaluation process is long-lasting and choice of evaluation methods is vast, we have decided to define our own evaluation methodology that presents synthetic approach that relates an experiment and an idea about checking system's efficiency in intermediate states. This methodology has been implemented in a system for automatic evaluation of e-learning systems' educational influence that enables easy efficiency calculation.

Further research is oriented towards application of the defined methodology, as well as the system for automatic evaluation, to evaluate efficiency of as much as possible e-learning systems. That would enable calculation of the overall effectiveness of different categories of e-learning systems using meta-analysis.

Key words:

e-learning, e-learning systems, intelligent e-learning systems, intelligent tutoring systems, evaluating educational influence, effect size, 2-sigma problem

ŽIVOTOPIS

Rođena sam 10.11.1978. godine u Splitu. Završila sam Prirodoslovno – matematičku gimnaziju (III. Gimnazija) u Splitu s odličnim uspjehom. Sudjelovala sam na natjecanjima iz matematike i informatike na kojima sam ostvarila značajne rezultate.

Nakon devetogodišnjeg učenja engleskog jezika u Centru za strane jezike u Splitu, položila sam ispit First Certificate in English. Bila sam članica organizacije Mensa Hrvatska.

Diplomirala sam 27.11.2001. godine s odličnim uspjehom na Fakultetu prirodoslovno – matematičkih znanosti i odgojnih područja Sveučilišta u Splitu i postigla visoku spremu i stručno zvanje profesor matematike i informatike. Tema mog diplomskog rada je „Model traganja – dijagnostička tehnika inteligentnih tutorskih sustava“.

Od 01.01.2002. zaposlena sam kao znanstveni novak na Fakultetu prirodoslovno – matematičkih znanosti i odgojnih područja Sveučilišta u Splitu, te sam od 2002. godine prijavljena kao suradnik na znanstvenoistraživačkom projektu 177110 „Računalni i didaktički aspekti inteligentnih autorskih alata u obrazovanju“.

Sveučilišni znanstveni poslijediplomski studij na Fakultetu elektrotehnike i računarstva smjer Jezgra računarstva upisala sam 28.02.2002. i položila ispite iz svih upisanih i odslušanih kolegija.

19. prosinca 2002. godine stekla sam istraživačko zvanje mlađeg asistenta (znanstveni novak) za znanstveno područje Tehničkih znanosti, polje Računarstva.

Sudjelovala sam kao suradnik na Tehnologijskom projektu TP-02/0177-01 od 2003. do 2005. godine.

Između ostalih pomagala sam u izvođenju praktičnog dijela nastave (vježbi) na sljedećim kolegijima: Uvod u računarstvo, Programiranje I, Računalni praktikum I (vizualno modeliranje), Računalni praktikum II (programiranje skriptnim jezicima).

Znanstveni i stručni radovi:

1. Grubišić, A., Stankov, S., Žitko, B. (2006). “An Approach to Automatic Evaluation of Educational Influence”, Proceedings of the 6th WSEAS International Conference on DISTANCE LEARNING and WEB ENGINEERING (DIWEB '06), Lisbon, Portugal, September 22-24, 2006, ISSN: 1790-5117, pp.20-25.
 2. Grubišić, A., Stankov, S., Žitko, B. (2006). “EVEDIN: A System for Automatic Evaluation of Educational Influence”, in International Journal WSEAS Transactions on Computers, Vol. 6, Issue 1, January 2007, ISSN: 1109-2750, pp.95-102.
 3. Grubišić, A., Stankov, S., Žitko, B. (2005). “Evaluating the educational influence of an e-learning system”, Proceedings of the International Conference, CEEPUS Summer School 2005, Intelligent Control Systems, 1 (2005), pp.151-156.
 4. Stankov, S., Žitko, B., Grubišić, A. (2005). “Ontology as a Foundation for Knowledge Evaluation in Intelligent E-learning Systems”, International Workshop on Applications of Semantic Web Technologies for E-Learning (SW-EL'05) in conjunction with 12th International Conference on Artificial Intelligence in Education (AI-ED 2005), Amsterdam, Netherlands, pp. 81-84.
-

-
5. Stankov, S., Grubišić, A., Žitko, B., Krpan, D. (2005). „Vrednovanje učinkovitosti procesa učenja i poučavanja u sustavima za e-učenje“, Školski Vjesnik - časopis za pedagoška i školska pitanja, 54 (2005) , 1-2; 21-31
 6. Stankov, S., Glavinić, V., Grubišić, A. (2004). „What is our effect size: Evaluating the Educational Influence of a Web-Based Intelligent Authoring Shell?“, IEEE International Conference on Intelligent Engineering Systems 2004 - INES 2004, Cluj-Napoca, Romania
 7. Stankov, S., Rosić, M., Granić, A., Maleš, L., Grubišić, A., Žitko, B. (2004). „Paradigma e-učenja & Inteligentni tutorski sustavi“, MIPRO-2004, Računala u obrazovanju, Opatija
 8. Amižić, A., Stankov, S., Rosić, M. (2002). „Model Tracing – A Diagnostic Technique in Intelligent Tutoring Systems“, CEEPUS Summer school - Modern Methods in Control Split 2002 Jointly with Fifth Symposium on Intelligent Systems, Split, Croatia (Proc. – CD ROM version Reprints of CEEPUS CZ-0103, ISBN: 953-96516-8-9).
 9. Amižić, A., Stankov, S., Rosić, M. (2002). „Model traganja – dijagnostička tehnika inteligentnih tutorskih sustava“, MIPRO-2002, Računala u obrazovanju, Opatija, str. 101 - 106.
 10. Amižić, A. (2001). „Model učitelja u inteligentnim tutorskim sustavima“, MIPRO-2001, Računala u obrazovanju, Opatija, str. 89-91.

CURRICULUM VITAE

I was born on the 10.11.1978. in Split. I have finished the Gymnasium for natural sciences and mathematics in Split with excellent marks. I have participated in many competitions in mathematics and computer science where I have achieved significant results.

After nine years of learning English language in the Center for foreign languages in Split, I have passed the First Certificate in English. I was a member of organization Mensa Croatia.

I graduated on 27.11.2001. with excellent success on the Faculty of natural sciences and kinesiology University of Split and I gained a degree for mathematics and computer science teacher. The theme of my graduate thesis was “Model tracing – a diagnostic technique of intelligent tutoring systems”.

Since 01.01.2002. I have been employed as a scientific novice at the Faculty of natural sciences and kinesiology University of Split. Since 2002. I have been collaborating on a scientific and research project 177110 „Computational and didactical aspects of intelligent authoring tools in education “.

I enrolled postgraduate study at the Faculty of Electrical Engineering and Computing on the 28.02.2002. and I have passed all exams.

I was collaborating on a technological project TP-02/0177-01 from 2003. to 2005.

Among others, I have helped realizing exercises of the following courses: Introduction to computer science, Programming I, Computational practicum I (visual modeling), Computational practicum II (script languages programming).

SADRŽAJ:

1. UVOD	1
2. E-UČENJE I INTELIGENTNI TUTORSKI SUSTAVI	6
2.1. INTELIGENTNI TUTORSKI SUSTAVI I INTELIGENTNE AUTORSKE LJUSKE	6
2.2. MODEL TEX-SYS I SUSTAV XTEX-SYS.....	12
2.2.1. Model <i>TeX-Sys</i>	12
2.2.2. <i>xTeX-Sys</i>	15
2.2.2.1 Oblikovanje baza područnog znanja.....	17
2.2.2.2 Oblikovanje nastavnog sadržaja i ocjenjivanje znanja učenika	19
2.2.2.3 Učenje, poučavanje i testiranje znanja učenika	21
2.3. PRIMJENE MODELA TEX-SYS – KRONIKA PROVEDENIH ISTRAŽIVANJA.....	25
2.3.1. Testiranja na sustavu <i>TeX-Sys</i>	26
2.3.2. Testiranja na sustavu <i>DTeX-Sys</i>	26
2.3.3. Testiranja na sustavu <i>xTeX-Sys</i>	30
2.3.4. Vrednovanja učinka	31
3. STATISTIČKE OSNOVE I METODE VREDNOVANJA SUSTAVA E-UČENJA	36
3.1. STATISTIČKE OSNOVE	36
3.1.1. <i>Temeljni pojmovi</i>	37
3.1.2. <i>Karakteristike distribucije frekvencija</i>	38
3.1.2.1 Mjere centralne tendencije ili srednje vrijednosti	38
3.1.2.2 Mjere varijabilnosti ili disperzije.....	41
3.1.3. <i>Položaj pojedinog rezultata u grupi</i>	45
3.1.3.1 Standardizirano odstupanje (z vrijednost)	45
3.1.3.2 Centili i decili	46
3.1.4. <i>Normalna i t-distribucija</i>	47
3.1.5. <i>Statistički testovi razlike</i>	48
3.1.5.1 t-test.....	48
3.1.5.2 Analiza varijance - ANOVA	49
3.1.6. <i>Korelacija</i>	52
3.1.7. <i>Grafičko prikazivanje rezultata</i>	54
3.1.7.1 Histogram i poligon frekvencija	54
3.1.7.2 Box and Whisker plot.....	56
3.1.7.3 Dijagram rasipanja	56
3.2. METODE VREDNOVANJA SUSTAVA E-UČENJA	58
3.2.1. <i>Modeli vrednovanja</i>	59
3.2.2. <i>Metode vrednovanja</i>	61
3.2.2.1 Pregled metoda vrednovanja	62
3.2.2.2 Klasifikacija metoda vrednovanja	66
3.2.3. <i>Postupci i instrumenti prikupljanja podataka</i>	68
3.2.3.1 Karakteristike instrumenata	68
3.2.3.2 Vrste postupaka i instrumenata.....	74
3.2.4. <i>Eksperiment kao metoda vrednovanja</i>	78
3.2.4.1 Zavisna i nezavisna varijabla.....	78
3.2.4.2 Vrste pogrešaka	80
3.2.4.3 Osnovni modeli eksperimentalnih postupaka	81
3.2.4.4 Složeni modeli eksperimentalnih postupaka.....	90
3.3. VELIČINA UČINKA	94
3.3.1. <i>Mjere postignuća</i>	94
3.3.2. <i>Mjere afekta</i>	95
3.3.3. <i>Vrste veličina učinka</i>	96
3.3.3.1 Razlika aritmetičkih sredina	96
3.3.3.2 Standardna razlika aritmetičkih sredina – glass Δ	96
3.3.3.3 Cohen d	97

3.3.3.4	Hedges g.....	98
3.3.4.	<i>Intervali pouzdanosti</i>	99
3.3.5.	<i>Interpretacija veličine učinka</i>	100
3.3.6.	<i>2-sigma problem</i>	102
3.3.7.	<i>Meta-analiza</i>	104
3.3.7.1	Meta-analize o učinku tehnologijom potpomognutog učenja i poučavanja.....	105
3.4.	PRIMJERI PROVEDENIH ISTRAŽIVANJA O VREDNOVANJU UČINKA SUSTAVA E-UČENJA.....	107
3.4.1.	<i>Spoken Conversational Tutor - SCoT</i>	107
3.4.2.	<i>AutoTutor</i>	108
3.4.3.	<i>Equation tutor - E-tutor</i>	109
3.4.4.	<i>Reusable Educational Design Environment and Engineering Methodology - REDEEM</i>	109
3.4.5.	<i>Learning by asking - LBA</i>	110
3.4.6.	<i>Conceptual Helper</i>	111
3.4.7.	<i>SQL- Tutor</i>	112
3.4.8.	<i>NORMIT</i>	112
3.4.9.	<i>System Dynamics Model</i>	113
4.	METODOLOGIJA ZA VREDNOVANJE UČINKA PROCESA UČENJA I POUČAVANJA U SUSTAVIMA E-UČENJA.....	116
4.1.	PROCES VREDNOVANJA.....	116
4.1.1.	<i>Oblikovanje eksperimenta</i>	117
4.1.1.1	Izrada i provedba inicijalnog testa.....	117
4.1.1.2	Definiranje ekvivalentnih grupa.....	117
4.1.1.3	Izrada i provedba testova provjere te završnog testa.....	121
4.1.2.	<i>Definiranje nul-hipoteza</i>	121
4.1.3.	<i>Izračun veličine učinka</i>	122
4.1.4.	<i>Analiza i interpretacija rezultata vrednovanja učinka</i>	123
4.2.	VREDNOVANJE UČINKA SUSTAVA XTEX-SYS.....	125
4.2.1.	<i>Definiranje ekvivalentnih grupa</i>	126
4.2.2.	<i>Definiranje nul-hipoteza</i>	131
4.2.3.	<i>Izračun veličine učinka</i>	134
5.	PROTOTIP PROGRAMSKOG ALATA ZA VREDNOVANJE UČINKA SUSTAVA E-UČENJA.....	138
5.1.	PRIMJERI ALATA ZA IZRADU I PROVEDBU ELEKTRONIČKIH TESTOVA ZNANJA.....	138
5.2.	FUNKCIONALNOST I ARHITEKTURA EVEDIN-A.....	139
5.2.1.	<i>Funkcionalnosti učitelja</i>	140
5.2.2.	<i>Funkcionalnosti učenika</i>	144
5.2.3.	<i>Struktura programskog rješenja</i>	144
5.2.3.1	Red podataka.....	146
5.2.3.2	Red programske logike.....	149
5.2.3.3	Prezentacijski red.....	149
6.	ZAKLJUČAK.....	159
7.	LITERATURA.....	161
8.	PRILOZI.....	169

Popis slika:

SLIKA 2. 1. KOMPONENTE ITS-A (PREMA [BURN1988]).....	8
SLIKA 2. 2. VERZIJE MODELA TEX-SYS	12
SLIKA 2. 3. KIBERNETIČKI MODEL SUSTAVA [STAN2001]	12
SLIKA 2. 4. MODEL VOĐENJA PROCESA STJECANJA ZNANJA I VJEŠTINA UČENIKA U INTELIGENTNOM TUTORSKOM SUSTAVU [STAN2001]	13
SLIKA 2. 5. SEMANTIČKA MREŽA S TRI ČVORA I DVIJE VEZE	13
SLIKA 2. 6. SEMANTIČKA MREŽA S OKVIRIMA.....	14
SLIKA 2. 7. SEMANTIČKA MREŽA S OKVIRIMA I STRUKTURNIM ATRIBUTOM TIPA SLIKA	14
SLIKA 2. 8. STRUKTURA AUTORSKE LJUSKE	16
SLIKA 2. 9. ARHITEKTURA SUSTAVA xTeX-Sys	17
SLIKA 2. 10. PRIJAVA NA SUSTAV xTeX-Sys	17
SLIKA 2. 11. FORMA ZA IZBOR POSTOJEĆEG PODRUČJA I PODPODRUČJA	18
SLIKA 2. 12. GRAFIČKI PRIKAZ POVEZANOSTI ČVORA ZNANJA S ČVOROVIMA TIPA RODITELJ I DIJETE	18
SLIKA 2. 13. RAŠČLANA NASTAVNOG SADRŽAJA KOLEGIJA	19
SLIKA 2. 14. ODNOS ELEMENATA STRUKTURE PODRUČNOG ZNANJA I NASTAVNOG SADRŽAJA U SUSTAVU xTeX-Sys.....	20
SLIKA 2. 15. ODREĐIVANJE MJESTA I VREMENA UČENJA, POUČAVANJA I TESTIRANJA ZNANJA	21
SLIKA 2. 16. IZBOR KOLEGIJA	22
SLIKA 2. 17. IZABRANI KOLEGIJ	22
SLIKA 2. 18. POGLED NA DIO SADRŽAJA NASTAVNE CJELINE	23
SLIKA 2. 19. PRIKAZ SADRŽAJA SCO-A	23
SLIKA 2. 20. AKTIVIRANJE TESTA	24
SLIKA 2. 21. SERIJA OD DVA PITANJA.....	24
SLIKA 2. 22. REZULTATI O IZVRŠENOM TESTIRANJU	25
SLIKA 3. 1. NEKE POVRŠINE ISPOD NORMALNE KRIVULJE	47
SLIKA 3. 3. HISTOGRAM I POLIGON FREKVENCIJA	55
SLIKA 3. 4. OBLICI KRIVULJA DISTRIBUCIJE: A) MEZOKURTIČNA, B) ŠILJASTA, C) SPLJOŠTENA.....	55
SLIKA 3. 5. BOX AND WHISKER PLOT.....	56
SLIKA 3. 6. OBLICI DIJAGRAMA RASIPANJA.....	57
SLIKA 3. 8. SHEMA ODNOSA ZAVISNE I NEZAVISNE VARIJABLE.....	79
SLIKA 3. 9. JEDNOSTAVAN MODEL EKSPERIMENTA	82
SLIKA 3. 10. EKSPERIMENT S JEDNOM GRUPOM	84
SLIKA 3. 11. EKSPERIMENT S PARALELNIM GRUPAMA	86
SLIKA 3. 12. EKSPERIMENT S TRI PARALELNE GRUPE	87
SLIKA 3. 13. EKSPERIMENT S ROTACIJOM FAKTORA.....	89
SLIKA 3. 14. LATINSKI 3X3 KVADRAT.....	92
SLIKA 3. 15. GRČKO-LATINSKI 4X4 KVADRAT	93
SLIKA 3. 16. DISTRIBUCIJA POSTIGNUĆA UČENIKA KOD TRADICIONALNOG UČENJA, UČENJA S PROVJERAVANJEM I TUTORSKOG UČENJA [BLOO1984]	103
SLIKA 3. 17. 2-SIGMA RAZLIKA	104
SLIKA 3. 18. VELIČINE UČINKA TEHNOLOGIJOM POTPOMOGNUTOG UČENJA I POUČAVANJA [FLET2003].....	106
SLIKA 3. 19. GRAFIČKI PRIKAZ REZULTATA EKSPERIMENTA [PONB2004].....	108
SLIKA 4. 1. VIZUALNA INTERPRETACIJA RAZLIKE IZMEĐU DVIJE GRUPE	125
SLIKA 4. 2. HISTOGRAM REZULTATA INICIJALNOG TESTA KONTROLNE GRUPE.....	127
SLIKA 4. 3. HISTOGRAM REZULTATA INICIJALNOG TESTA PROŠIRENE EKSPERIMENTALNE GRUPE.....	127
SLIKA 4. 4. ZAJEDNIČKI HISTOGRAM KONTROLNE I EKSPERIMENTALNE GRUPE	130
SLIKA 4. 5. HISTOGRAM I BOX-AND-WHISKER PLOT ZA PRVI TEST PROVJERE	132
SLIKA 4. 6. HISTOGRAM I BOX-AND-WHISKER PLOT ZA DRUGI TEST PROVJERE	133
SLIKA 4. 7. HISTOGRAM I BOX-AND-WHISKER PLOT ZA ZAVRŠNI TEST	134

SLIKA 4. 8. DIJAGRAM TOKA METODOLOGIJE ZA VREDNOVANJE	137
SLIKA 5. 1. FUNKCIONALNI MODEL EVEDIN-A.....	140
SLIKA 5. 2. ARHITEKTURA EVEDIN-A	145
SLIKA 5. 3. EVEDIN – DIJAGRAM KOMPONENTI.....	145
SLIKA 5. 4. STRUKTURA BAZE PODATAKA	146
SLIKA 5. 5. EVEDIN – RED PODATAKA	148
SLIKA 5. 6. EVEDIN – RED PROGRAMSKE LOGIKE	149
SLIKA 5. 7. FORMA ZA PRIJAVU NA SUSTAV	150
SLIKA 5. 8. POČETNA FORMA ZA UČITELJA	150
SLIKA 5. 9. FORMA ZA DODAVANJE I BRISANJE EKSPERIMENTA, TE MIJENJANJE POSTAVKI EKSPERIMENTA	151
SLIKA 5. 10. FORMA ZA DODAVANJE, MIJENJANJE I BRISANJE UČENIKA, TE PRIDRUŽIVANJE UČENIKA U EKSPERIMENT	151
SLIKA 5. 11. FORMA ZA DODAVANJE, MIJENJANJE I BRISANJE TESTA	152
SLIKA 5. 12. FORMA ZA PRIDRUŽIVANJE TESTOVA U EKSPERIMENT	152
SLIKA 5. 13. FORMA ZA DODAVANJE, MIJENJANJE I BRISANJE PITANJA, TE PREGLED I BRISANJE PITANJA U TESTU.....	153
SLIKA 5. 14. FORMA ZA PREGLED TESTA.....	153
SLIKA 5. 15. FORMA ZA RUČNO UNOŠENJE REZULTATA TESTA.....	154
SLIKA 5. 16. FORMA ZA UKUPNI PREGLED REZULTATA TESTA ZA ODABRANI EKSPERIMENT	154
SLIKA 5. 17. FORMA ZA IZJEDNAČAVANJE GRUPE.....	155
SLIKA 5. 18. FORMA ZA ISPITIVANJE NUL-HIPOTEZA I VREDNOVANJE UČINKA	155
SLIKA 5. 19. POČETNA FORMA ZA UČENIKA	156
SLIKA 5. 20. FORMA ZA RJEŠAVANJE TESTA.....	156
SLIKA 5. 21. FORMA ZA PREGLED REZULTATA POJEDINOG TESTA ZA JEDNOG UČENIKA	157
SLIKA 5. 22. FORMA ZA ODJAVU IZ SUSTAVA.....	157

Popis tablica:

TABLICA 2. 1. KATEGORIJE ITS AUTORSKIH ALATA (PREMA [MURR2003])	11
TABLICA 2. 2. BODOVNI KRITERIJI	15
TABLICA 2. 3. OCJENA TESTA	15
TABLICA 2. 4. ZNAČENJE SIMBOLA U AUTORSKOM OKRUŽENJU ZA OBLIKOVANJE BAZE PODRUČNOG ZNANJA	18
TABLICA 2. 5. ZNAČENJE SIMBOLA U OKRUŽENJU ZA OBLIKOVANJE NASTAVNOG SADRŽAJA	19
TABLICA 2. 6. ZNAČENJE SIMBOLA U OKRUŽENJU ZA UČENJE I POUČAVANJE.....	21
TABLICA 2. 7. BROJ TESTIRANJA REALIZIRANIH NA MODELU TEX-SYS PO AKADEMSKIM GODINAMA	34
TABLICA 2. 8. BROJ UČENIKA KOJI SU SUDJELOVALI U TESTIRANJIMA NA MODELU TEX-SYS PO AK. GODINAMA.....	34
TABLICA 2. 9. PREGLED BAZA PODRUČNIH ZNANJA NA KOJIMA SU SE VRŠILA TESTIRANJA.....	34
TABLICA 3. 1. PODJELA MJERA DISPERZIJA	41
TABLICA 3. 2. ODNOS MJERA DISPERZIJE I SREDNJIH VRIJEDNOSTI	41
TABLICA 3. 3. RASPON REZULTATA	42
TABLICA 3. 4. POGREŠKE PRI ZAKLJUČIVANJU IZ UZORKA NA POPULACIJU.....	49
TABLICA 3. 5. TABLICA ANALIZE VARIJANCE.....	51
TABLICA 3. 6. MODELI VREDNOVANJA (PREMA [ALEX1994]).....	59
TABLICA 3. 7. ODNOS FAZA RAZVOJA SUSTAVA I VREDNOVANJA SUSTAVA (PREMA [ALEX1994])	61
TABLICA 3. 8. LATINSKI 3X3 KVADRAT	93
TABLICA 3. 9. GRČKO-LATINSKI 4X4 KVADRAT	93
TABLICA 3. 10. INTERPRETACIJA VELIČINE UČINKA (PROMIENJENO PREMA [COE2000] I [BACK2000]).....	101
TABLICA 3. 11. NEKE VELIČINE UČINKA DOBIVENE KROZ ISTRAŽIVANJA [COE2000]	102
TABLICA 3. 12. UČINAK NEKIH FAKTORA NA POUČAVANJE [BLOO1984]	104
TABLICA 3. 13. KORELACIJE RAZMJERA STUDENTOVA ZNANJA I DIJALOŠKIH AKCIJA	108
TABLICA 3. 14. REZULTATI INICIJALNOG I ZAVRŠNOG TESTIRANJA.....	110
TABLICA 3. 15. GRUPE U EKSPERIMENTU.....	110
TABLICA 3. 16. ANALIZA REZULTATA PO GRUPAMA (RAZLIKA BODOVA IZMEĐU ZAVRŠNOG I INICIJALNOG TESTA) ...	111
TABLICA 3. 17. ANALIZA REZULTATA IZMEĐU GRUPE (RAZLIKA BODOVA IZMEĐU ZAVRŠNOG I INICIJALNOG TESTA) ..	111
TABLICA 3. 18. ARITMETIČKE SREDINE REZULTATA EKSPERIMENTA (U ZAGRADAMA SU STANDARDNE DEVIJACIJE)..	113
TABLICA 3. 19. ORGANIZACIJA EKSPERIMENTA	113
TABLICA 3. 20. REZULTATI EKSPERIMENTA ZA SVAKU GRUPU I SVAKU ZAVISNU VARIJABLU (1)	114
TABLICA 3. 21. REZULTATI EKSPERIMENTA ZA SVAKU GRUPU I SVAKU ZAVISNU VARIJABLU (2)	114
TABLICA 4. 1. REZULTATI INICIJALNOG TESTA	118
TABLICA 4. 2. PAROVI ISPITANIKA NAKON IZJEDNAČAVANJA	118
TABLICA 4. 3. IZJEDNAČAVANJE U RASPONU	119
TABLICA 4. 4. PODACI POTREBNI ZA PRIHVAĆANJE ILI ODBACIVANJE NUL-HIPOTEZE	122
TABLICA 4. 5. PODACI POTREBNI ZA IZRAČUN PARCIJALNIH VELIČINA UČINKA	123
TABLICA 4. 6. REZULTATI ANALIZE	125
TABLICA 4. 7. ETAPE EKSPERIMENTA	126
TABLICA 4. 8. TABLICA NORMALNOSTI ZA KONTROLNU GRUPU.....	127
TABLICA 4. 9. TABLICA NORMALNOSTI ZA PROŠIRENU EKSPERIMENTALNU GRUPU	128
TABLICA 4. 10. FREKVENCije BODOVA INICIJALNOG TESTA KONTROLNE I PROŠIRENE EKSPERIMENTALNE GRUPE	128
TABLICA 4. 11. IZJEDNAČAVANJE GRUPE PO RASPONU	129
TABLICA 4. 12. FREKVENCije EKSPERIMENTALNE I KONTROLNE GRUPE	129
TABLICA 4. 13. PODACI POTREBNI ZA ISPITIVANJE STATISTIČKI ZNAČAJNE RAZLIKE	130
TABLICA 4. 14. PODACI POTREBNI ZA PRIHVAĆANJE ILI ODBACIVANJE NUL-HIPOTEZE H1	131
TABLICA 4. 15. PODACI POTREBNI ZA PRIHVAĆANJE ILI ODBACIVANJE NUL-HIPOTEZE H2	132
TABLICA 4. 16. PODACI POTREBNI ZA PRIHVAĆANJE ILI ODBACIVANJE NUL-HIPOTEZE H3	133
TABLICA 4. 17. REZULTATI ISPITIVANJA NUL-HIPOTEZA.....	134
TABLICA 4. 18. PODACI POTREBNI ZA IZRAČUN PARCIJALNIH VELIČINA UČINKA	134
TABLICA 5. 1. USPOREDBA ALATA ZA IZRADU I PROVEDBU ELEKTRONIČKIH TESTOVA.....	139

1. UVOD

S obzirom na rastuće interese za razvoj sustava koji podupiru proces e-učenja i svekoliku prisutnost e-učenja u svijetu obrazovanja, potrebno je osigurati valjano i učinkovito okruženje u kojem bi se e-učenje odvijalo. *E-učenje* u središte stavlja učenika za razliku od tradicionalne nastave u razredu koja u središte stavlja učitelja. Najčešće korištena definicija e-učenja jest da je e-učenje skup aplikacija i procesa, kao što su učenje temeljeno na Web-u (eng. Web-based learning), učenje temeljeno na računalu (eng. computer-based learning), virtualni razredi (eng. virtual classrooms) i digitalna suradnja (eng. digital collaboration), koji omogućavaju pristup nastavnim sadržajima pomoću različitih elektroničkih medija (CD-ROM, Internet, intranet, extranet, audio i video, satelit, itd.) [ASTD2001].

Sustavi e-učenja omogućavaju pristup elektroničkim izvorima za učenje bilogdje, bilo kada i bilo kome [ALBE2001]. Danas se skoro na svakom fakultetu na kojem postoji odjel za računarstvo, pristupa oblikovanju i implementaciji sustava e-učenja. Iako se ti sustavi razvijaju u svrhu potpore nastavi, mali je broj sustava koji podupiru proces e-učenja i koji su vrednovani uporabom neke od metodologija za vrednovanje. Naime, vrednovanje sustava, a pogotovo vrednovanje učinka sustava, zahtijeva dugotrajan angažman kako nastavnika koji te sustave koriste u nastavi, tako i studenata koji prisustvuju toj nastavi.

Glavni cilj sustava e-učenja je učenje, a glavni način vrednovanja učinka sustava e-učenja je ispitivanje da li učenici učinkovito uče uz pomoć tog sustava (prema [MARK1993]). *Vrednovanje* se koristi za ispitivanje kako različiti i novi načini uporabe informacijske i komunikacijske tehnologije podupiru proces učenja i poučavanja. Programaska podrška koja se koristi u obrazovne svrhe treba biti vrednovana prije korištenja u procesu učenja i poučavanja. Dobro oblikovano vrednovanje trebalo bi dati dokaz da li je određeni pristup uspješan i da li ima potencijalnu vrijednost za druge [DEMP2004].

Upravo zbog činjenice da sustavi e-učenja zbog svoje prisutnosti u procesu učenja i poučavanja, utječu na taj isti proces, kao i na postignuća učenika, smatramo da se u nastavi mogu koristiti samo oni sustavi čija je učinkovitost provjerena. Pošto je proces vrednovanja učinka dugotrajan, a izbor metoda za vrednovanje velik, odlučili smo pristupiti definiranju vlastite metodologije za vrednovanje učinka sustava e-učenja te izgraditi prototip programskog alata koji će omogućiti jednostavno vrednovanje učinka za po volji odabran sustav e-učenja.

Izbor *metodologije vrednovanja* ovisi o izboru pitanja na koje želimo dobiti odgovor, jer ne postoji jedinstven model za vrednovanje učinka sustava e-učenja. Zatim se pristupa definiranju *nul-hipoteza* koje se moraju moći provjeriti da bi ih se moglo potvrditi ili odbiti na temelju određenih uvjeta i rezultata. Potom slijedi definiranje metodologije za vrednovanje kojom će se ispitati istinitost nul-hipoteza. Nadalje, pristupa se provođenju istraživanja te analiziranju podataka. U idealnim okolnostima, ako dobiveni rezultati ne potvrde nul-hipotezu, istraživači bi trebali objasniti razloge zbog kojih su dobiveni rezultati takvi kakvi jesu.

Složeni sustavi, kao što su sustavi e-učenja, mogu se promatrati u terminima cjelovitih sustava, komponenti sustava i specifičnih značajki. Tehnike koje su prikladne za vrednovanje cjelovitih sustava ne odgovaraju u potpunosti vrednovanju komponenti ili značajki sustava i obrnuto (interno i eksterno vrednovanje). Metoda vrednovanja ovisi o svrsi istraživanja, tj. prvo, treba usporediti koliko je učinkovit sustav u odnosu na tradicionalni način poučavanja i drugo, treba odrediti značajke sustava koje su važne za vrednovanje sustava. Ako želimo vrednovati učinak sustava e-učenja, onda vrednujemo cijeli sustav, a ne samo jedan njegov dio, pa se radi o eksternom vrednovanju. Eksperimentalno istraživanje je prikladno za vrednovanje sustava e-učenja jer omogućava ispitivanje odnosa između poučavanja i učenikovih rezultata te dobivanje kvantitativnih mjera značajnosti tih odnosa.

Eksperiment predstavlja metodološki način pristupa istraživanja pedagoške pojave kojim se nastoje uočiti uzročno-posljedične veze među pojavama, služeći se pri tom raznim postupcima prikupljanja podataka, kao što je promatranje, testiranje, razgovor i sl. [MUŽI1977]. Osnovna je karakteristika pedagoškog eksperimenta plansko ispitivanje učinka procesa učenja i poučavanja. Osnovni cilj eksperimenta jest utvrditi da li postoji, koliki je i kakav je utjecaj određene nezavisne varijable na određenu zavisnu varijablu. U našem pristupu zavisna varijabla je znanje učenika, a nezavisna varijabla je sustav e-učenja čiji učinak želimo ispitati.

Jedno od najpoznatijih istraživanja u području vrednovanja učinka procesa učenja i poučavanja je definitivno Bloom-ovo istraživanje o razlikama između tradicionalnog učenja (eng. conventional learning), učenja s provjeravanjem (eng. mastery learning) i tutorskog učenja (eng. tutoring learning) [BLOO1984]. Studenti su bili po slučajnom uzorku raspoređeni u neku od ove tri grupe, ali su imali slična prethodna postignuća i sličnu razinu znanja koja je vezana za danu materiju. Vrijeme koje je utrošeno za nastavu jednako je u svim trima grupama, osim za korektivni postupak u grupi učenja s provjeravanjem i grupi s tutorskim učenjem. Bloom je u svom istraživanju ustvrdio da prosječan student u tutorskoj grupi za oko dvije standardne devijacije bolji od prosječnog studenta u tradicionalnoj grupi. Dobivenu razliku od dvije standardne devijacije Bloom je nazvao 2-sigma razlika.

Bloom je smatrao da je ova razlika od dvije standardne devijacije ideal prema kojem proces učenja i poučavanja treba stremiti, te je postizanje ove razlike u procesu učenja i poučavanja nazvao *2-sigma problem*. Stoga su se mnogi istraživači uhvatili ukoštac s ovim problemom ne bi li pronašli način kako iskoristiti informacijsku i komunikacijsku tehnologiju u svrhu postizanja učinka individualizirane nastave pod praktičnijim i realnijim uvjetima nego što je jedan na jedan učenje, koje je preskupo da bi se primijenilo na veći broj studenata..

Naša metodologija predstavlja sintetički pristup vrednovanju učinka kojim se povezuju eksperiment kao najčešće korištena metoda vrednovanja i ideja o provjeravanju učinka sustava u što većem broju međustanja. Ostale metodologije vrednovanja učinka sustava e-učenja promatraju učinak sustava samo na kraju eksperimenta, čime gube uvid u utjecaj sustava na postignuće učenika tijekom samog procesa učenja i poučavanja.

U svrhu našeg istraživanja definirali smo jednu varijantu eksperimentalnog istraživanja s paralelnim grupama [MUŽI1977]. Kod eksperimenta s paralelnim grupama postoje dvije grupe ispitanika, od kojih je svaka nosilac svog eksperimentalnog faktora - kontrolna grupa se poučava na tradicionalan način, a eksperimentalna grupa se poučava uz pomoć sustava e-učenja. Kod osnovnog modela eksperimenta s paralelnim grupama se pomoću inicijalnog i završnog testa mjeri inicijalno i završno znanje ispitanika. Kod modela eksperimentalnog istraživanja s

paralelnim grupama kojeg ovdje opisujemo, osim inicijalnog i završnog stanja, ispituje se znanje ispitanika i u međustanjima uporabom testova provjere. Stoga smo ovakav model nazivali *eksperiment s paralelnim grupama uz parcijalno ispitivanje stanja*.

Svi testovi koji se koriste u jednom eksperimentu trebaju biti potpuno komparativni. To znači da moraju biti bodovani na istoj ljestvici i da moraju ispitivati poznavanje istog područnog znanja. *Inicijalni test* se koristi zbog toga što u eksperimentu sudjeluju ispitanici s različitim vještinama i predznanjima, te je potrebno prikupiti informacije o postojanju statistički značajnih razlika u predznanju ispitanika. Moramo utvrditi početno stanje njihovog poznavanja i razumijevanja određenog područnog znanja da bi mogli kvantificirati veličinu njegove promjene. Obje grupe moraju pisati nekoliko 45-minutnih *testova provjere*, kao i 45-minutni *završni test* na kraju eksperimenta. Broj testova provjere se određuje prema trajanju eksperimenta (najmanje jedan mjesечно). Spomenuti testovi omogućavaju prikupljanje informacija o postojanju statistički značajnih razlika između dviju grupa tijekom trajanja eksperimenta. Raspon bodova svih testova je na ljestvici 0-100.

Način na koji se bira *uzorak ispitanika* nad kojim ćemo provesti istraživanje o učinkovitosti, utjecat će na sakupljene informacije. Ako sami biramo uzorak, imamo mogućnost odabira ispitanika koji su najspremniji za suradnju ili imaju najprikladnije sposobnosti. Zato se trebamo služiti slučajnim odabirom na osnovu rezultata inicijalnog testa, da bismo mogli dobiti što precizniju i bespogovornu razliku u učincima eksperimentalnih faktora. Ukoliko se pokaže da su prethodno definirane grupe statistički značajno različite (prema t-vrijednosti, razini značajnosti i stupnjevima slobode), potrebno ih je pokušati ponovno presložiti metodom slučajnog odabira i izjednačiti u rasponu, a zatim opet provjeriti da li su ekvivalentne.

U eksperimentima koji prate ovdje predloženu metodologiju definira se n hipoteza koje se odnose na n testova provjere, te jednu hipotezu koja se odnosi na završni test. Te *hipoteze* se tijekom ispitivanja statističke značajnosti rezultata odgovarajućeg testa, prihvaćaju ili odbacuju. Nul-hipoteza u statistici znači „nema razlike“ među pojavama koje mjerimo.

Dok testovi značajnosti govore o vjerojatnosti kojom se rezultati na testovima eksperimentalne grupe razlikuju od slučajnih rezultata, *veličina učinka* (eng. effect size) govori o veličini utjecaja eksperimentalnog faktora. Veličina učinka kvantificira veličinu razlike između dvije grupe te se stoga smatra stvarnom mjerom značajnosti te razlike, te može biti pozitivna ili negativna. Pozitivna je kada je eksperimentalna grupa bolja, a negativna je kada je kontrolna grupa bolja.

Veličina učinka se može izračunati uz pomoć različitih formula i pristupa, i njena vrijednost se može zbog toga uvelike razlikovati. Postoje četiri *vrste veličina učinka*: razlika aritmetičkih sredina (eng. mean difference), standardna razlika aritmetičkih sredina (eng. standardized mean difference) ili Glass-ova Δ , Cohen-ov d, Hedges-ov g. Za utvrđivanje razlika između grupa u eksperimentalnom istraživanju Mohammad [MOHA1998] preporuča uporabu standardne razlike aritmetičkih sredina.

Jedna od glavnih prednosti uporabe veličine učinka je u tome što se mogu koristiti različite veličine učinka dobivene kroz ponavljanja istog eksperimenta u svrhu dobivanja najbolje procjene veličine učinka. Ovaj proces uspoređivanja kvalitativnih rezultata eksperimenata i izračunavanje jedne vrijednosti veličine učinka, naziva se *meta-analiza*. Razvio ju je statističar Glass [GLAS1981] koji ju smatra analizom analiza i danas se koristi u svim granama znanosti.

Fletcher [FLET2003] je u svojoj meta-analizi proučio je 233 istraživanja o učinku nastave temeljene na računalima koja je uključivala računalne prezentacije s tekstem, slikama, jednostavnim animacijama, kao i malu razinu individualizacije. Pokazalo se da nastava temeljena na računalima ima veličinu učinka 0.39. Nastava pomoću interaktivne multimedije uključuje zvuk, složenije animacije i video, te je pokazala veličinu učinka 0.50. Inteligentni tutorski sustavi nastoje oponašati učenje i poučavanje jedan-na-jedan i pokazuju veličinu učinka od 0.84. Suvremeni inteligentni tutorski sustavi pokazuju obećavajućih 1.05.

U našem pristupu računamo *parcijalne veličine učinka* prema rezultatima svakog testa kao Glass-ove Δ , u kojima koristimo aritmetičke sredine razlika između rezultata testova provjere i završnog testa te inicijalnog testa za obje grupe studenata, kao i standardnu devijaciju odgovarajuće razlike za kontrolnu grupu. Zbog usporedbe, izračunat ćemo i Hedges-ov g koristeći zajedničku standardnu devijaciju. *Ukupna veličina učinka* računa se kao aritmetička sredina parcijalnih veličina učinka.

Nakon definiranja metodologije za vrednovanje učinka javila se potreba za stvaranjem programskog alata za automatsko vrednovanje učinka sustava e-učenja. Naglasak je pri tom na riječ „automatsko“, jer se učinak može izračunati „ručno“, koristeći programsku podršku za statističku obradu podataka, oslanjajući se na neku od metodologija vrednovanja, što može biti zamoran i dugotrajan posao. Programski alat za automatsko vrednovanje učinka omogućava jednostavno izračunavanje učinka. Samim time omogućava jednako tako jednostavnu provedbu meta-analiza nad rezultatima vrednovanja učinka dobivenih primjenom spomenutog programskog alata, jer su ti rezultati dobiveni primjenom iste metodologije vrednovanja. Upravljanje testovima (eng. test management), tj. izrada, isporuka i ocjenjivanje testova je nužno, ali ne i dovoljno, za vrednovanje učinka sustava e-učenja.

Metodologija vrednovanja koja je prethodno opisana, implementirana je i tvori središnji dio programskog alata EVEDIN (EValuation of EDucational INfluence) za automatsko vrednovanje učinka sustava e-učenja. EVEDIN omogućava upravljanje i izvođenje eksperimenata, tj. upravljanje kontrolnom i eksperimentalnom grupom, izradu i provođenje svih potrebnih testova, njihovo automatsko ispravljanje i ocjenjivanje, te, što je najvažnije, omogućava izračun veličine učinka eksperimentalnog faktora.

U drugom poglavlju prikazan je odnos e-učenja i inteligentnih tutorskih sustava. *Inteligentni sustavi e-učenja* imaju sposobnost prikladnog ponašanja u neizvjesnim situacijama koje se javljaju u procesu učenja i poučavanja. Posebna klasa inteligentnih sustava e-učenja su *inteligentni tutorski sustavi* (ITS) koji su namijenjeni potpori i poboljšanju procesa učenja i poučavanja u odabranom područnom znanju, poštujući individualnost učenika kao što je to slučaj kod tradicionalnog poučavanja jedan-na-jedan. U ovom poglavlju je opisan model *Tutor-Expert System* (TEx-Sys) [STAN1997] koji predstavlja model inteligentne hipermedijske autorske ljuske za izgradnju inteligentnih tutorskih sustava u po volji odabranom područnom znanju. Prva implementacija ovog modela je „on-site“ TEx-Sys (1992-2001), nakon čega je uslijedila inteligentna autorska ljuska temeljena na računalnim mrežama Distributed Tutor-Expert System (1999-2003, DTEEx-Sys) [ROSI2000] te, naposljetku, sustav temeljen na Web servisima *Extended Tutor-Expert System* (2003-2005, xTEx-Sys). Na kraju ovog poglavlja je dana kronika testiranja obavljenih nad spomenutim modelom.

U prvom dijelu trećeg poglavlja objašnjeni su neki od osnovnih statističkih pojmova kao što su varijabla, populacija, uzorak, distribucija, nul-hipoteza, statistička značajnost. Opisane su

karakteristike distribucije frekvencija, položaj pojedinog rezultata u grupi, t-test, korelacija, te na kraju, grafički način prikazivanja rezultata. U drugom dijelu ovog poglavlja objašnjavaju se različiti pristupi vrednovanju sustava e-učenja, daje se pregled i klasifikacija metoda vrednovanja, objašnjavaju se postupci i instrumenti prikupljanja podataka, kao i karakteristike tih instrumenata, te se posebno objašnjava eksperiment kao metoda vrednovanja. U trećem dijelu je objašnjeno što je to veličina učinka i kako se može izračunati te je naglašena njena primjena kod meta-analiza. Na kraju je ukratko opisano nekoliko primjera vrednovanja sustava e-učenja.

U četvrtom poglavlju se, na temelju spoznaja iz literature, pristupa definiranju vlastite metodologije za vrednovanje učinka procesa učenja i poučavanja u sustavima e-učenja. U tu svrhu potrebno je odrediti vrstu veličine učinka koja će se računati, zatim definirati strukturu eksperimenta, postaviti nul-hipoteze, pokazati kako se provodi ispitivanje statističke značajnosti, te kako analizirati i interpretirati rezultate. Na kraju će se pokazati kako se definiranom metodologijom pristupilo vrednovanju učinka konkretnog sustava koji pripada klasi inteligentnih tutorskih sustava, a samim time i sustavima e-učenja.

U petom poglavlju je opisana funkcionalnost i arhitektura *sustava za automatsko vrednovanje učinka sustava e-učenja*. Programski alat za automatsko vrednovanje učinka sustava e-učenja nazvan je EVEDIN. EVEDIN izračunava veličinu učinka na temelju informacija o postignućima učenika jer se znanje učenika utvrđuje primjenom nekoliko „online“ testova: inicijalnog testa, testova provjere te završnog testa. Rezultati spomenutih testova prolaze kroz EVEDIN-ov mehanizam za statističku analizu, te kao rezultat EVEDIN daje izračunatu veličinu učinka sustava e-učenja.

2. E-UČENJE I INTELIGENTNI TUTORSKI SUSTAVI

E-učenje predstavlja presjek svijeta informacijske i komunikacijske tehnologije (eng. information and communication technology, ICT) i svijeta obrazovanja [STAN2004a]. U usporedbi s tradicionalnom nastavom u razredu koja u središte stavlja učitelja i njegovu kontrolu nad razredom, nastavnim sadržajem i procesom učenja i poučavanja, e-učenje u središte stavlja učenika kojem omogućava interaktivno učenje vlastitim ritmom, u jednostavnom, fleksibilnom, distribuiranom okruženju za učenje [KHAN2001]. Najčešće korištena definicija e-učenja jest da je e-učenje skup aplikacija i procesa, kao što su učenje temeljeno na Web-u (eng. Web-based learning), učenje temeljeno na računalu (eng. computer-based learning), virtualni razredi (eng. virtual classrooms) i digitalna suradnja (eng. digital collaboration), koji omogućavaju pristup nastavnim sadržajima pomoću različitih elektroničkih medija (CD-ROM, Internet, intranet, extranet, audio i video, satelit, itd.) [ASTD2001].

Sustavi e-učenja, stoga, omogućavaju pristup elektroničkim izvorima za učenje bilo gdje, bilo kada i bilo kome [ALBE2001]. *Inteligentni sustavi e-učenja* imaju sposobnost prikladnog ponašanja u neizvjesnim situacijama koje se javljaju u procesu učenja i poučavanja. Posebna klasa inteligentnih sustava e-učenja su *inteligentni tutorski sustavi* (ITS).

U ovim vremenima informacijske i komunikacijske tehnologije i Interneta, ITS-ovi ponovno dobivaju na značenju jer su istraživači uočili važnost i vezu pedagogijske paradigme ovih sustava s “2-sigma” problemom karakterističnim za učinak individualnog rada u usporedbi s grupnim radom u procesu učenja i poučavanja [STAN2004a].

2.1. INTELIGENTNI TUTORSKI SUSTAVI I INTELIGENTNE AUTORSKE LJUSKE

Kroz povijest računala mnogo se pažnje posvetilo ideji o uporabi računala kao inteligentnog suradnika, odnosno privatnog učitelja, koji će nam otkriti i objasniti zamršena područna znanja. Složenost tog zadatka je uočena nakon što se za konačni cilj u realizaciji te ideje uzelo *inteligentno poučavanje* (eng. intelligent tutoring). *Umjetna inteligencija* (eng. artificial intelligence, AI), kao područje koje spaja računala i inteligentno ponašanje, javlja se krajem 50-tih i početkom 60-tih godina prošlog stoljeća sa začetnicima Alanom Turing-om, Marvinom Minsky-m, Johnom McCarthy-iem i Allenom Newell-om [URBA1996]. AI je u suštini orijentirana na prikaz znanja (eng. knowledge representation), razumijevanje prirodnog jezika

(eng. natural language understanding), te rješavanje problema (eng. problem solving), a sve to je jednako važno za razvoj koncepta inteligentnog poučavanja [BECK1996].

Šezdesetih godina prošlog stoljeća znanstvenici su razvili brojne sustave za realizaciju *nastave pomoću računala* (eng. Computer Assisted Instructional, CAI) [UHRL1969]. U uporabi je i termin *vježbanje temeljeno na računalu* (eng. Computer Based Training, CBT) koji obuhvaća sve sustave koji omogućavaju „neinteligentni“ način učenja i poučavanja. Spomenuti CAI sustavi bi učenika upoznali s problemom, primili i zapamtili učenikov odgovor. Oni se nisu pretjerano bavili problematikom kako učenici uče, već su smatrali da ako je učeniku dana informacija, on će je usvojiti. Veći problem tada su bili tehnički izazovi u programiranju na skupim i glomaznim „mainframe“ računalima.

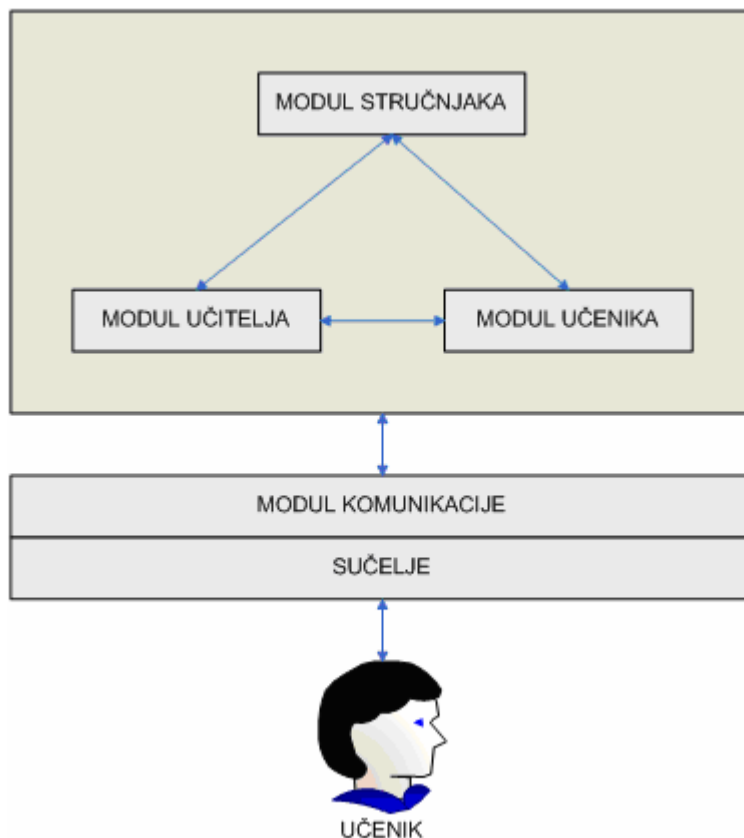
Kasnih 60-tih i ranih 70-tih godina prošlog stoljeća znanstvenici su prešli s pukog prezentiranja zadataka učenicima na uvažavanje učenika, kao faktora koji utječe na proces učenja i poučavanja (Suppes, 1967, prema [UHRL1969]). Neki od razvijenih sustava mijenjali su način prikazivanja nastavnih sadržaja ovisno o učenikovim odgovorima. Programeri su stoga morali unaprijed znati sve moguće učenikove odgovore i odlučiti koju će informaciju prezentirati sustav. Ovo predstavlja početak modeliranja učenika, iako je u ovom razdoblju promatrano samo učenikovo ponašanje, a nije modelirano znanje učenika. Kao jedan od prvih istraživača u području učenja uz pomoć računala i od računala, Carbonell je spojio CAI i AI implementirajući sustav SCHOLAR, koji se smatra jednim od prvih ITS-ova [CARB1970].

Godine 1982., Sleeman i Brown su u [SLEE1982] dali pregled posljednjih dostignuća u području nastave potpomognute računalima i prvi put su uporabili složenicu *inteligentni tutorski sustavi* (eng. Intelligent Tutoring Systems, ITS) da bi opisali nove sustave i da bi ih mogli razlikovati od starih CAI sustava. Učenje se kod ovih novih sustava vrši na principu „učim kroz rad“ (eng. learning-by-doing). Sleeman i Brown su klasificirali postojeće ITS-ove u četiri klase: (1) nadziratelji rješavanja problema (eng. problem-solving monitors), (2) poučavatelji (eng. coaches), (3) laboratorijski instruktori (eng. laboratory instructors) te (4) savjetnici (eng. consultants). Rickel u [RICK1989] navodi da su „inteligentni“ CAI sustavi, odnosno ICAI sustavi, oni CAI sustavi koji imaju područno znanje stručnjaka i znanje pedagoga, te ih smatra jednakima ITS-ovima.

Dakle, inteligentni tutorski sustavi su računalni sustavi namijenjeni potpori i poboljšanju procesa učenja i poučavanja u odabranom područnom znanju, poštujući individualnost učenika, kao što je to slučaj kod tradicionalnog poučavanja jedan-na-jedan [BURN1988]. ITS-ovi svojim modelima područnog znanja određuju što će se poučavati, a strategijama poučavanja određuju kako će se poučavati ([WENG1987], [OHLS1987], [SHUT1996]). To su sustavi računalom kontroliranog poučavanja u koje je integrirana umjetna inteligencija, odnosno to su sustavi koji moraju imati razvijenu komponentu „razmišljanja“, tj. rješavanja problema onako kako bi to učinili ljudi, a uz to su i tutorski sustavi s razvijenom pedagoškom paradigmom koja se temelji na činjenici da ITS-ovi, da bi ispunili svoju funkciju, moraju osposobiti učenika za rješavanje problema u danom područnom znanju.

Temeljne odrednice ovih sustava su: (1) znanje koje sustav ima o područnom znanju; (2) principi pomoću kojih sustav poučava i metode pomoću kojih primjenjuje te principe; (3) metode i tehnike za modeliranje učenika tijekom stjecanja znanja i umijeća. Ove temeljne odrednice uvjetuju sljedeću strukturu inteligentnog tutorskog sustava: (1) baza područnog znanja (modul stručnjaka); (2) modul za vođenje poduke (modul učitelja); (3) modul za obuhvat i procjenu

znanja učenika (modul učenika); (4) modul interakcije učenika, učitelja i područnog znanja (modul komunikacije). Tipična arhitektura inteligentnih tutorskih sustava je prikazana na Slika 2. 1.



Slika 2. 1. Komponente ITS-a (prema [BURN1988])

Modul stručnjaka izvršava dvije važne funkcije. Prvo, on služi kao izvor znanja koje se predstavlja učeniku. Drugo, on predstavlja standard za ocjenjivanje učenikovih performansi. Modul stručnjaka je okosnica svakog inteligentnog tutorskog sustava jer sadrži bazu područnog znanja, a znanje predstavlja ključ za inteligentno ponašanje [ANDE1988]. Znanstvenici su tijekom 80-tih godina prošlog stoljeća počeli razmišljati o prikazu znanja u ITS-ovima. Prikaz znanja je prevođenje, kodiranje ili formaliziranje znanja u format pogodan za rad računala. Brojne su tehnike za prikaz znanja, a koja će se u konkretnom slučaju upotrijebiti ovisi o vrsti znanja koje prevladava u promatranom područnom znanju. U načelu postoje tri modela prema kojima se u modul stručnjaka integrira znanje [ANDE1988]: (1) model crne kutije (eng. black box model) kod kojeg su vidljivi samo ulazno – izlazno podaci i ponašanje, ali ne i unutrašnji procesi obrade koji uzrokuju takvo ponašanje, (2) model prozirne kutije (eng. glass box model) koji, za razliku od modela crne kutije, objašnjava način svog zaključivanja, te (3) spoznajni model (eng. cognitive model) koji oponaša način rješavanja problema „živog“ stručnjaka.

Funkcija *modula učenika* je u bilježenju učenikova razumijevanja ili nerazumijevanja područnog znanja, tj. on bilježi učenikovo napredovanje. Sleeman i Brown u [SLEE1982] prvi put spominju model učenika kao pojam koji predstavlja apstraktni prikaz učenika u sustavu. Oni su klasificirali modele učenika u tri klase: (1) model prekrivanja (eng. overlay) – promatra učenikovo znanje kao podskup znanja stručnjaka, (2) model razlike (eng. differential) – promatraju se razlike između znanja učenika i znanja stručnjaka, te (3) model zbrke (eng. perturbation) – promatra se učenikovo pogrešno shvaćanje (eng. misconceptions) u usporedbi sa znanjem stručnjaka. Prvi pokušaji modeliranja učenika zasnivaju se na „buggy“ modelu kojeg su predložili Brown i Burton [BROW1978], a Burton ga je implementirao u sustav DEBUGGY [BURT1982].

Modul učitelja je usko povezan s modulom učenika te odlučuje o načinu poučavanja svakog učenika ponaosob. Modul učitelja, kao sastavni dio ITS-a koji preuzima ulogu stvarnih učitelja, mora zadovoljiti trima karakteristikama: (1) tutor mora imati kontrolu nad izborom i redoslijedom nastavnih sadržaja, (2) tutor mora odgovarati na učenikova pitanja., te (3) tutor mora prepoznati kada učeniku treba pomoć i kakva pomoć mu treba. Te zadatke modul učitelja rješava pomoću tzv. scenarija učenja. Scenarij učenja je situacija u kojoj se odvija učenje, tj. načini na koje ITS poučava učenika [RICK1989]:

1. Računalo trener (eng. computer coach) - Na vaš zahtjev „inteligentni trener“ gleda preko vašeg ramena dok vi pokušavate riješiti zadatak nudeći vam povremene, ali nenametljive savjete. Učitelj je na raspolaganju onda kada je to učeniku istinski potrebno.
2. Učenje kroz igru (eng. gaming environment) - Kombinira treniranje i učenje putem otkrivanja. Učenik ovladava umijećima primjenjujući ih kroz igru, a može ih i otkriti putem njihovog pozitivnog utjecaja na uspješnost u igri. Trener treba voditi učenika i prepoznati da li je njegov odgovor ili postupak optimalan, te zaključiti koji su učenikovi nedostaci.
3. Dijalog podijeljene inicijative (eng. mixed initiative method) - Učeniku se postavlja problem i on mora upotrijebiti novo znanje. Učitelj ističe nedostatke u učenikovu znanju te ga „dovodi“ u situacije otkrivanja pogrešnih zamisli. Tutor mora izraziti svoje znanje odgovarajući na pitanja.
4. Sokratov dijalog (eng. Socratic teaching method) – Slično dijalogu podijeljene inicijative. Sokrat je smatrao da učenik ne može naučiti pasivno čitajući, već mora sudjelovati u rješavanju problema.
5. Razgovijetni stručnjak (eng. articulate expert paradigm) - ITS je projektiran tako da pouči učenika u pronalaženju pogreški omogućujući mu da promatra kako tutor rješava probleme prije nego li sam dobije zadatak. Učeniku je dopušteno uvesti pogrešku i ITS traži pogrešku postupno objašnjavajući svoj način traženja pogreške. Učenik zatim rješava zadatak, dok ITS promatra i povremeno pruža pomoć i savjet.
6. Simulacija procesa (eng. simulation based instruction) - Učenik bi trebao znati kada i za što su pojedini postupci prikladni. ITS mora stoga prikazati proceduralna umijeća na različitim razinama apstrakcije. U poučavanju koje se temelji na simulaciji nije dovoljno samo predstaviti simulaciju kao crnu kutiju s kojom se učenik može igrati, već informacija mora biti predstavljena učeniku tako da mu omogući izgradnju pravilnog spoznajnog konstrukta.
7. Učenje pomoću otkrivanja (eng. discovery learning) - Učenik ima potpunu kontrolu i može dopustiti svojoj znatiželji i interesima vođenje i tako će usvojiti principe koje je samostalno otkrio.

Modul komunikacije kontrolira interakciju između ITS-a i učenika. Ono predstavlja sučelje učenika i okruženje nastavnog procesa odnosno interakciju učenik-učitelj-znanje. Modul komunikacije upravlja interakcijom sustava s učenikom koja se ostvaruje preko dijaloga i grafičkog korisničkog sučelja. Ova komponenta nije toliko izučena kao i ostale, ali postoje značajni pomoci na ovom području, pogotovo na području interakcije računala i čovjeka (eng. Human Computer Interaction, HCI) [CARD1983].

Glavni problemi kod razvoja ITS-ova su velika cijena i velik utrošak vremena. Da bi se prevladali spomenuti problemi pristupa se izgradnji određenog ITS-a uz pomoć *autorskih ljuški* ili, kako neki autori kažu, autorskih sustava, tj. autorskih alata ([MURR1996], [AAAI1997], [ITSW1996]). Autorske ljuške omogućavaju osobama koje nemaju programerska znanja i vještine, izgradnju ITS-ova u određenom područnom znanju.

Murray je u [MURR2003] dao kategorizaciju autorskih alata (Tablica 2. 1.):

1. Sustavi za nizanje i planiranje nastavnog sadržaja - Organiziraju elemente nastavnog sadržaja kao hijerarhiju kolegija, modula, lekcija, te prezentacija koji su međusobno povezani vezama preduvjeta i posjedovanja. Nastavni sadržaj se posprema kao zatvoren i ograničen tekst s elementima grafike. Nizanje sadržaja se određuje u odnosu na postignuće učenika. Ovakvi sustavi su pogodni za izgradnju tutorskih sustava u kojima se uči i poučava deklarativno područno znanja. Nisu pedagoški snažni kako bi se pomoću njih vodio proces učenja i poučavanja proceduralnog znanja ili vještina rješavanja problema.
2. Sustavi sa strategijama poučavanja - Karakteristika ovih sustava je što imaju sposobnost izvođenja različitih strategija učenja gdje se u danoj situaciji izabire najprikladnija. Slični su sustavima s nizanjem i planiranjem nastavnog sadržaja. Osim što nižu elemente nastavnog sadržaja, također daju objašnjenja, sažetke i primjere učeniku.
3. Simulacije i uvježbavanje rada na uređajima - Učeniku se predočava dio opreme i od njega se traži da identificira njegove komponente, izvrši operativne korake, izvede korake održavanja ili da ustanovi nepravilno ponašanje, te popravi ili zamijeni određene dijelove. Znanje stručnjaka o smještaju komponenata i operativnim skriptama je doslovno preneseno na model baš kao i nadgledanje izvršavanja operacija. Učenici koji uče na ovakvim sustavima se uče kroz rad. Ovakvi tutorski sustavi su zapravo okruženja učenja kod kojih se uvježbava vještina učenika.
4. Sustavi područnog znanja - Ovakvi tutorski sustavi prate ponašanje učenika i grade spoznajne modele učenikova znanja koji se mogu usporediti s modelom stručnjaka. Učenici koji koriste ove sustave obično rješavaju probleme unutar ciljanog prostora i dobivaju povratne informacije u slučaju kada se njihovo znanje razlikuje od modela stručnjaka. Za razliku od svih prethodno navedenih kategorija tutorskih sustava, ovakvi sustavi imaju profinjene modele procjene i u slučaju da učenik „zapne“, on može pitati sustav da izvrši sljedeći korak ili da riješi cijeli problem.
5. Sustavi s višestrukim tipovima znanja - Klasificiraju znanje i zadatke kao odvojene kategorije, te propisuju metode učenja i poučavanja za svaku od njih. Oni se ograničavaju na tipove znanja koji se mogu jednostavno definirati, kao što su činjenice, koncepti i procedure. Variranjem tipova znanja, variraju i metode poučavanja, koje se u slučaju nizanja nastavnog sadržaja i vježbi definiraju odvojeno za svaki tip znanja.

6. Sustavi specijalne namjene - Specijalizirani su za određene zadatke ili područja. Fokusirani su na više određene i manje općenite zadatke.
7. Sustavi s inteligentnom i prilagodljivom hipermedijom - Kao i kod većine sustava zasnovanih na Web-u, nivo interakcije sustava i učenika je vrlo nizak. Za razliku od sustava iz drugih kategorija, ovi sustavi moraju koristiti hiperveze za navigaciju kroz nastavni sadržaj. Veze vidljive učeniku se mogu filtrirati i sortirati ovisno o modelu učenika.

Tablica 2. 1. Kategorije ITS autorskih alata (prema [MURR2003])

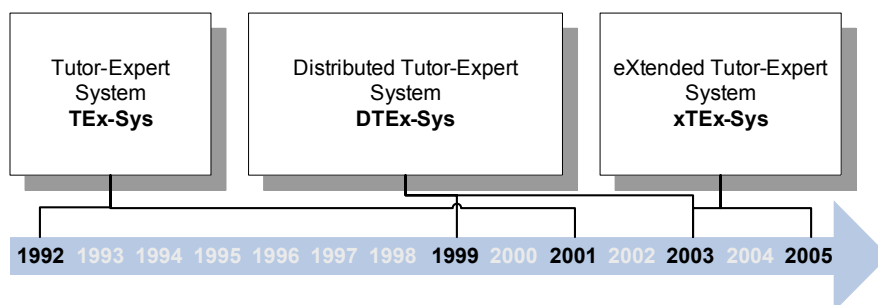
KATEGORIJA	PRIMJER SUSTAVA
Nizanje i planiranje nastavnog sadržaja (eng. Curriculum Sequencing and Planning)	DOCENT [WINN1988], IDE [RUSS1988], ISD Expert [MERI2001]
Strategije poučavanja (eng. Tutoring Strategies)	Eon [MURR1998], GTE [VANM1998], REDEEM [AINS2003]
Simulacije i uvježbavanje rada na uređajima (eng. Device Simulation and Equipment Training)	RIDES [MUNR1997], SIMQUEST [VANJ1996], XAIDA [REDF1996]
Sustavi područnog znanja (eng. Domain Expert System)	Demonstr8 [BLES1997], D3 Trainer [SCHE1999], Training Express [CLAN1988]
Višestruki tipovi znanja (eng. Multiple Knowledge Types)	CREAM-Tools [NKAM1996], DNA [SHUT2000], IRIS [ARRU1997]
Specijalna namjena (eng. Special Purpose)	IDLE-Tool/IMap [QIUR2003], LAT [SPAR1999]
Inteligentna / prilagodljiva hipermedija (eng. Intelligent/adaptive Hypermedia)	CALAT [KIYA1997], GETMAS [WONG1997], InterBook [BRUS1996]

Obrazovanje na daljinu (eng. distance education) je se proces učenja i poučavanja u kojem su osobe koje sudjeluju u tom procesu fizički odvojene pa je zbog toga potrebno osigurati komunikacijska sredstva kojima se provodi prijenos nastavnog sadržaja i obavlja interakcija svih sudionika. Inteligentni tutorski sustavi kojima se pristupa korištenjem standardnih Web pretraživača nazivaju se *Web orijentiranim inteligentnim tutorskim sustavima* (eng. Web oriented intelligent tutoring systems). Oni omogućavaju ugodno, korisniku orijentirano, distribuirano okruženje za učenje i poučavanje [BRUS1997].

Više od četrdeset godina razvoja dijeli današnje inteligentne tutorske sustave od prvih pristupa i pokušaja razvoja. Mnoštvo sustava je razvijeno, implementirano i testirano u okviru nastavnog procesa u školi i na fakultetima, te je postignut je jednoglasan dogovor o arhitekturi ITS-ova. U suglasju s postignutim spoznajama i dogovorima, a u svrhu realizacije što kvalitetnijeg obrazovanja, Slavomir Stankov [STAN1997] je prije desetak godina pristupio razvoju ideje o inteligentnom poučavanju temeljenom na kibernetičkom modelu sustava. Definirani i implementirani model inteligentnog tutora, kao i sustavi temeljeni na tom modelu, opisani su u sljedećem poglavlju.

2.2. MODEL TEx-Sys I SUSTAV xTEx-Sys

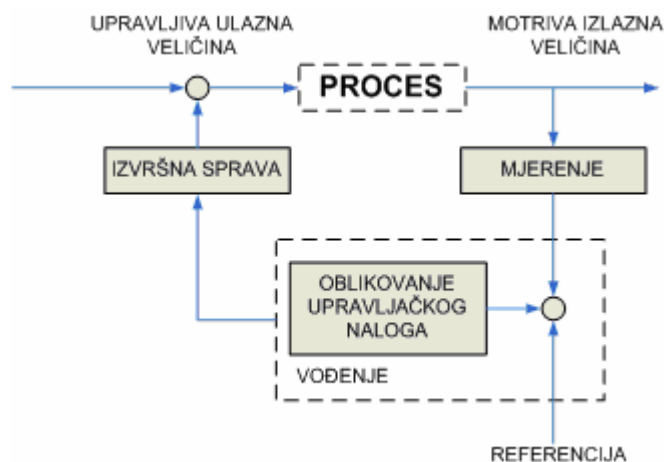
Tutor-Expert System (TEx-Sys) [STAN1997] je model inteligentne hipermedijske autorske ljuske za izgradnju inteligentnih tutorskih sustava u po volji odabranom područnom znanju. Prva implementacija ovog modela je „on-site“ TEx-Sys (1992-2001), nakon čega je uslijedila inteligentna autorska ljuska temeljena na računalnim mrežama Distributed Tutor-Expert System (1999-2003, DTEEx-Sys) [ROSI2000] te, naposljetku, sustav temeljen na Web servisima Extended Tutor-Expert System (2003-2005, xTEx-Sys) [STAN2003].



Slika 2. 2. Verzije modela TEx-Sys

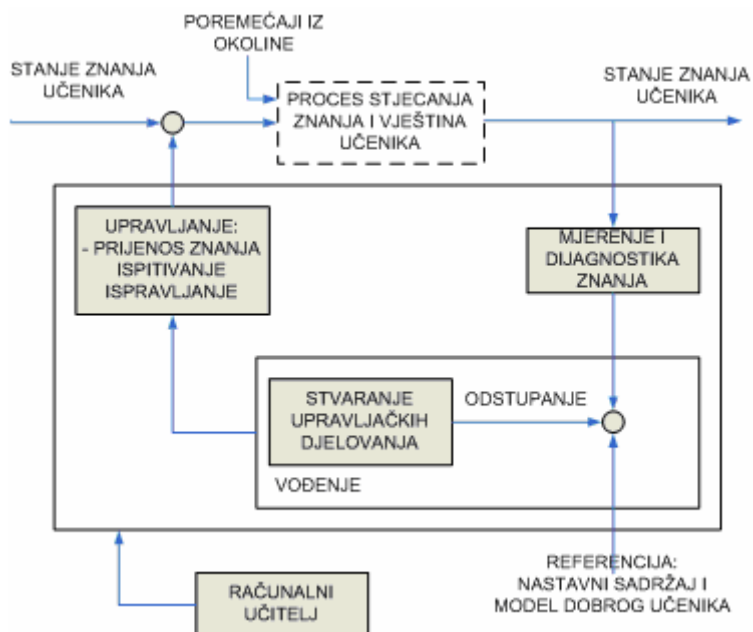
2.2.1. MODEL TEx-Sys

Model TEx-Sys se temelji na *kibernetičkom modelu sustava* (Slika 2. 3.) ([BOŽI2001], [PASK1965], [WIEN1948]) u kojem djeluju tri informacijske strukture: učenik, učitelj i nastavni sadržaji. U takvom se sustavu vodi proces stjecanja znanja i vještina učenika.



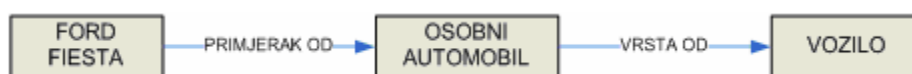
Slika 2. 3. Kibernetički model sustava [STAN2001]

Sustav TEx-Sys je strukturiran u suglasju s kibernetičkim modelom (Slika 2. 4.), pa je u vezi s tim stjecanje znanja i vještina učenika vođeni proces, a referenca nastavni sadržaj i model "dobrog" učenika [STAN2001]. Stanje znanja učenika upravljiva je ulazna i izlazna veličina. Upravljačka funkcija sustava temelji se na mjerenju i dijagnosticiranju znanja učenika, određivanju razlike stvarnog znanja učenika i referentnog modela, te na ocjenjivanju znanja učenika s preporukom za daljnji rad. Računalni učitelj, kao zamjena "živom" učitelju, motri, tj. mjeri i provodi dijagnostiku znanja učenika, određuje odstupanja učenikova stvarnog znanja od referentnog modela, oblikuje upravljačko djelovanje i prijenos novog znanja te ispravljanje (upućivanje, pružanje dodatnih informacija, dopunu stečenog znanja). Pretpostavljeno je da poremećaj djeluje samo neposredno na učenika.



Slika 2. 4. Model vođenja procesa stjecanja znanja i vještina učenika u inteligentnom tutorskom sustavu [STAN2001]

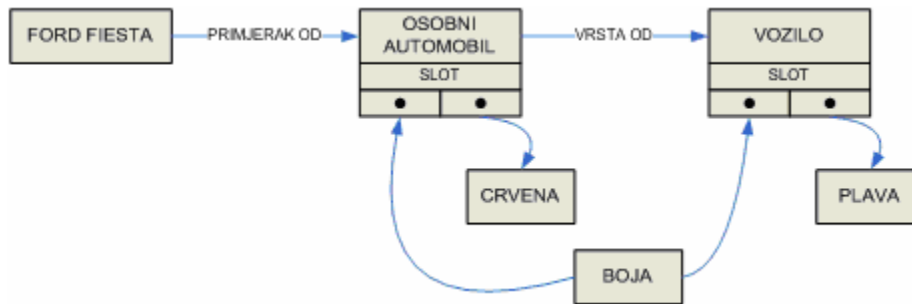
U modelu TEx-Sys znanje je prikazano pomoću *semantičke mreže s okvirima* čiji su osnovni elementi čvorovi (eng. nodes) i veze (eng. links). Čvorovi se koriste za prikaz objekata područnog znanja, a veze za prikaz odnosa između tih objekata (Slika 2. 5.). Veza je u biti tvrdnja da je za izvjesni objekt nešto istina u odnosu na drugi objekt.



Slika 2. 5. Semantička mreža s tri čvora i dvije veze

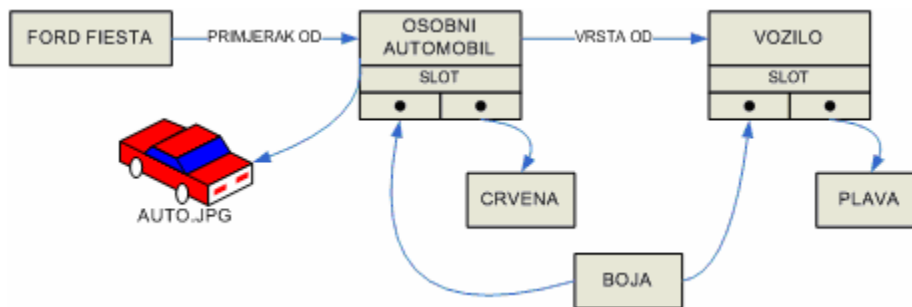
Model omogućava uporabu svojstava i okvira koji se sastoje od atributa i pripadajućih im vrijednosti. Koriste se za iskazivanje znanja o svojstvima objekata u danom područnom znanju dodavanjem novih čvorova i pridruživanjem relacije sa značenjem svojstva (Slika 2. 6.). Primjenjuje se zato shema Minsky-og [TOUR1992] u kojoj je znanje enkodirano u tzv. okvirima

koji su ukomponirani u mrežu s mogućnošću pretraživanja. Sve skupa se zato naziva sustav temeljen na okvirima gdje se okvir se obično pridružuje čvoru. Objekt ima proizvoljni broj "otvora" (eng. slot), pomoću kojih mu se dodaju atributi <SLOT> i pripadne im vrijednosti <FILLER>. Ovakve sustave s okvirima Cherniak i McDermot [CHER1985] nazivaju "slot-and-filler".



Slika 2. 6. Semantička mreža s okvirima

Čvorovi u bazi znanja pored naziva i veza prema ostalim čvorovima mogu imati i jedan od strukturnih atributa i to: tekstualni opis, identifikacijski simbol (eng. icon), sliku i animaciju (pokretnu sliku i zvuk) kao i URL adresu (Slika 2. 7.).



Slika 2. 7. Semantička mreža s okvirima i strukturnim atributom tipa slika

U modelu TEx-Sys jedna od funkcija kojoj je posvećena posebna pozornost je modeliranje učenika pri učenju i poučavanju odabranog područnog znanja, što podrazumijeva izgradnju modela učenika te dijagnosticiranje znanja kao osnovica za ocjenjivanje znanja. Ocjenjivanje znanja provodi se prekrivanjem znanja učenika sa znanjem učitelja (eng. overlay technique), te pomoću *dinamičkog kviza*. Dinamički su oni kvizovi za testiranje i ocjenjivanje znanja kod kojih se pitanja generiraju pomoću, na računalu utemeljenom i oblikovanom, algoritmu testiranja i ocjenjivanja, a pitanja se generiraju uz pomoć generatora slučajnih brojeva pri zahtjevu za rješavanjem kviza.

U dinamičkom kvizu se učeniku distribuira skup pitanja uz pridružene ponuđene odgovore koji mogu biti točni ili netočni. Test se rješava obilježavanjem odgovora za koje učenik smatra da su točni. Za svako pitanje mogu biti svi ponuđeni odgovori točni, može biti nekoliko točnih odgovora, može biti samo jedan točan odgovor, ali mogu i svi ponuđeni odgovori biti netočni. Hipermedijski pristup u izgradnji baze znanja omogućava da pitanja u kvizu popratimo i

određenim multimedijским strukturnim atributima, da bi učeniku pomogli pri rješavanju problema. Nakon rješavanja kviza ocjenjuju se odgovori učenika i po potrebi upućuje prema pojmovima područnog znanja za koje se utvrdilo da nije dobro naučio.

Pitanja su u kvizu grupirana su u tri težinske kategorije, a u svakoj kategoriji su po četiri težinske forme. Ukupno to je dvanaest formi pitanja koje učenik dobiva tijekom testiranja svojeg znanja. Ako učenik odgovori točno na oba pitanja, u sljedećoj seriji dobiti će pitanja iz teže skupine, odnosno ako odgovori netočno na oba pitanja sljedeća će biti iz lakše skupine. Time se sustav prilagođava iskazanom učenikovom znanju. Za ocjenjivanje učenika razvijen je bodovni kriterij povezan s težinskim kategorijama pitanja u kvizu (Tablica 2. 2).

Tablica 2. 2. Bodovni kriteriji

Bodovna vrijednost	
Težinska kategorija pitanja	Broj bodova
Prva	1
Druga	2
Treća	3

Ostvareni bodovi se zbrajaju nakon svake serije pitanja, te se na kraju dobiva ukupan rezultat. Ocjena je predstavljena rangom kao odnos konačnog zbroja ostvarenih bodova i maksimalnog broja bodova. Rang je kvantificiran u rasponu od 30 do 91 bod, a određuje se temeljem izraza:

$$\text{rang} = \frac{\text{ostvareni broj bodova}}{\text{maksimalan broj bodova}} \cdot 100$$

Tablica 2. 3. Ocjena testa

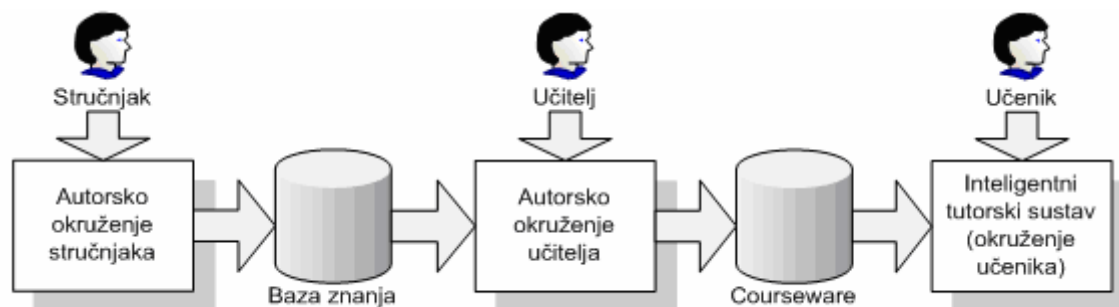
Bodovi kriterij ranga	
rang	Ocjena
rang ≤ 30	Nedovoljan
31 ≤ rang ≤ 50	Dovoljan
51 ≤ rang ≤ 70	Dobar
71 ≤ rang ≤ 90	Vrlo dobar
91 ≤ rang	Odličan

Generira se lista svih pitanja s informacijom o točnosti učenikovog odgovora te se omogućava pristup čvorovima iz onih pitanja na koje nije dan potpuno točan odgovor. Ostvareni rezultat testiranja učenika smješta se u bazu učenika, da bi se sustav mogao prilagoditi njegovom znanju kod sljedećeg pristupa testiranju ili daljnjem učenju.

2.2.2. xTEx-Sys

Zamisao Web orijentirane inteligentne autorske ljuske xTEx-Sys se zasniva na udruženim i proširenim funkcionalnostima sustava TEx-Sys i DTEEx-Sys. Web orijentirana inteligentna autorska ljuska xTEx-Sys predstavlja specijalizirano okruženje prilagođeno potrebama stručnjaka

za razvoj baza područnog znanja, učiteljima za izgradnju courseware-a i učenicima za učenje, poučavanje i testiranje znanja. Struktura ljuške je prikazana na Slika 2. 8.



Slika 2. 8. Struktura autorske ljuške

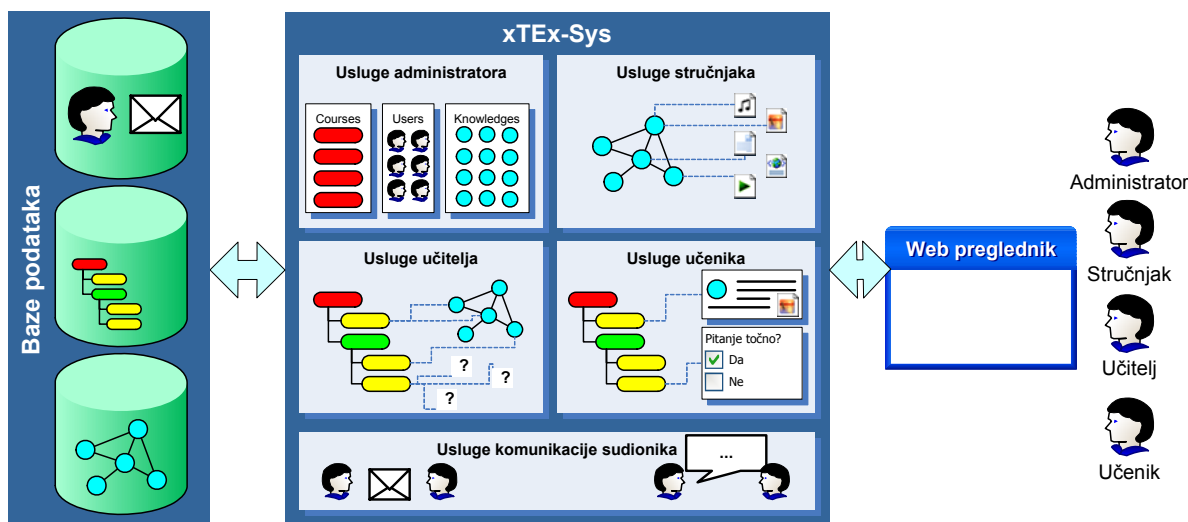
Sudionici ovog programskog sustava su:

- Učenici radi stjecanja znanja i vještina. Uče i poučavaju se uz pomoć izabrane baze znanja za odabrano područno znanje.
- Učitelji radi didaktičkog oblikovanja nastavnih sadržaja. Prikupljaju i organiziraju način realizacije nastavnih sadržaja za područno znanje.
- Stručnjaci za područna znanja radi izgradnje baza znanja.
- Administrator sustava radi nadzora funkcija sustava, korisnika sustava kao i načina korištenja sustava.

Temeljne funkcije autorske ljuške su:

- oblikovanje baza znanja za različita područna znanja,
- oblikovanje i pristup nastavnim sadržajima temeljenim na bazama područnih znanja,
- učenje i poučavanje te testiranje i ocjenjivanje znanja
- nadzor sustava.

Web orijentirana inteligentna autorska ljuška xTeX-Sys je zasnovana na trorednoj arhitekturi (Slika 2. 9.) sa redom korisničkog sučelja, redom aplikacijske logike i redom podataka. Korisničko sučelje je implementirano tako da omogući dobru komunikaciju svih sudionika i programskog sustava. Red aplikacijske logike i red podataka implementiraju funkcionalnosti sustava i omogućavaju raspolaganje sa svim potrebnim podacima o sudionicima sustava, bazama znanja i bazama oblikovanih nastavnih sadržaja namijenjenih učeniku za učenje, poučavanje i vrednovanje znanja. Na Slika 2. 10. prikazana je stranica za prijavu na sustav xTeX-Sys.



Slika 2. 9. Arhitektura sustava xTEx-Sys

xTEx-Sys Web orijentirana inteligentna autorska ljuska, verzija 1.2

[Odjava](#)

Tehnolojski projekt TP-02/0177-01

Uz primjenu objektno orijentirane metodologije programskog inženjerstva kao i po načelima daljinskog učenja pristupa se implementaciji Web orijentirane autorske ljuske za izgradnju inteligentnih tutorskih sustava namijenjenih vođenju procesa učenja i poučavanja u po volji odabranom područnom znanju. Temeljne funkcije autorske ljuske bit će: oblikovanje i pristup nastavnim sadržajima, razvoj i pristup bazama znanja s područnim znanjima potrebnim za oblikovanje odnosnih nastavnih sadržaja, testiranje i ocjenjivanje znanja te nadzor sustava.

Prijava na sustav

korisničko ime:

zaporka:

Napomena: Za dobivanje korisničke oznake i zaporke kliknite [ovdje](#).

0 projektu Zadnja izmjena: 8.3.2005 Pomoć














Slika 2. 10. Prijava na sustav xTEx-Sys

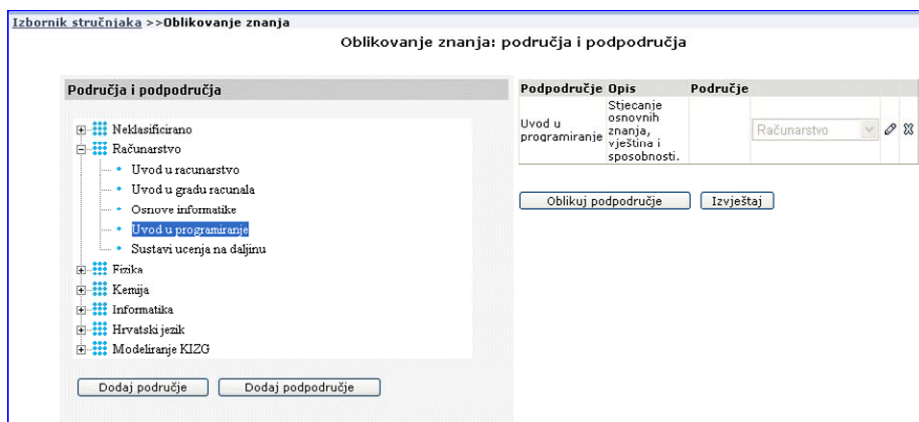
2.2.2.1 OBLIKOVANJE BAZA PODRUČNOG ZNANJA

Stručnjak aktivnosti na oblikovanju baze znanja realizira u autorskom okruženju za stručnjaka područnog znanja radi oblikovanja baza područnog znanja, primjenom specijaliziranih autorskih alata i uz pomoć posebno definirane ontologije za prikaz i strukturiranje znanja.

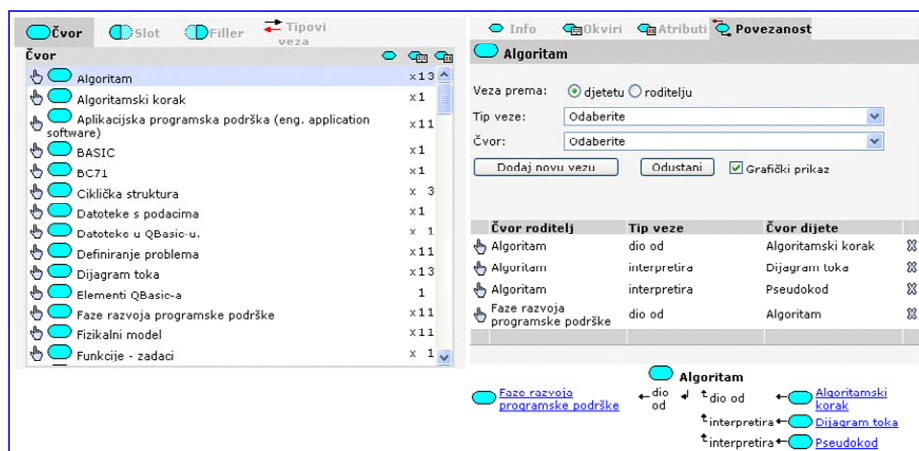
Oblikovanje baza područnog znanja i oblikovanje nastavnih sadržaja ostvareno je s tehnikom za prikaz znanja koja je u sustavu xTEx-Sys ostvarena primjenom semantičkih mreža s okvirima. Globalna razina znanja iskazana je s nazivom *područja*, a dalje svako područje ima pripadna *podpodručja* (Slika 2. 11.), te konačno podpodručja su grupirana u *elementarne objekte znanja* zastupljene sa čvorovima znanja u semantičkoj mreži (Slika 2. 12.). Izgrađena baza područnog znanja je ulazna komponenta procesa oblikovanja nastavnog sadržaja i ocjenjivanja znanja.

Tablica 2. 4. Značenje simbola u autorskom okruženju za oblikovanje baze područnog znanja

Stručnjak	
Oznaka	Značenje
	Čvor
	Slot
	Filler
	Tipovi veza
	Povezanost
	Info
	Okvir
	Atribut
	Video zapis
	Aplikacija
	Audio zapis
	Slika
	URL



Slika 2. 11. Forma za izbor postojećeg područja i podpodručja




Slika 2. 12. Grafički prikaz povezanosti čvora znanja s čvorovima tipa roditelj i dijete

2.2.2.2 OBLIKOVANJE NASTAVNOG SADRŽAJA I OCJENJIVANJE ZNANJA UČENIKA

Učitelj svoje aktivnosti obavlja u autorskom okruženju za učitelja radi oblikovanja nastavnih sadržaja, primjenom specijaliziranih autorskih alata. Oblikovanje nastavnog sadržaja na najvišoj razini provodi se za grupu učenika, a grupa dalje sadrži niz kolegija. Kolegij je strukturiran u obliku *courseware*-a koji predstavlja didaktički oblikovan nastavni sadržaj nekog područnog znanja za izvođenje na računalu. Engleska riječ *courseware* je nastala kombinacijom riječi „course“ (kurs, tečaj, kolegij) i riječi „ware“ (npr. eng. software, hardware) i samom svojom strukturom upućuje na svoju vezu s računalima. U našem pristupu *courseware* kolegija je višerazinski strukturiran na elemente nastavnog sadržaja (Slika 2. 13.): *nastavne cjeline* (NC), *nastavne teme* (NT), *nastavne jedinice* (NJ), *nastavne pojmove* (NP) kao i *elemente za testiranje i ocjenjivanje znanja* učenika.

Nastavna cjelina u načelu uključuje više nastavnih tema, dok dalje nastavna tema uključuje više nastavnih jedinica i konačno nastavna jedinica uključuje više nastavnih pojmova. Elementi za testiranje i ocjenjivanje znanja učenika se pridružuju nastavnoj cjelini, nastavnoj temi ili nastavnoj jedinici kao i kolegiju u cjelini. Nastavni pojam smo uveli kao sadržajno najmanji element strukture nastavnog sadržaja (dalje nedjeljiv) koji odgovara čvoru znanja u bazi područnog znanja. Temeljna zadaća učitelja je raščlamba nastavnih sadržaja prema ovoj strukturi kao i priprema i generiranje testova i ispita za provjeru znanja učenika.

Tablica 2. 5. Značenje simbola u okruženju za oblikovanje nastavnog sadržaja

Učitelj	
Oznaka	Značenje
	Kolegij
	Agregacija
	SCO
	SCO kviz
	Čvor znanja
	Pitanje (element kviza)

Oblikovanje nastavnog sadržaja za kolegij Programiranje 1 po predlošku 4

Struktura nastavnog sadržaja

-  Programiranje 1
-  Programiranje - temeljni pojmovi
-  Osnove razvoja programske podrške
-  Algoritamske strukture s primjerima
-  Vodič kroz QBASIC
-  Metode programiranja
-  Rječnik pojmova u programiranju
-  Programiranje 1

 Izradi novi SCO

Akcije nad elementom nastavnog sadržaja

Izabrani element:  Kolegij: Programiranje 1

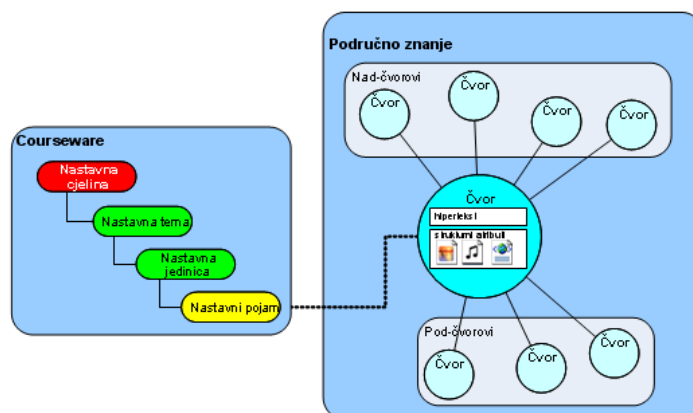
Opis:

-  Dodaj nastavnu cjelinu
-  Pridruži nastavnu cjelinu
-  Izradi i pridruži nastavnu cjelinu
-  Pridruži nastavnu cjelinu kao kviz
-  Izradi i pridruži nastavnu cjelinu kao kviz

Slika 2. 13. Raščlana nastavnog sadržaja kolegija

Oblikovanje nastavnog sadržaja provedeno je promišljajući i na SCORM (Shareable Content Object Referent Model) referentni model [SCOR2004]. Usuglašavajući pristup u programskom sustavu xTEx-Sys sa SCORM referentnim modelom ističemo da SCO (Shareable Content Object) kao temeljni te djeljivi objekt nastavnog sadržaja može biti svaki od navedenih elemenata nastavnog sadržaja, a to znači nastavna cjelina, nastavna tema, nastavna jedinica i nastavni pojam te konačno i element za ocjenjivanje znanja. Međutim, agregacija može biti nastavna cjelina, nastavna tema i nastavna jedinica. Istaknimo da je nedjeljivi element nastavnog sadržaja SCO prema modelu SCORM kojega ističe kao elementarnu jedinicu – granulu nastavnog sadržaja.

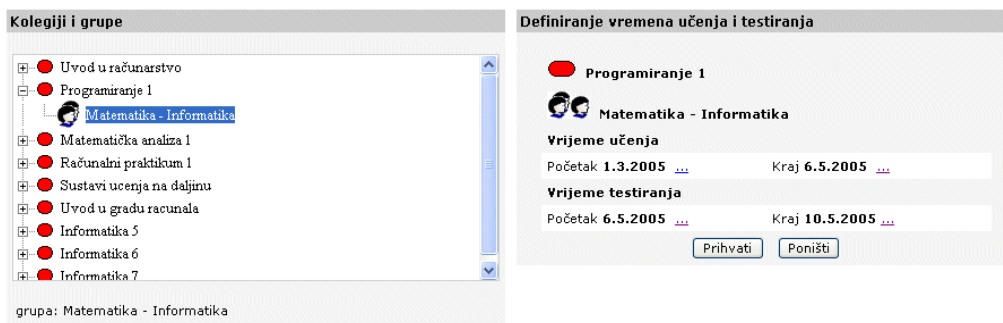
Svaka agregacija ima samo naziv, a SCO osim svog naziva ima i dodatke (eng. asset) koji su u ovom slučaju *čvorovi područnog znanja* (Slika 2. 14). Svaki element nastavnog sadržaja, osim nastavnog pojma, kao SCO može imati proizvoljan broj dodataka (čvorova područnog znanja), dok nastavni pojam kao SCO može imati samo jedan dodatak (čvor područnog znanja). Dodaci SCO-a (čvorovi područnog znanja) imaju i svoje dodatke koji su u stvari *strukturni atributi* određenih čvorova znanja: tekstualni opis, slikovni zapis (u raznim formatima), prezentacija, animirani zapis sa slikom i zvukom (u raznim formatima), zvukovni zapis, URL adresa.



Slika 2. 14. Odnos elemenata strukture područnog znanja i nastavnog sadržaja u sustavu xTEx-Sys

Učitelj nakon provedenog postupka oblikovanja nastavnog sadržaja u njegovoj strukturi određuje elemente kojima pridružuje dinamički način testiranja znanja kao seriju pitanja na koje će učenik odgovarati, a time uči i u proces ocjenjivanja znanja. Elementi za testiranje i ocjenjivanje znanja učenika mogu se pridružiti svakom elementu nastavnog sadržaja osim nastavnom pojmu. Učenik nakon izbora područnog znanja pristupa procesu testiranja znanja.

Učitelj ima i mogućnost određivanja vremena (mjesec i dan) za učenje i poučavanje učenika kao i vrijeme (mjesec i dan) za testiranje znanja učenika (Slika 2. 15.) za odabrani kolegij. U vezi s tim brojni su scenariji učitelju na raspolaganju. Osim navedenog učitelj ima mogućnost i pogleda na rezultate koje su učenici postigli tijekom testiranja i ocjenjivanja znanja.




Slika 2. 15. Određivanje mjesta i vremena učenja, poučavanja i testiranja znanja

2.2.2.3 UČENJE, POUČAVANJE I TESTIRANJE ZNANJA UČENIKA

Okruženje učenika radi učenja, poučavanja i testiranja znanja nema autorske značajke, ali omogućava izbor područnog znanja kao i „navigaciju“ kroz odabrano područno znanje putem didaktički oblikovanih nastavnih sadržaja koje je napravio učitelj.

Učenje, poučavanje i testiranje znanja učenika započinje izborom opcije Učenje, poučavanje i testiranje nastavnog sadržaja. Učenik dobiva formu kao na Slika 2. 16. s popisom kolegija.

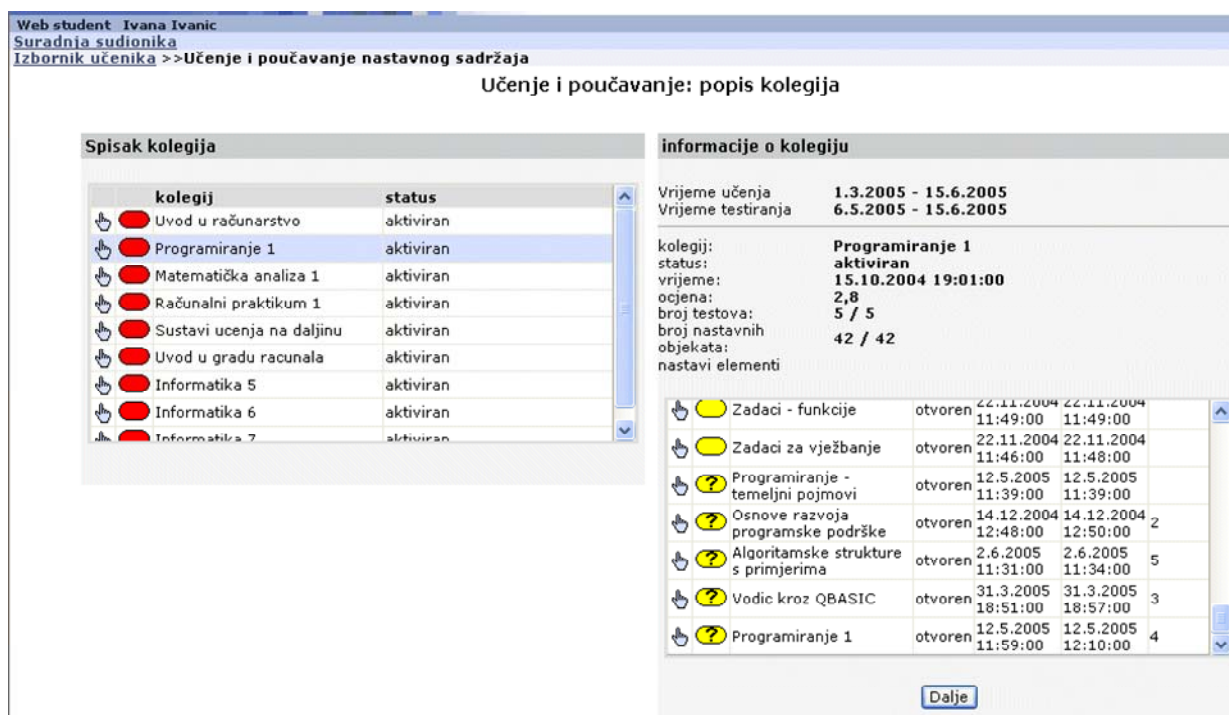
Tablica 2. 6 Značenje simbola u okruženju za učenje i poučavanje

Učenik	
Oznaka	Značenje
	Kolegij
	Agregacija
	SCO
	SCO kviz
	Čvor znanja
	Naprijed-nazad
	Video zapis
	Aplikacija
	Audio zapis
	Slika
	URL



Slika 2. 16. Izbor kolegija

Izborom kolegija učenik dobiva sve potrebne informacije: vrijeme određeno za učenje, vrijeme određeno za testiranje, postignute ocjene u prethodnim testiranjima (ako ih je bilo), broj SCO-ova kao elemenata strukture nastavnih sadržaja kolegija. U ovom se popisu posebno označavaju testovi kao SCO s oznakom „?“ i nazivom kojeg je odredio učitelj (Slika 2. 17).



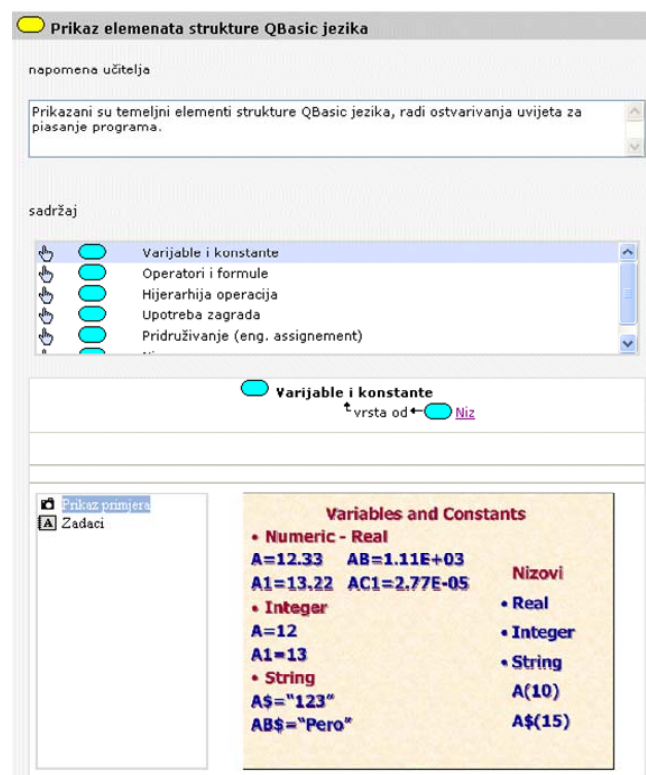
Slika 2. 17. Izabrani kolegij

Nakon odabira elementa strukture nastavnog sadržaja dobiva se njegov prikaz (Slika 2. 18).



Slika 2. 18. Pogled na dio sadržaja nastavne cjeline

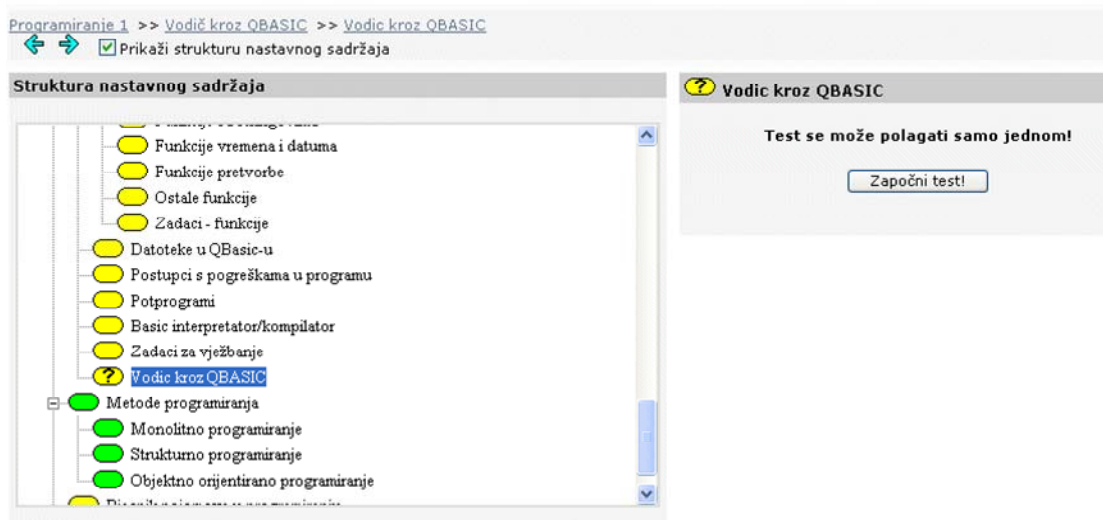
Ako se odabere prikaz SCO-a, dobije se njegov prikaz u obliku niza čvorova znanja sa pripadnim strukturnim atributima, okvirima i opisima stručnjaka (Slika 2. 19). Osim toga na svim formama postoje i napomene učitelja koje je učitelj napisao s namjerom da didaktički i metodički dodatno oblikuje nastavne sadržaje namijenjene učeniku u okviru zadanog kolegija.



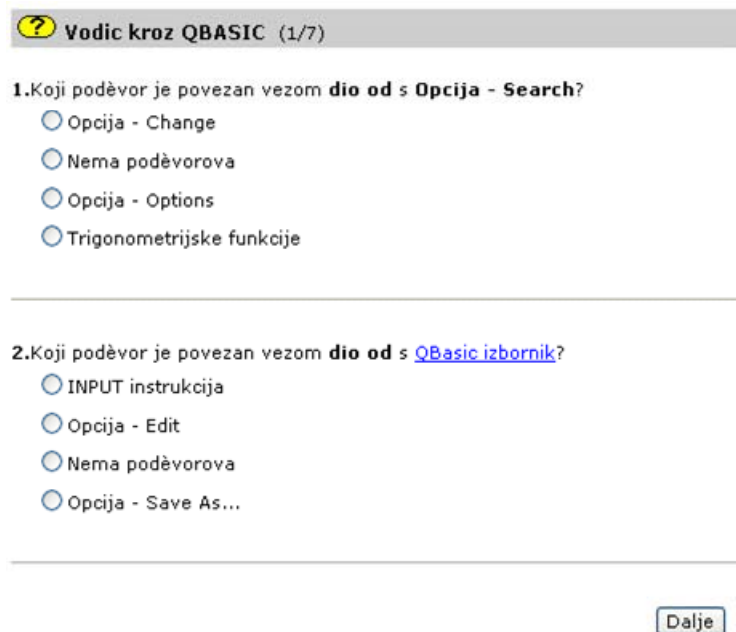
Slika 2. 19. Prikaz sadržaja SCO-a

Tijekom učenja i poučavanja učenik u jednom trenutku donese odluku da započne testiranje svojeg znanja. U vezi s tim odabire SCO test i započinje proces testiranja (Slika 2. 20). U toku testiranja učenik rješava nekoliko serija od po dva pitanja (Slika 2. 21).

Učenje nastavnog sadržaja

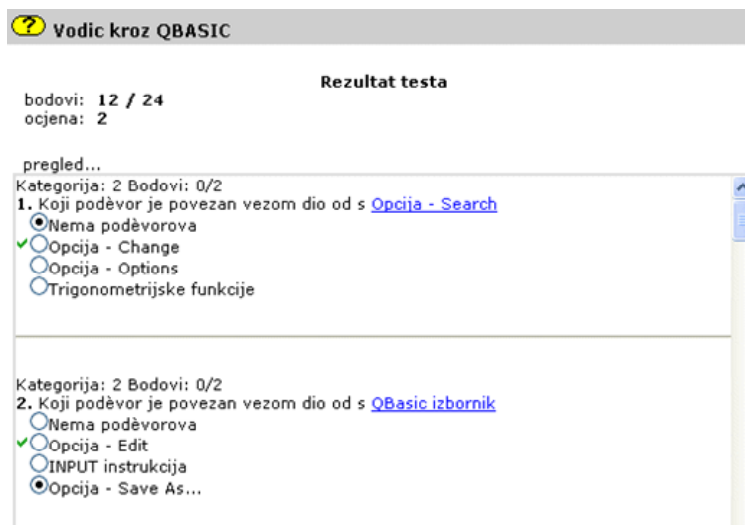


Slika 2. 20. Aktiviranje testa



Slika 2. 21. Serija od dva pitanja

Na kraju učenik dobije izvješće o rezultatima testiranja (Slika 2. 22). Rezultati testa prikazuju „sliku“ bodova kao ocjenu koja je relevantna postignutom broju bodova. Osim toga prikazana su i sva pitanja učeniku postavljena tijekom testiranja s naznakom točnog i netočnog odgovora kao i veze prema čvorovima znanja koji su sudjelovali u pitanju.



Slika 2. 22. Rezultati o izvršenom testiranju

2.3. PRIMJENE MODELA TEx-Sys – KRONIKA PROVEDENIH ISTRAŽIVANJA

Sustavi temeljenih na modelu TEx-Sys primjenjivani su kao nadopuna nastavi koja se realizirala kako u ustanovama primarnog obrazovanja, tako i u ustanovama akademskog obrazovanja. Testiranja provedena u svrhu vrednovanja cjelokupnog modela su, dakle, obuhvatila učenike po svim kronološkim dobima, od onih u primarnom obrazovanju do onih na akademskoj razini. U realizaciji spomenutih testiranja korištene su oblikovane baze znanja te na njima zasnovani i oblikovani nastavni sadržaji kolegija.

Pored testiranja provedenih na sustavu xTEx-Sys, provedena su i testiranja na ostalim postojećim sustavima, tj. na inicijalnom sustavu TEx-Sys i na sustavu DTEEx-Sys. Testiranja su opisana područnim znanjem, uzorkom i brojem ispitanika, te vremenom testiranja. Kod svakog istraživanja, svaki ispitanik se učio, poučavao i testirao na svakoj bazi područnog znanja, tako da kod svakog istraživanja se ukupni broj testova dobije kao umnožak broja ispitanika i broja pripadnih baza područnog znanja. Položenim se smatra onaj test u kojem je student ostvario tzv. 0-1-9 kriterij, tj. bio je 0 puta u I. težinskoj kategoriji, 1 put u II. težinskoj kategoriji i 9 puta u III. težinskoj kategoriji.

2.3.1. TESTIRANJA NA SUSTAVU TEx-Sys

Testiranja na sustavu TEx-Sys su se provodila za područna znanja o korištenju energije u kućanstvu s učenicima u primarnom obrazovanju, područna znanja iz fizike o gibanju planeta, općeg zakona gravitacije i primjene zakona gravitacije s učenicima u srednjoj školi.

2001/2002					
PODRUČNO ZNANJE	UZORAK	BROJ ISPITANIKA	VRIJEME	BAZE PODRUČNOG ZNANJA	UKUPAN BROJ TESTOVA
Uvod u računarstvo	1. godina studijskih grupa M, MI, ITK FPMZ i OP Sveučilište u Splitu	18	Ak.god. 2001/02	Računalo kao sustav (48 čvorova, 46 veza) Win98 (83 čvorova, 72 veza) Programiranje (79 čvorova, 82 veze) Računalne mreže (68 čvorova, 59 veze)	72
2004/2005					
PODRUČNO ZNANJE	UZORAK	BROJ ISPITANIKA	VRIJEME	BAZE PODRUČNOG ZNANJA	UKUPAN BROJ TESTOVA
Korištenje energije u kućanstvu	7. razred OŠ „Ravne njive“, Split	94	Ožujak, 2005.	<i>Korištenje energije u kućanstvu</i> (110 čvorova, 150 veza)	94
Fizika za područja gibanja planeta, općeg zakona gravitacije i primjene zakona gravitacije	1. razred Gimnazija „Antun Vrančić“, Šibenik	25	Travanj, 2005.	Gibanje planeta (37 čvorova, 27 veza) Opći zakon gravitacije (29 čvorova, 27 veza) Primjena zakona gravitacije (110 čvorova, 150 veza)	75
UKUPNO					241

2.3.2. TESTIRANJA NA SUSTAVU DTEx-Sys

Testiranja na sustavu DTEx-Sys su se provodila za područna znanja o geometrijskim likovima te životinjama prema staništima s učenicima u primarnom obrazovanju, modeliranje i vođenje procesa, uvoda u računarstvo te objektno orijentirane paradigme sa studentima na fakultetima.

2002/2003

PODRUČNO ZNANJE	UZORAK	BROJ ISPITANIKA	VRIJEME	BAZE PODRUČNOG ZNANJA	UKUPAN BROJ TESTOVA
Uvod u računarstvo	1. godina studijskih grupa M, MI, ITK, FI FPMZ i OP Sveučilište u Splitu	72	Ak.god. 2002/03	Računalo kao sustav (48 čvorova, 46 veza) Commanderi MS DOS (95 čvorova, 88 veza) Win98 (83 čvorova, 72 veza) MS Word (80 čvorova, 77 veze) MS PowerPoint (93 čvorova, 111 veze) MS Excel (116 čvorova, 113 veze) Programiranje (79 čvorova, 82 veze) Računalne mreže (68 čvorova, 59 veze)	648

2003/2004

PODRUČNO ZNANJE	UZORAK	BROJ ISPITANIKA	VRIJEME	BAZE PODRUČNOG ZNANJA	UKUPAN BROJ TESTOVA
Modeliranje i vođenje procesa u kemijskom inženjerstvu	2. godina smjera Kemijsko inženjerstvo Fakultet Kemijskog inženjerstva i tehnologije Sveučilišta u Zagrebu	63	Ak.god. 2003/04	MS Windows (117 čvorova, 120 veza) Računalo kao sustav (48 čvorova, 46 veza) Programiranje (79 čvorova, 82 veze)	169
	4. godina smjera Kemijsko inženjerstvo Fakultet Kemijskog inženjerstva i tehnologije Sveučilišta u Zagrebu	14	Ak.god. 2003/04	Značajke mjernih pretvornika (64 čvorova, 62 veza) Modeliranje i vođenje procesa (140 čvorova, 155 veza) Dvostupnjeviti izmjenjivač topline (109 čvorova, 104 veze)	42
Uvod u računarstvo	1. godina studijskih grupa M, MI, IF, ITK, FTK FPMZ i OP Sveučilište u Splitu	76	Ak.god. 2003/04	Računalo kao sustav (48 čvorova, 46 veza) Brojevi sustavi i logičke funkcije (220 čvorova, 105 veza) MS DOS (95 čvorova, 88 veza) MS Windows (83 čvorova, 72 veza) Programiranje (79 čvorova, 82 veze)	380

2004/2005

PODRUČNO ZNANJE	UZORAK	BROJ ISPITANIKA	VRIJEME	BAZE PODRUČNOG ZNANJA	UKUPAN BROJ TESTOVA
Područno znanje o geometrijskim likovima „Kocke i kvadri“	4. razred OŠ „Spinut“, Split	10	Listopad, 2004.	Kocke i Kvadri (44 čvorova, 77 veza)	10
Područno znanje o životinjama prema staništima	4. razred OŠ „Spinut“, Split	10	Listopad, 2004.	Zivotinje (111 čvorova, 107 veza)	10

Uvod u računarstvo	1. godina studijskih grupa M, MI, IF, ITK, FTK FPMZ i OP Sveučilište u Splitu	114	Ak.god. 2004/05	Računalo kao sustav (48 čvorova, 46 veza) Brojevni sustavi i logičke funkcije (220 čvorova, 105 veza) MS DOS (95 čvorova, 88 veza) MS Windows (83 čvorova, 72 veza) MS Word (80 čvorova, 77 veze) MS PowerPoint (93 čvorova, 111 veze) MS Excel (116 čvorova, 113 veze) MS Access (110 čvorova, 102 veze) Računalne mreže (68 čvorova, 59 veze)	1026
Objektno orijentirana paradigma	2. godina studijskih grupa M, MI, ITK FPMZ i OP Sveučilište u Splitu	31	Listopad, 2004.	OOP (202 čvorova, 189 veza)	31
UKUPNO					2316

Posebno je zanimljiva anketa koja je provedena tijekom testiranja provedenog u akademskoj godini 2002/03 u okviru područnog znanja Uvod u računarstvo. Navodimo analizu rezultata te ankete:

1. Na pitanje je li rad u sustavu DTEEx-Sys pripomogao boljem razumijevanju gradiva kolegija Uvod u računarstvo 51% studenata je odgovorilo potvrdno. Iznosimo neka od obrazloženja:

DA:

- neke pojmove koje nisam znala imala sam mogućnost naučiti; bili bi bolje da su pojmovi malo opširnije opisani
- točno se zna hijerarhija pojmova pojedinog područja
- zbog preglednosti i lakšeg načina učenja
- pomaže u onome što se nije dobro naučilo
- jednostavniji način učenja
- pomoglo mi je pri shvaćanju veza između različitih pojmova
- sažeto i jasno prikazuje glavne pojmove gradiva
- zbog sustavnog predočavanja i objašnjavanja veza između pojmova i njihovog značenja
- sistematizacija znanja, tj. stavljanja svega na svoje mjesto
- lakše se pronade ono što nas zanima
- bolje razumijevanje i povezivanje pojmova
- koliko zabavan toliko i koristan

NE:

- pomaže mi samo za sistematizaciju znanja, a za razumijevanje su mi potrebne knjige
- teško mi se naviknuti na takav način učenja, gradivo je razbacano, nema nekog reda
- komplicirano je i teško za snaći se
- treba puno vremena da se uhoda u takav način učenja, a poslije je bilo sve napornije nego uzeti knjigu i učiti
- samo sistematizira postojeće znanje, ali ne daje praktično znanje
- standardno učenje je puno bolje
- smatram da je lakše učiti iz knjiga, ovako se gubi dosta vremena
- odnosi među pojmovima su vrlo nerazumni i nedefinirani

2. Na pitanje bi li koristili modul Learning za usvajanje NOVOG gradiva 51% studenata je odgovorilo potvrdno. Iznosimo neka obrazloženja:

DA:

- jednostavnost; mnogo je lakše učiti i zatim ispitati svoje znanje
- zbog preglednosti i lakšeg načina učenja
- zanimljivo je učiti na taj način
- zbog sustavnog predočavanja i objašnjavanja veza između pojmova i njihovog značenja

- vidljivi su odnosi između pojmova i njihovo značenje što olakšava učenje
- lakoća korištenja
- lakše se uči
- mogu se zadržati koliko hoću, ne moram nikoga pitati da mi ponavlja već se jednostavno vratim na ono što nisam razumjela
- inače učim na sličan način
- zbog preglednosti pojmova te bržeg i boljeg snalaženja te mogućnosti pronalaženja nepoznatih pojmova
- ovaj način je zabavan, a puno naučimo

NE:

- pojmovi nisu dobro opisani i konfuzno su povezani; komplicirano je
- lakše se snalazim u učenju iz knjiga
- lakše je povezivati pojmove i učiti kad su rečenice suvisle, kao mali sastavi, a ne gomila pojmova povezanih s dvije-tri veze; ne može se sve opisati s tim vezama
- čini mi se malo težim načinom učenja
- jedino uz dodatni udžbenik i to samo za sistematizaciju
- možda za rekapitulaciju već naučenog
- nema nikakvih objašnjenja, iznesene su samo činjenice
- jer ne pojašnjava najbolje neke stvari; ima previše linkova
- za ponavljanje već stečenog znanja, jer lekcije u sustavu nisu zanimljive, komplicirane su i nemaju toliku jasnoću da bi se iz njih učilo

3. Na pitanje bi li sustav DTEEx-Sys, radi pomoći pri učenju i poučavanju na odabranom područnom znanju, preporučili svojem kolegi/kolegici 61% studenata je odgovorilo potvrdno. Iznosimo neka obrazloženja:

DA:

- sustav je dobro osmišljen
- daje više znanja i lakši pristup učenju i shvaćanju gradiva
- vrlo je poučno
- jednostavan i praktičan sustav za učenje
- omogućava brzo i jednostavno učenje nepoznatog gradiva
- zbog sustavnog predočavanja i objašnjavanja veza između pojmova i njihovog značenja
- olakšava učenje, zanimljivije je od standardnog učenja iz knjiga
- pronađe se puno zanimljivih podataka
- zbog sveobuhvatnosti i povezanosti sadržaja
- jednostavnije je i lakše za učiti, vrlo je praktično
- sve informacije su približene putem računala koje je spojeno na Internet
- zbog lakšeg povezivanja pojmova te utvrđivanja njihovih odnosa, radi lakše preglednosti i mogućnosti bržeg učenja
- to je nešto novo za nas, a za razliku od knjiga ovo je mnogo zanimljivije

NE:

- radije bih im preporučila neke dobre knjige i skripte
- sustav je samo sistematizirao gradivo, ali ga nije baš objasnio
- samo kao ponavljanje

4. Na pitanje smatra li da je modul Quiz dobar način ispitivanja učenikova znanja 57% studenata je odgovorilo potvrdno. Iznosimo neka obrazloženja:

DA:

- kroz pitanja provjeravamo svoje znanje, ali uz to i učimo pogotovo na svojim greškama
- odmah daje rezultate i učenik tako lakše uči
- jednostavno i zabavno
- nemam tremu kad odgovaram i vidim odmah rezultate
- brzo i jednostavno; ili znaš ili ne znaš
- odmah nakon rješavanja se dobije ocjena i vide se pogreške; profesori ne moraju ručno ispravljati ispite
- lakše je od pisanja
- pitanja su temeljita i pomoću njih se bolje uči i provjerava svoje znanje
- imaš određeno pitanje i ako znaš, znaš, a što je najbolje nema onih pitanja, "zašto ovako, zašto onako"

NE:

- nije dovoljno da saznamo koliko netko zna
- dobro je za profesore jer dobiju na vremenu
- zanimljiv je, ali samo to
- pitanja su glupo postavljena i uopće ih ne razumiješ
- ne može se provjeriti konkretno znanje jer učenici mogu lako varati
- na testiranjima većina studenata vara pa ovakvo učenje nema svrhe
- koncentracija pred ekranom se uvelike razlikuje od opuštenosti uz knjigu i komunikaciju s profesorom

5. Neke primjedbe na realizaciju modula Quiz:

- pitanja se znaju ponavljati više puta

- pitanja nisu dobro postavljena
- ponavlja se i više točnih odgovora na pitanje koje dopušta samo jedan odgovor
- baze su pomiješane
- rečenice su nejasne i ponekad nije baš jasno što se traži
- neka pitanja su toliko nesuvisla da se na njih može lako odgovoriti i da ništa ne znaš
- pomiješa pitanja iz drugih baza
- malo bolja organizacija i povezanost te smislenost pitanja i ponuđenih odgovora
- pitanja nisu baš jasna

2.3.3. TESTIRANJA NA SUSTAVU xTeX-Sys

Testiranja na sustavu xTeX-Sys su se provodila za područna znanja iz informatike i hrvatskog jezika s učenicima u primarnom obrazovanju, modeliranje i vođenje procesa te građe računala sa studentima na fakultetima.

Oblikovanje baza područnih znanja omogućila je testiranje okruženja za rad stručnjaka, dok je oblikovanje nastavnih sadržaja kolegija omogućilo testiranje okruženja učitelja. Uloge stručnjaka i učitelja obavljali su nastavnici i studenti na Fakultetu prirodoslovno matematičkih znanosti Sveučilišta u Splitu te studenti Visoke učiteljske škole Sveučilišta u Splitu.

Okruženje za rad učenika je testirano s ciljem utvrđivanja upotrebljivosti sustava, kao i s ciljem utvrđivanja postignuća učenika u procesu učenja i poučavanja.

2004/2005					
PODRUČNO ZNANJE	UZORAK	BROJ ISPITANIKA	VRIJEME	BAZE PODRUČNOG ZNANJA	UKUPAN BROJ TESTOVA
Katalog znanja iz hrvatskog jezika za osnovnu školu	4. godina učiteljskog studija Visoke učiteljske škole Sveučilišta u Splitu	8	Ak.god. 2004/05	Riječi (102 čvorova, 10 veza)	8
Modeliranje i vođenje procesa u kemijskom inženjerstvu	4. godina smjera Kemijsko inženjerstvo na Fakultetu Kemijskog inženjerstva i tehnologije Sveučilišta u Zagrebu	11	Ak.god. 2004/05	Modeliranje i vođenje procesa (140 čvorova, 155 veza) Dvostupnjeviti izmjenjivač topline (109 čvorova, 104 veza) LabTeX-Sys (57 čvorova, 53 veze)	33
Katalog znanja iz informatike za osnovnu školu	5., 6. i 8. razred OŠ „Spinut“, Split	54	Ak.god. 2004/05	Osnove informacijske i komunikacijske tehnologije (103 čvorova, 7 veza) Strojna i programska oprema (114 čvorova, 5 veza) Multimediji (120 čvorova, 6 veza) Obrada teksta (146 čvorova, 4 veza) Proračunske tablice i baze podataka (61 čvor, 6 veza) Izrada prezentacija (172 čvora, 5 veza) Izrada Web stranica (24 čvora, 4 veza) Rješavanje problema i programiranje (111 čvora, 3 veza) Internet (109 čvora, 4 veza)	486

2005/2006					
PODRUČNO ZNANJE	UZORAK	BROJ ISPITANIKA	VRIJEME	BAZE PODRUČNOG ZNANJA	UKUPAN BROJ TESTOVA
Područno znanje programiranju	5. i 7. razred OŠ „Spinut“, Split	115	Svibanj, 2006.	Programiranje	115
Područno znanje o prirodi i društvu	1., 2., 3. i 4. razred OŠ „Spinut“, Split	143	Svibanj, 2006.	Priroda i društvo	143
Područno znanje o fizici	8. razred OŠ „Ravne njive“, Split	92	Ak god. 2005/06	Optika	92
Uvod u računarstvo	1. godina studijskih grupa M, MI, IF, ITK, FTK FPMZ i OP Sveučilište u Splitu	102	Ak god. 2005/06	Računalo kao sustav (48 čvorova, 46 veza) Brojevi sustavi i logičke funkcije (220 čvorova, 105 veza) MS DOS (95 čvorova, 88 veza) MS Windows (83 čvorova, 72 veza) MS Word (80 čvorova, 77 veza) MS PowerPoint (93 čvorova, 111 veza) MS Excel (116 čvorova, 113 veza) MS Access (110 čvorova, 102 veza) Računalne mreže (68 čvorova, 59 veza)	918
Područno znanje o kemiji	7. i 8. razred OŠ „Fra oca Petra Perice“, Makarska	100	Ak god. 2005/06	Kemija	100
UKUPNO					1895

2.3.4. VREDNOVANJA UČINKA

Od svih prethodno navedenih testiranja modela TEx-Sys samo su dva provedena u svrhu ispitivanja učinka. Radi se o eksperimentima provedenima na sustavu DTEEx-Sys. Iskustva stečena tijekom realizacije tih eksperimenata uvelike su utjecala na metodologiju za vrednovanje učinka koja je opisana u četvrtom poglavlju.

PRVI EKSPERIMENT

U prvom eksperimentu, koji je detaljno opisan u [STAN2004b], sudjelovalo je 22 studenta prve godine studijskih grupa matematike, matematike i informatike, informatike i fizike, informatike i tehničke kulture te fizike i tehničke kulture FPMZ i OP Sveučilišta u Splitu, i 11 studenata prve godine Kemijsko-tehnološkog fakulteta (KTF) Sveučilišta u Splitu koji su upisali kolegij *Uvod u računarstvo* u akademskoj godini 2004/2005. Studenti s FPMZ i OP bili su podijeljeni u dvije grupe: eksperimentalnu (11 studenata) i tutorsku (11 studenata), dok su studenti s KTF-a pripadali kontrolnoj grupi.

Kontrolna grupa je sudjelovala u tradicionalnom procesu učenja i poučavanja, eksperimentalna grupa je koristila sustav DTEEx-Sys, a tutorska grupa je poučavana od strane predmetnih profesora i asistenata kolegija Uvod u računarstvo (studenti iz tutorske grupe su podijeljeni u četiri podgrupe s kojima su radili dodijeljeni im tutori). Studenti su bili uključeni u pripadajuće im oblike procesa učenja i poučavanja dva sata tjedno tijekom cijelog semestra (2 sata x 15 tjedana = 30 sati po semestru). Eksperimentalna grupa se učila i poučavala na sljedećim bazama područnog znanja u okviru sustava DTEEx-Sys: baze znanja: Računalo kao sustav (48 čvorova, 46 veza), Brojevnih sustavi i logičke funkcije (220 čvorova, 105 veza), MS DOS (95 čvorova, 88 veza), MS Windows (83 čvorova, 72 veza), MS Word (80 čvorova, 77 veze), MS PowerPoint (93 čvorova, 111 veze), MS Excel (116 čvorova, 113 veze).

Sve tri grupe su pisale 45-minutni inicijalni test na početku semestra. Također su sve tri grupe pisale 60-minutni završni test dva tjedna nakon završetka semestra. Oba testa su bodovana na ljestvici od 0 do 100 bodova. Inicijalni test je dao informaciju o postojanju statistički značajne razlike između grupa s obzirom na predznanje studenata. Završni test je dao informaciju o postojanju statistički značajnih razlika između grupa s obzirom na utjecaj različitih načina poučavanja.

Vrednovanje učinka sustava DTEEx-Sys je pokazalo da je strategija učenja i poučavanja učinkovita. Naime, eksperimentalna grupa je pokazala bolje rezultate nego kontrolna grupa kod svakog statističkog testa. Značajna statistička razlika između kontrolne i eksperimentalne grupe je pokazala da postoje prednosti učenja i poučavanja u sustavu DTEEx-Sys naspram tradicionalnog učenja. Nepostojanje statistički značajne razlike između eksperimentalne i tutorske grupe je pokazalo da je sustav DTEEx-Sys sposoban zamijeniti živog učitelja. Veličina učinka sustava DTEEx-Sys je 0.82 što je malo manje od 0.84, što je standardna vrijednost za učinak inteligentnih tutorskih sustava (prema [FLET2003]).

DRUGI EKSPERIMENT

U drugom eksperimentu koji je detaljno opisan u [STAN2005], sudjelovao je 31 student druge godine studijskih grupa matematika i informatika, matematika, te informatika i tehnička kultura FPMZ i OP Sveučilišta u Splitu, koji su upisali kolegij *Računalni praktikum I* u akademskoj godini 2004/2005. Studenti su podijeljeni na eksperimentalnu (20 studenata) i kontrolnu grupu (11 studenata) slučajnim odabirom.

Kontrolna grupa je sudjelovala u tradicionalnom procesu učenja i poučavanja, eksperimentalna grupa je koristila sustav DTEEx-Sys na bazi područnog znanja *Objektno-orijentirana paradigma OOP* (202 čvorova, 189 veza).

Svi studenti su 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 0-100, da bi se vidjelo ima li kakvih razlika u predznanju između kontrolne i eksperimentalne grupe. Eksperimentalnoj grupi je slijedio jedan tjedan učenja i poučavanja o objektno-orijentiranom programiranju na sustavu DTEEx-Sys. Kontrolna grupa nije imala pristup, te su oni nastavili učiti iz ranije dobivenih tekstualnih materijala. Na kraju tog tjedna svi studenti su 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 grupe nakon korištenja sustava. Test je također ocjenjivan bodovima od 0-100.

Vrednovanje učinka sustava DTEEx-Sys je pokazalo da nema statistički značajnih razlika između kontrolne i eksperimentalne grupe u rezultatima završnog ispitivanja. Dakle, možemo zaključiti da tretman eksperimentalne grupe tj. učenje na sustavu DTEEx-Sys nije imalo značajnog utjecaja na njihov uspjeh. Bilo kakva razlika između eksperimentalne i kontrolne grupe je mogla nastati kao posljedica greške. Veličina učinka sustava DTEEx-Sys koji je dobiven istraživanjem je 0.12. Kako smo očekivali značajnije rezultate (s obzirom na veličinu učinka dobivenu u prethodno opisanom eksperimentu), bilo je potrebno ispitati što je moglo utjecati na istraživanje. U tu svrhu proveli smo provjeru karakteristika mjernih instrumenata tj. testova koji su se koristili za ispitivanje znanja studenata [MUŽI1977].

Valjanost: 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.

Pouzdanost: 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. Pouzdanost za inicijalno ispitivanje i završno ispitivanje približno jednaka i iznosi 0.85 što je dovoljno visok koeficijent prema (Mužić, 1997), što znači da se možemo osloniti na rezultate testova.

Objektivnost i osjetljivost: Kod našeg istraživanja nije ispitana objektivnost testova iz praktičnih razloga. Naime, ovi su se testovi trebali upotrijebiti samo jedan put te se ne bi isplatilo okupiti veći broj ocjenjivača koji bi morali poznavati specifično područno znanje i biti upoznati s istraživanjem. Ova karakteristika je primjerenija ako se testove planira 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.

Diskriminativna vrijednost zadataka: Pojedini zadaci u testu su više valjani što se više slaže uspjeh učenika na tom zadatku s uspjehom na čitavom testu. Prema tome, valjani zadaci omogućuju razlikovanje uspješnih i neuspješnih učenika. Pearson-ov koeficijent korelacije je izračunat 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. ($r_{pb}=0.12$) završnom ispitivanju). Analizom rezultata bez tih zadataka dobije se veličina učinka od 0.18 standardnih devijacija, što je neznatno povećanje u odnosu na rizik gubitka pouzdanosti i osjetljivosti ukoliko bi se ti zadaci izbacili iz testa.

Prilikom provedbe ovog eksperimenta nismo dobili visoku veličinu učinka, 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. Samo istraživanje je trajalo relativno kratko na manjem broju studenata, a vjerojatno bi se kroz dulje vrijeme studenti bolje priviknuli na rad sa sustavom DTEEx-Sys. Potrebno je prilikom analize istraživanja uzeti u obzir i koliko vremena su studenti učili on-line na sustavu DTEEx-Sys, te ispitati imaju li svi odgovarajuću tehnologiju za pristup (pristup Internetu, računalo i odgovarajuću programsku podršku).

Temeljem prethodno iznesenog, kao zaključak ovog poglavlja iznosimo rekapitulaciju svih testiranja. Tijekom rada na razvoju i implementaciji modela TEx-Sys, te njegovih izvedenica DTeX-Sys i xTEx-Sys, u razdoblju od 2001. do 2006. godine, provedeno je ukupno 4452 testiranja znanja na različitim područnim znanjima u kojima je sudjelovalo ukupno 1134 učenika i studenata (Tablica 2. 7., Tablica 2. 8., Tablica 2. 9.).

Tablica 2. 7. Broj testiranja realiziranih na modelu TEx-Sys po akademskim godinama

	TEx-Sys	DTeX-Sys	xTEx-Sys	UKUPNO
2001/2002	72	-	-	72
2002/2003	-	648	-	648
2003/2004	-	591	-	591
2004/2005	169	1077	527	1773
2005/2006	-	-	1368	1368
UKUPNO	241	2316	1895	

Tablica 2. 8. Broj učenika koji su sudjelovali u testiranjima na modelu TEx-Sys po akademskim godinama

	TEx-Sys	DTeX-Sys	xTEx-Sys	UKUPNO
2001/2002	18	-	-	18
2002/2003	-	72	-	72
2003/2004	-	153	-	153
2004/2005	119	165	73	357
2005/2006	-	-	552	552
UKUPNO	137	390	625	

Tablica 2. 9. Pregled baza područnih znanja na kojima su se vršila testiranja

	TEx-Sys	DTeX-Sys	xTEx-Sys
Uvod u računarstvo	1		
Win98	1	1	
Programiranje	1	3	2
Računalne mreže	1	2	2
Korištenje energije u kućanstvu	1		
Gibanje planete	1		
Opći zakon gravitacije	1		
Primjena zakona gravitacije	1		
Računalo kao sustav		4	1
Commanderi		1	
MSDOS		3	1
MSWord		2	2
MSPowerPoint		2	2
MSExcel		2	2
MSWindows		3	1
Značajke mjernih pretvornika		1	
Modeliranje i vodenje procesa		1	1
Dvostupnjeviti izmjenjivač topline		1	1
Brojevni sustavi i logičke funkcije		2	1
Kocke i kvadri		1	
Životinje		1	
MSAccess		1	1
Objektno-orijentirano programiranje		1	
Riječi			1
LabTEx-Sys			1
Osnove informacijske i komunikacijske tehnologije			1
Strojna i programska oprema			1
Multimediji			1
Izrada Web stranica			1
Priroda i društvo			1
Optika			1
Kemija			1
UKUPNO	8	32	27

Kod nekih testiranja provodila se i anketa o utiscima učenika vezanima za rad na sustavima. Kvalitativni rezultat te ankete pokazali su da su učenici, u globalu, zadovoljni s radom na sustavu i prilično dobro prihvaćaju taj način učenja i poučavanja, ali nisu dali nikakvu kvantitativnu veličinu koja bi nam mogla okarakterizirati njihov učinak. Stoga smo proveli dva eksperimenta vrednovanja učinka sa sustavom DTeX-Sys koja su pokazala različite veličine učinka jednog te istog sustava. To je bio razlog zašto smo pristupili strogo definiranju metodologije za vrednovanje, jer smo smatrali da će uniformnost pristupa eliminirati sve vanjske utjecaje koji bi mogli dovesti do krivih pokazatelja o učinku sustava. Nova metodologija za vrednovanje primijenit će se za vrednovanje učinka sustava xTeX-Sys.

3. STATISTIČKE OSNOVE I METODE VREDNOVANJA SUSTAVA E-UČENJA

U ovom poglavlju dan je pregled osnovnih statističkih pojmova, te su opisani statistički instrumenti potrebni za analizu podataka. Smatramo da je potrebno dati jedan takav repetitorij pojmova i postupaka koji će se spominjati u raspravi o metodama vrednovanja i analiziranju podataka, zbog činjenice da je još uvijek mali broj ljudi „statistički pismen“. Ovdje opisujemo pojmove kao što su varijabla, distribucija, statistički značajna razlika, testovi razlike, korelacija, grafičko prikazivanje rezultata.

Nadalje su opisani postupci i instrumenti prikupljanja podataka te metode vrednovanja sustava e-učenja. Posebno se opisuje eksperiment kao najčešće upotrebljavana metoda vrednovanja. Objašnjeno je što je to veličina učinka i kako se može izračunati, te je naglašena njena primjena kod meta-analiza. Na kraju je ukratko objašnjeno nekoliko primjera vrednovanja sustava e-učenja.

3.1. STATISTIČKE OSNOVE

Znanstveno područje koje se bavi metodama za istraživanje masovnih pojava s pomoću brojčanog izraza, tj. opisivanjem konkretnih rezultata, dobivenih prilikom nekog ispitivanja ili mjerenja, naziva se *statistikom*. Njena je zadaća opisati podatke sređivanjem i sažimanjem da bi bili što pregledniji. Grana posebne statistike koja se bavi masovnim pojavama na području odgoja i obrazovanja naziva se *pedagoškom statistikom* [MUŽI1977]. Dva osnovna načina primjene statistike su: primjena iscrpnog promatranja (deskriptivna statistika) u kojem se obuhvaćaju sve pojave koje su predmetom istraživanja i primjena metode uzoraka (induktivna statistika) u kojoj se promatra samo reprezentativni dio, tj. uzorak cjeline.

U ovom poglavlju bit će objašnjeni neki od osnovnih statističkih pojmova kao što su varijabla, populacija, uzorak, distribucija, nul-hipoteza, statistička značajnost. Opisat će se karakteristike distribucije frekvencija, položaj pojedinog rezultata u grupi, t-test, korelacija, te na kraju, grafički način prikazivanja rezultata. Pri razradi ovog poglavlja korištena su u najvećoj mjeri sljedeća dva literaturna izvora: Mužić, V. „Metodologija pedagoškog istraživanja“ [MUŽI1977] i Petz, B. „Osnovne statističke metode za nematematičare“ [PETZ2004] te se na njih neće više u tekstu posebno referencirati.

3.1.1. TEMELJNI POJMOVI

Jedinice neke pedagoške pojave se razlikuju (ili ne razlikuju) po raznim svojim karakteristikama (učenici neke škole variraju po spolu, uspjehu u pojedinim predmetima, itd.). Svaku takvu karakteristiku bilo koje pedagoške pojave pri kojoj promatramo razlike koje se u njoj javljaju nazivamo *varijablom* [MUŽI1977].

Pod *populacijom* razumijevamo sve članove neke grupe s određenom karakteristikom koju mjerimo. Ponekad je populaciju nemoguće izmjeriti, ili je to skupo i komplicirano pa zato većinom mjerimo ograničeni broj slučajeva koji nazivamo *uzorkom*.

Podatke nad kojima želimo provesti statističku obradu treba najprije srediti prema nekom načelu, a zatim i grupirati, odnosno prikazati u *tablici*. Tablice povećavaju preglednost podataka i čine ih prikladnijima za daljnju statističku obradu. Postoje dvije tehnike grupiranja podataka: određivanje koliko se puta pojavljuje svaki pojedini rezultat ili grupiranje u *razrede* (nekoliko rezultata ulazi u istu grupu). U svaki razred tablice dolazi, po pravilu, isti raspon mogućih rezultata. Isti podatak ne smije ući u dva različita razreda. U tablici ne smije biti praznih polja (- nema pojave, ... ne raspolaze se s podatkom). *Razredna granica* je ona veličina kojom razred u tablici počinje, kao i ona kojom razred završava. *Razredni interval* je veličina razreda, tj. prostor koji se nalazi između donje i gornje granice nekog razreda. *Razredna sredina* je ona točka koja se nalazi točno na sredini između donje i gornje razredne granice.

Kad se kod sređivanja podataka govori o broju neke pojave, onda se zapravo govori o čestoci, tj. *frekvenciji* kojom se u pojedinom razredu tablice pojavljuju podaci. Frekvencije koje se nalaze u pojedinim razredima tablice zapravo se u tim razredima raspoređuju, tj. *distribuiraju*. Prema tome, podaci koje imamo u tablicama zapravo predstavljaju distribuciju frekvencija u određenoj varijabli.

Nul-hipoteza se često spominje u znanosti uopće, a posebno u statistici, a znači „nema razlike“ među pojavama koje mjerimo. Npr. ako nas zanima razlikuju li se u prosječnom uspjehu u nekom testu muškarci od žena, onda nul – hipoteza glasi: „Među njima nema statistički značajne razlike“. Upotrebom statističkih testova provjeravamo postavljenu hipotezu: ako test pokaže da razlika među aritmetičkim sredinama nije statistički značajna, time je nul-hipoteza potvrđena. Naprotiv, ako je razlika statistički značajna, nul-hipoteza je oborena.

Za *razinu značajnosti* (eng. significance level) vrijednosti p se obično uzima vrijednost 0.05 kao minimalna vrijednost za koju je razlika dovoljno velika da bi bila "statistički značajna" (eng. statistically significant). Uzima se razina značajnosti od 5%, što znači: ako zapravo ne postoji nikakva razlika između dviju aritmetičkih sredina, onda bi se takva konkretna razlika koju smo dobili, mogla slučajno dogoditi samo pet puta u 100 mjerenja, a to je malo vjerojatno, pa zato možemo uzeti da je razlika statistički značajna. Prema tome, razina značajnosti od 5% znači zapravo šansu od 5% da smo pogriješili. Drugim riječima, ako se razlika može dogoditi potpuno slučajno u više od 5 slučajeva od 100 ($p > 0.05$), onda tu razliku ne smatramo statistički značajnom razlikom. Što je p manja, razlika je značajnija. Ako je npr. $p = 0.001$ to znači da postoji mogućnost od 1 prema 1000 da se razlika dogodila sasvim slučajno.

Granična vrijednost α je vrijednost za koje se nul-hipoteza odbacuje. Dakle, ako je p vrijednost manja od α , odbacuje se nul-hipoteza, a ako je p vrijednost veća od α , prihvaća se nul-hipoteza.

3.1.2. KARAKTERISTIKE DISTRIBUCIJE FREKVENCIJA

Sređivanjem podataka, izradom tablica i konstrukcijom poligona frekvencija ili histograma uspijevamo podatke izraziti na pregledniji način, ali to nije uvijek dovoljno. Grafičkim se prikazivanjem mogu zornije prikazati razlike između dvije ili tri distribucije, osobito ako su te razlike dovoljno velike. Za točniju usporedbu, a naročito kada su razlike među distribucijama relativno male ili kad se radi o više od tri, potrebni su nam jedinstveni izrazi koji će na najkraći način, npr. nekim jedinstvenim brojem, okarakterizirati čitavu distribuciju. Spomenuti jedinstveni izrazi nazivaju se *karakteristikama distribucije frekvencija*. Dvije su osnovne grupe:

- Srednje vrijednosti – to su karakteristike koje pokazuju na opću razinu čitave distribucije. Srednja je vrijednost, dakle, ona vrijednost oko koje se svrstavaju sve ostale vrijednosti.
- Mjere disperzije – pojedini podaci oko srednje vrijednosti u raznim distribucijama raspoređuju se, odnosno distribuiraju na različite načine. Mjera disperzije upućuje upravo na tu karakteristiku raspršenosti.

3.1.2.1 MJERE CENTRALNE TENDENCIJE ILI SREDNJE VRIJEDNOSTI

Postoji više vrsta srednjih vrijednosti. Npr. pri ocjenjivanju učenika obično se neka ocjena češće pojavljuje od bilo koje druge ocjene. To je tzv. *najčešća vrijednost ili mod*. Ili, tri se učenika natječu u trčanju. Pobjednik stiže prvi, posljednji je treći, a jedan je u sredini, tj. drugi. Njegov rezultat je *središnja ili centralna vrijednost ili medijan*. Ili, *prosjek ili aritmetička sredina* broja bodova svih učenika u nekom testu.

Analiza distribucije frekvencija započinje izračunavanjem srednjih vrijednosti, odnosno mjera centralne tendencije. Pri analizi mjera centralne tendencije je cilj usporediti dvije distribucije frekvencija karakterizirajući svaku distribuciju jednim jedinim brojem (parametrom, karakteristikom). Praktična uporaba srednjih vrijednosti u statistici podrazumijeva određivanje načina izračunavanja srednje vrijednosti koja će okarakterizirati centralnu tendenciju neke distribucije frekvencija. Kako je poznato više mjera centralne tendencije, pred istraživačima se uvijek postavlja pitanje koju srednju vrijednost uporabiti? Na navedenu odluku će utjecati specifična svojstva srednjih vrijednosti, koja će odrediti koja će se srednja vrijednost u danom slučaju uporabiti.

Osnovne karakteristike srednjih vrijednosti su:

- utjecaj ekstremnih obilježja na srednje vrijednosti,
- utjecaj frekvencija u distribuciji frekvencija na srednje vrijednosti,
- utjecaj svih obilježja koja su različita od srednje vrijednosti na tu srednju vrijednost,
- odnos promatrane srednje vrijednosti i drugih obilježja.

Zahtjevi kojima moraju odgovarati sve srednje vrijednosti:

- mogućnost utvrđivanja srednje vrijednosti objektivnim računskim pravilom na jedinstveni način,
- srednja vrijednost mora biti sadržana između najmanje i najveće vrijednosti obilježja,
- ako su sve vrijednosti obilježja jednake, i srednja vrijednost mora biti jednaka toj vrijednosti.

DOMINANTNA VRIJEDNOST ILI MOD

Mod je najčešća vrijednost koja se pojavljuje u nekoj distribuciji. Ako imamo, npr. neku distribuciju rezultata koje su učenici postigli na nekom testu, onda je mod onaj rezultat koji se kod učenika najčešće pojavljuje, onaj koji ima najveću frekvenciju. Ako imamo podatke koji su sređeni u razrede (a da ne znamo originalne podatke), onda, nije moguće odrediti onaj pojedinačni podatak koji predstavlja mod, nego se može odrediti samo modalni razred. Dakle, mod je najčešći pojedinačni podatak, a modalni razred je razred koji kod grupiranih podataka ima najveću frekvenciju. Koliko ima modova, toliko je modalna i distribucija (unimodalna, bimodalna, multimodalna).

CENTRALNA VRIJEDNOST ILI MEDIJAN

Druga srednja vrijednost, koja se često upotrebljava, upućuje na podatak koji je u središtu, tj. centru svih slučajeva kad ih poredamo po rangu. Naziva se *medijan*. Mjesto gdje se medijan u distribuciji nalazi, računa se po sljedećoj formuli:

$$\text{mjesto } M = \frac{N+1}{2} \quad (3.1)$$

gdje je N broj rezultata.

Pri određivanju i izračunavanju medijana polazi se od nakane pronalaženja one vrijednosti numeričkog obilježja koja frekvencije statističkog niza dijeli na dva jednaka dijela. Medijan je, dakle, numerička vrijednost koja niz dijeli na dva jednaka dijela. U jednom dijelu se nalaze elementi koji imaju vrijednost numeričkog obilježja jednaku ili manju od medijana, dok se u drugom dijelu nalaze oni elementi koji imaju vrijednost numeričkog obilježja jednaku ili veću od medijana.

Pri analizi numeričkoga niza uvijek provjeriti jesu li podaci grupirani u razrede (distribucija frekvencija) ili negrupirani, odnosno, je li riječ o negrupiranom nizu sa parnim, ili negrupiranom nizu sa neparnim brojem članova. Nadalje, na izračunavanje medijana će utjecati i podatak o tomu je li numerički niz kontinuiran ili diskontinuiran.

Medijan za negrupirane statističke nizove se računa po sljedećoj formuli:

$$Me = \begin{cases} X_r, & \frac{N}{2} \neq INT, & r = INT\left(\frac{N}{2}\right) + 1 \\ \frac{X_r + X_{r+1}}{2}, & \frac{N}{2} = INT, & r = \frac{N}{2} \end{cases} \quad (3.2)$$

gdje je INT cjelobrojni dio razlomka, X_r razredna sredina r -tog razreda.

Prednosti medijana u odnosu na izračunate srednje vrijednosti (aritmetička sredina, harmonijska sredina) je u tome što je manje osjetljiv na ekstreme, te u tome što je reprezentativnost medijana u izrazito asimetričnim distribucijama frekvencija veća od reprezentativnosti aritmetičke sredine.

ARITMETIČKA SREDINA ILI PROSJEK

Aritmetička sredina ili prosjek je srednja vrijednost koja se najčešće upotrebljava. Aritmetička sredina izračunava se tako da se svi podaci zbroje, pa se zatim zbroj podijeli s brojem podataka:

$$\bar{X} = \frac{\sum X}{N} \quad (3.3)$$

Kad su rezultati grupirani, aritmetička sredina se dobije tako što se množi vrijednost svakog rezultata sa svojom frekvencijom. Formula za izračunavanje aritmetičke sredine grupiranih rezultata je malo složenija:

$$\bar{X} = \frac{\sum fX}{N} \quad (3.4)$$

gdje je f frekvencija.

Algebarski zbroj odstupanja originalnih vrijednosti numeričkog obilježja od aritmetičke sredine jednak je nuli. Zbroj kvadrata odstupanja originalnih vrijednosti numeričkog obilježja od aritmetičke sredine jednak je minimumu. Aritmetička sredina uvijek se nalazi između najmanje i najveće vrijednosti numeričkog obilježja varijable X_i . Ako je vrijednost numeričke varijable X_i jednaka konstanti C , aritmetička sredina te varijable jednaka je konstanti C . Aritmetička sredina je sklona ekstremima.

Pogreška koja se veže uz svaku aritmetičku sredinu uzorka bit će veća što je pojava koju mjerimo više varijabilna i što je uzorak manji. Pogreška ne opada proporcionalno povećanju broja mjerenja, nego proporcionalno drugom korijenu broja mjerenja. Formula za izračunavanje *standardne pogreške* aritmetičke sredine glasi:

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{N}} \quad (3.5)$$

Jedan od najčešćih slučajeva pri eksperimentiranju i statističkoj obradi podataka je uspoređivanje dviju aritmetičkih sredina i testiranje razlike među njima. Ako kažemo da je neka razlika statistički značajna, ustvrdili smo da dobivena razlika nije slučajna, već da razlika vrlo vjerojatno postoji i među populacijama. Ako neka razlika nije statistički značajna, znači da ona može biti i slučajna posljedica variranja uzoraka, a da među populacijama, kojima ti uzorci pripadaju, možda i nema takve razlike.

Standardna pogreška razlike među aritmetičkim sredinama računa se prema formuli:

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}} \quad (3.6)$$

Koliko je puta neka razlika veća od svoje pogreške možemo ustanoviti računanjem *t-vrijednosti*:

$$t = \frac{\text{razlika}}{\text{stan dardna pogreška te razlike}} = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} \quad (3.7)$$

Ako je neka razlika između dvije aritmetičke sredine barem dva puta (točnije 1,96 puta) veća od svoje pogreške, onda se može smatrati statistički značajnom, jer je vrlo malo vjerojatno da će se tako velika razlika dogoditi slučajno.

3.1.2.2 MJERE VARIJABILNOSTI ILI DISPERZIJE

Kod mjerenja mnogih pojava možemo opaziti da se rezultati grupiraju i skupljaju oko jedne srednje vrijednosti koje na nekakav način reprezentiraju rezultate. No, srednje vrijednosti nam ne mogu reći sve. Ako su vrijednosti nekog niza mjerenja gusto grupirane oko srednje vrijednosti, onda nam ta srednja vrijednost dobro reprezentira te rezultate (vrijedi i obrat).

Kao što ima više vrsta srednjih vrijednosti, tako ima i više vrsta *mjera disperzije*. Mjere disperzije su karakteristike distribucije frekvencija koje izražavaju različitu raspršenost vrijednosti numeričkog obilježja oko srednjih vrijednosti (Tablica 3. 1.). Relativne mjere disperzije omogućavaju usporedbu raznorodnih distribucija. Svaka od mjera disperzije se logički veže uz neku od srednjih vrijednosti (Tablica 3. 2.).

Tablica 3. 1. Podjela mjera disperzija

Apsolutne mjere disperzije	Relativne mjere disperzije
1. Raspon varijacija (rang)	1. Koeficijent varijacije
2. Interkvartil	2. Koeficijent kvartilne devijacije
3. Kvartilno odstupanje	
4. 10-90 rang percentila	
5. Srednje apsolutno odstupanje	
6. Varijanca	
7. Standardna devijacija	

Tablica 3. 2. Odnos mjera disperzije i srednjih vrijednosti

Mjera disperzije	Kratica	Odgovarajuća srednja vrijednost
RASPON VARIJACIJE	RV	Mod
KVARTILNO ODSTUPANJE	Q	Medijan
PROSJEČNO ODSTUPANJE	PO	Aritmetička sredina
STANDARDNA DEVIJACIJA	σ	

RASPON VARIJACIJE

Raspon varijacije je najjednostavnija od svih mjera disperzije i predstavlja prvu indikaciju o disperziji. Pri njenom određivanju dolaze u obzir samo ekstremni podaci, tj. najviši i najniži. Kao što sam naziv kaže, raspon varijacije je raspon u kojem mogu varirati podaci unutar neke postojeće distribucije, tj. raspon varijacije je broj raznih podataka koji se mogu pojaviti od najvišeg do najnižeg. Računa se po formuli:

$$R = (X_{\max} - X_{\min}) + 1 \quad (3.8)$$

Kod tablice s grupiranim podacima uz i -ti interval veći od 1 nismo mogli odrediti koji rezultat odgovara modu, nego smo odredili samo modalni razred. Pri određivanju raspona varijacije postupit ćemo analogno. Budući da ne znamo koji je točno najviši, odnosno najniži rezultat, uzet ćemo ekstremne rezultate ekstremnih razreda i s njima primijeniti postupak za određivanje raspona varijacije.

KVARTILNO ODSUPANJE

Kvantili su vrijednosti numeričkog obilježja koje niz uređen po veličini dijele na q jednakih dijelova. Broj kvantila p je za jedan manji od njegova reda q :

- a) Medijan = kvartil reda $q = 2$, $p = 1$,
- b) Kvartili = kvantili koji red dijele na 4 jednaka dijela ($p = 3$):
 - Q_1 - prvi ili donji kvartil
 - Me - drugi kvartil ili medijan
 - Q_3 - treći ili gornji kvartil
- c) Decili = kvantili koji red dijele na 10 jednakih dijelova (broj decila je 9; $p = 9$),
- d) Percentili = kvantili koji red dijele na 100 jednakih dijelova ($p = 99$),

Kvartili su ona mjesta u nekoj distribuciji podataka poredanih po rangu koja dijele tu distribuciju na četiri dijela (Tablica 3. 3.).

Tablica 3. 3. Raspon rezultata

Najviši postignuti rezultat	X_{\max}
Treći kvartil	Q_3
Drugi kvartil	$Q_2 = \text{medijan}$
Prvi kvartil	Q_1
Najniži postignuti rezultat	X_{\min}

Mjesto gdje se Q_1 i Q_3 distribuciji nalaze, računa se po sljedećim formulama:

$$\text{mjesto } Q_1 = \frac{1}{4}(N + 1) \quad (3.9)$$

$$\text{mjesto } Q_3 = \frac{3}{4}(N + 1) \quad (3.10)$$

Kvartili za negrupirane statističke nizove:

$$Q_1 = \begin{cases} X_r, & \frac{N}{4} \neq INT, & r = INT\left(\frac{N}{4}\right) + 1 \\ \frac{X_r + X_{r+1}}{2}, & \frac{N}{4} = INT, & r = \frac{N}{4} \end{cases} \quad (3.11)$$

$$Q_3 = \begin{cases} X_r, & \frac{3N}{4} \neq INT, & r = INT\left(\frac{3N}{4}\right) + 1 \\ \frac{X_r + X_{r+1}}{2}, & \frac{3N}{4} = INT, & r = \frac{3N}{4} \end{cases} \quad (3.12)$$

Upoznali smo određivanje kvartila kod negrupiranih podataka. Sada ćemo pomoću njih izračunati mjeru disperzije, tj. *kvartilno odstupanje*. Kvartilno odstupanje je polovina raspona između trećeg i prvog kvartala. Računa se po formuli:

$$Q = \frac{Q_3 - Q_1}{2} \quad (3.13)$$

Interkvartil je razlika između prvog i trećeg kvartila, tj. raspon numeričkog obilježja kojim je obuhvaćeno 50% elemenata statističkog niza. Računa se po formuli:

$$I_Q = Q_3 - Q_1 \quad (3.14)$$

Razlika između 10-tog i 90-tog percentila računa se po formuli:

$$I_{P_{90}-P_{10}} = P_{90} - P_{10} \quad (3.15)$$

$$P_{90} \rightarrow r = INT\left(\frac{90 \cdot N}{100}\right) + 1 \quad (3.16)$$

$$P_{10} \rightarrow r = INT\left(\frac{10 \cdot N}{100}\right) + 1 \quad (3.17)$$

Koeficijent korelacije računa se po formuli:

$$V_Q = \frac{Q_3 - Q_1}{Q_3 + Q_1} \quad (3.18)$$

PROSJEČNO ILI SREDNJE ODSUPANJE

Prosječno odstupanje prikazuje stupanj raspršenosti rezultata oko srednje vrijednosti. Prosječno odstupanje je aritmetička sredina odstupanja svih rezultata distribucije oko odgovarajuće srednje vrijednosti te distribucije. Pri računanju prosječnog odstupanja najprije ćemo odrediti odstupanje svakog pojedinog rezultata od aritmetičke sredine, a zatim ta odstupanja zbrojiti (izostavljajući predznak) i podijeliti brojem slučajeva. Dakle, prosječno odstupanje je aritmetička sredina apsolutnih vrijednosti svih rezultata distribucije od aritmetičke sredine te distribucije. Računa se po formuli:

$$PO = \frac{\sum |X - \bar{X}|}{N} \quad (3.19)$$

Princip rada s grupiranim podacima sličan je onome kod negrupiranih, a razlika je analogna razlici između računanja aritmetičke sredine grupiranih i aritmetičke sredine negrupiranih podataka. Ono što je kod negrupiranih podataka pojedini rezultat, to je kod grupiranih podataka razredna sredina pomnožena frekvencijom. Računamo po formuli:

$$PO = \frac{\sum f |X_i - \bar{X}_i|}{N} \quad (3.20)$$

STANDARDNA DEVIJACIJA

Pri izračunavanju prosječnog odstupanja ne vodimo brige o predznaku što i nije baš „legalno“. Znamo da je kvadrat svakog broja (pozitivnog ili negativnog) pozitivan broj. Dakle, rješenje leži u tome da se odstupanje svakog pojedinog rezultata od aritmetičke sredine kvadrira, a zatim se iz zbroja tih odstupanja izračuna drugi korijen. *Standardnu devijaciju* računamo po formuli:

$$\sigma = \sqrt{\frac{\sum x^2}{N}} \quad (3.21)$$

gdje je $x = X - \bar{X}$.

Standardnu devijaciju za grupirane podatke računamo po formuli:

$$\sigma = \sqrt{\frac{\sum f x_i^2}{N}} \quad (3.22)$$

gdje je $x_i = X_i - \bar{X}_i$.

Standardna devijacija predstavlja standard za mjerenje varijabiliteta rezultata. Ako aritmetičkoj sredini (i samo njoj) na jednu i na drugu stranu dodamo, odnosno oduzmemo vrijednost standardne devijacije, obuhvatit ćemo oko $\sigma \cdot 100\%$ rezultata. Kod normalne distribucije u intervalu $\bar{X} \pm 1\sigma$ nalazi se 68,26% svih rezultata, a u intervalu $\bar{X} \pm 2\sigma$ 95,44% rezultata, a u intervalu $\bar{X} \pm 3\sigma$ 99,73% rezultata.

Ako iz jedne populacije uzimamo mnogo uzoraka iste veličine i od svakog uzorka izračunamo aritmetičku sredinu, pojedine aritmetičke sredine većinom će se više ili manje razlikovati od prave aritmetičke sredine populacije, ali će se te aritmetičke sredine uzoraka ipak grupirati oko prave aritmetičke sredine i to tako da će ih najviše biti u neposrednoj blizini aritmetičke sredine, dakle, uočava se normalna distribucija aritmetičkih sredina uzoraka. Standardna devijacija te distribucije je manja što su uzorci veći.

Varijanca je prosječno kvadratno odstupanje od prosjeka. Računa se pomoću standardne devijacije.

KOEFICIJENT VARIJABILNOSTI

Kad su nam poznate aritmetička sredina i standardna devijacija nekih rezultata, onda su ti rezultati potpuno definirani i možemo ih uspoređivati s nekim drugim rezultatima. Ako je kod dvije pojave aritmetička sredina ista, onda je kod onog rezultata kod kojeg je standardna devijacija manja, odgovarajuća aritmetička sredina bolji reprezentant rezultata.

Da bismo mogli međusobno uspoređivati varijabilnost različitih pojava i svojstava, služimo se tzv. *koeficijentom varijabilnosti* koji pokazuje koliki postotak vrijednosti aritmetičke sredine iznosi vrijednost standardne devijacije. Pitanje koje se postavlja s tim u vezi glasi: „Kolika je disperzija (ili varijabilnost) neke pojave s obzirom na njenu srednju vrijednost?“. Koeficijent varijabilnosti računa se po formuli:

$$V = \frac{\sigma}{\bar{X}} 100 \quad (3.23)$$

i predstavlja postotnu vrijednost standardne devijacije prema aritmetičkoj sredini.

Koeficijent varijabilnosti vrlo je korisna mjera u svim onim slučajevima kad želimo znati u kojem svojstvu neka grupa varira više, a u kojem manje, te koja od grupa varira više, a koja manje u istom svojstvu.

3.1.3. POLOŽAJ POJEDINOG REZULTATA U GRUPI

3.1.3.1 STANDARDIZIRANO ODSUPANJE (z VRIJEDNOST)

Budući da aritmetička sredina i standardna devijacija potpuno definiraju distribuciju nekih rezultata, to je za svaki rezultat moguće izračunati na koji dio standardne devijacije on pada, a kad znamo taj podatak, znamo odmah i koliko imamo rezultata većih i koliko manjih od tog rezultata. Na taj način možemo točno odrediti položaj pojedinca u grupi.

Standardizirano odstupanje (z vrijednost) pokazuje za koliko je standardnih devijacija neki rezultat viši ili niži od aritmetičke sredine. Drugim riječima, z vrijednost predstavlja odstupanje nekog rezultata od aritmetičke sredine izražen u jedinicama standardne devijacije. Tako rezultat koji je za jednu standardnu devijaciju viši od aritmetičke sredine bit će z vrijednost +1, a ako je za tri standardne devijacije niži od aritmetičke sredine, nego z = - 3. Računa se po formuli:

$$z = \frac{X - \bar{X}}{\sigma} = \frac{x}{\sigma} \quad (3.24)$$

Standardizirano odstupanje se koristi kad želimo prikazati koliko „vrijedi“ pojedini rezultat nekog testa iskazanog bodovima. Naime, ako netko dobije 70 bodova na testu, u kojem ispitanici po rezultatima variraju od 10 do 100, taj mu dobiveni rezultat znači manje nego kada bi rezultati varirali od 40 do 80. Potrebno je, dakle, točno odrediti položaj svakog rezultata među ostalim rezultatima.

Tablica z-vrijednosti koja govori o površini ispod normalne krivulje za određenu z – vrijednost (P = površina standardne normalne krivulje od zadane vrijednosti X do bližeg kraja krivulje), tj. koliko se rezultata nalazi do kraja krivulje. Veličina površine ujedno znači i vjerojatnost.

3.1.3.2 CENTILI I DECILI

Položaj pojedinca u grupi može se izraziti i tzv. *centilom*. Centilne točke (percentilom eng. percent = postotak) su one točke u distribuciji nekog većeg broja ispitanika koje je dijele na stotinu dijelova (postotaka) s jednakim brojem jedinica (npr. učenika). Svaki taj dio se naziva centil. Analogna definicija vrijedi za decil. Decilne točke su one točke u distribuciji nekog većeg broja ispitanika koje je dijele na deset dijelova s jednakim brojem jedinica (npr. učenika). Svaki taj dio se naziva decil.

Prvi centil obuhvaća 1% najslabijih, drugi centil 1% idućih najslabijih, ... deseti centil obuhvaća onaj jedan posto koji je na desetom mjestu od najslabijeg, ... sedamdeseti centil znači da je 70% rezultata jednako ili slabije, a 30% rezultata bolje, itd. Iza tablice za z – vrijednosti možemo za svaki rezultat očitati i njegov položaj u centilima. Npr. ako neki rezultat pada na 1,75 z, znači da se taj rezultat nalazi u 96. centilu, jer ima samo oko 4% rezultata koji su bolji od njega. Decile interpretiramo na način da prvi decil obuhvaća najslabijih (ili najboljih) 10%, drugi decil idućih 10%, itd.

Računski centil nekog rezultata možemo izračunati pomoću formule:

$$\text{centil nekog rezultata} = \frac{\text{rang rezultata}}{N} \cdot 100 \quad (3.25)$$

Ako nas zanima obratni postupak, tj. ako nas zanima koji rezultat odgovara nekom određenom centilu, koristit ćemo sljedeću formulu:

$$X_C = D + \frac{\frac{C \cdot N}{100} - f_D}{f_R} \cdot i \quad (3.26)$$

gdje je X_C = rezultat koji pada u željeni centil, D = prava donja granica razreda od X , C = zadani centil, f_D = ukupni broj rezultata ispod razreda od X , f_R = broj rezultata u razredu od X , i = interval.

Medijan se može povezati uz 50. centil, odnosno uz 5. decil (jer je medijan točno na polovici svih podataka), a kvartili sa 75, odnosno 25. centilom, jer su oni točno na četvrtinama svih podataka.

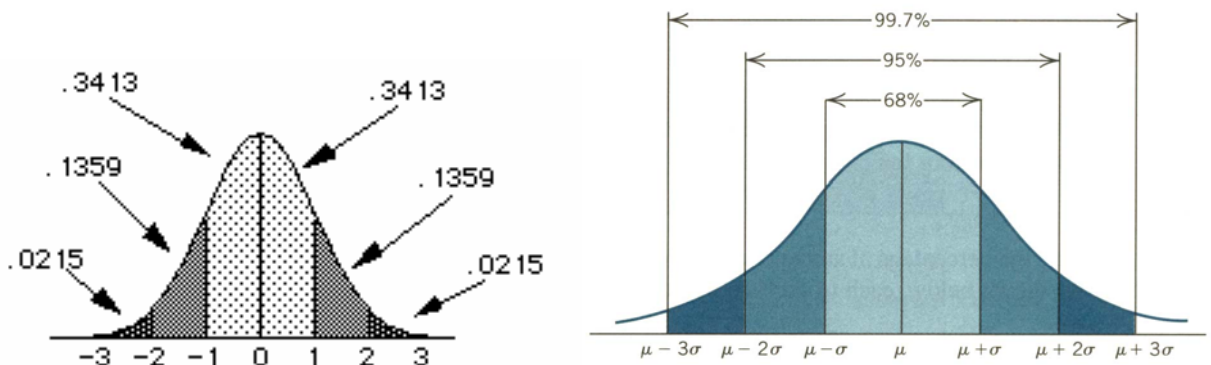
Prednost z – vrijednosti pred centilima sastoji se u tome što z – vrijednosti možemo zbrajati i tražiti njihov prosjek, dok kod centila takvo zbrajanje nije dopušteno. Prednost centila pred z – vrijednostima je u tome što centili ne zahtijevaju normalnu distribuciju: može se iz bilo kakve distribucije odrediti koji rezultat postiže 10% najboljih i li 35% najlošijih ispitanika.

3.1.4. NORMALNA I t -DISTRIBUCIJA

Mnogo puta u praksi se pojavi pitanje kakav je oblik distribucije neke pojave koja nas zanima, tj. kako se ta pojava raspoređuje u prirodi. Ako nam je poznato kako se neka pojava distribuira, možemo s većim razumijevanjem interpretirati dobivene rezultate. Najčešća distribucija jest tzv. *normalna distribucija* ili *Gaussova distribucija* u kojoj nalazimo najviše srednjih, a najmanje ekstremno malih ili ekstremno velikih rezultata. Ta je distribucija simetrična. Normalna distribucija je potpuno definirana ako joj znamo aritmetičku sredinu i standardnu devijaciju. Jedna od osnovnih karakteristika normalne distribucije je to da se tzv. „mjesto infleksije“ nalazi iznad $\pm 1\sigma$.

Pomoću Kolmogorov-Smirnov testa normalnosti podataka ispitujemo koliko se distribucija uzorka s kojim radimo razlikuje od normalne distribucije. Ako za razinu značajnosti $\alpha=0.05$ dobijemo kod K-S testa $p>0.05$, onda možemo smatrati da su podaci normalno distribuirani, ako je $p<0.05$, onda podaci nisu normalno distribuirani.

U svakoj normalnoj distribuciji određeni njezin dio se nalazi između dviju granica: aritmetičke sredine i neke z -vrijednosti. Na Slika 3. 1. je označeno da se između aritmetičke sredine i $z=+1$ (tj. rezultata koji je za 1 standardnu devijaciju veći od aritmetičke sredine) nalazi 34,13 % rezultata, pa je spomenuta površina 0,3413. Površina između $-1z$ i $+1z$ iznosi $0,3413 \times 2 = 0,6826$ od čitave površine. Drugim riječima, rezultati između $-1z$ i $+1z$ obuhvaćaju nešto više od dvije trećine rezultata. Područje između $-2z$ i $+2z$ zauzima preko 95%, a ono između $-3z$ i $+3z$ gotovo čitavu površinu (99,73%).



Slika 3. 1. Neke površine ispod normalne krivulje

t – *distribucija* (ili Studentova distribucija) je slična normalnoj, tj. i ona je simetrična i ima „zvonašt“ oblik, ali je ona pri dnu šira od normalne distribucije. Drugim riječima, kod t -distribucije „krajevi“ su širi što je uzorak manji. Ovdje se kriterij mijenja prema broju rezultata, odnosno t -vrijednosti ovisne su o tzv. *stupnjevima slobode*. Stupnjevi slobode su broj rezultata smanjen za jedan, odnosno $N-1$. Kod malih uzoraka koristi se t -tablica za očitavanje granične vrijednosti t , tj. koliko puta – uz određenu veličinu uzorka - mora razlika biti veća od svoje pogreške da bismo je mogli smatrati statistički značajnom.

U t -tablici npr. piše uz 18 stupnjeva slobode na nivou značajnosti od 1% ($P=0,01$) broj 2,88. To je „granična t -vrijednost“ i znači: kada između dvije populacije ne bi zapravo postojala razlika, a mi bismo iz tih populacija uzimali uzorke, pretpostavimo, veličine $N_1 = N_2 = 10$ (10-1 plus 10-1=18) i izračunali svaki put t -vrijednost, onda bismo t , veličine 2,88 i veće mogli slučajno dobiti samo u jednom od 100 izračunavanja (1%).

3.1.5. STATISTIČKI TESTOVI RAZLIKE

U praksi često nailazimo na podatke čija nam je distribucija nepoznata ili znamo da distribucija nije normalna. One metode koje se služe mjerljivim podacima, koji se distribuiraju normalno, nazivaju se *parametrijskim metodama*. One metode kod kojih nije važno je li populacija normalno distribuirana, a mjerni rezultati su izraženi u frekvencijama nekih kvaliteta, nazivaju se *neparametrijske metode*. Neparametrijske se metode upotrebljavaju prvenstveno kod mjerljivih podataka izraženih nominalnim i ordinalnim skalama, a parametrijske metode kod rezultata izraženih intervalnim i omjernim skalama. Ako između dvije populacije postoji razlika, to ćemo obično preciznije i uspješnije ustanoviti pomoću parametrijskih nego neparametrijskih testova.

3.1.5.1 t-TEST

T-test je statistički test razlike (eng. *test of difference*) koji pomaže pri provjeri postoji li stvarna razlika između pojava koje se testiraju. Računanjem t -testa dobiju se vrijednosti "t" i "p". Za izračunavanje se koristi formula:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}} \quad (3.27)$$

gdje su N_1 i N_2 veličine uzoraka, $\overline{X}_1, \overline{X}_2$ aritmetičke sredine uzoraka, te σ_1^2, σ_2^2 varijance uzoraka. Stupnjevi slobode se računaju po formuli:

$$v = \frac{\left(\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}\right)^2}{\frac{\left(\frac{\sigma_1^2}{N_1}\right)^2}{N_1 - 1} + \frac{\left(\frac{\sigma_2^2}{N_2}\right)^2}{N_2 - 1}} \quad (3.28)$$

Prilikom ispitivanja nul-hipoteze mogu se pojaviti dva tipa grešaka (Tablica 3. 4.):

- Pogreška tipa 1 – uz blaži kriterij izlažemo se riziku da proglasimo da se dvije aritmetičke sredine razlikuju, a zapravo među aritmetičkim sredinama populacije nema razlike. Odbacuje se nul-hipoteza koja je zapravo točna.
- Pogreška tipa 2 – uz stroži kriterij izlažemo se riziku da neku razliku ne proglasimo značajnom (tj. smatramo da dvije aritmetičke sredine pripadaju istoj populaciji) iako među populacijama razlika postoji. Nije se uspelo odbaciti nul-hipotezu koja nije točna.

Tablica 3. 4. Pogreške pri zaključivanju iz uzorka na populaciju

Odluka	Stanje u populaciji	
	Nema razlike između dvije aritmetičke sredine	Postoji razlika između dvije aritmetičke sredine
Odbacujemo nul-hipotezu	<i>Pogreška tipa 1</i>	<i>Nema pogreške</i>
Prihvaćamo nul-hipotezu	<i>Nema pogreške</i>	<i>Pogreška tipa 2</i>

Ispitivanje razlika između standardnih devijacija uzoraka ispituje se pomoću tzv. F-testa po formuli:

$$F = \frac{\text{veće } s^2}{\text{manja } s^2} \quad (3.29)$$

Ako su uzorci jednako veliki, nije ni potrebno testirati F-testom da li im se varijance statistički značajno razlikuju ili ne, jer će se – ako se one razlikuju – pogreška u računu biti neznatna. Ako se te grupe razlikuju u veličini, onda treba ispitati statističku značajnost razlike između aritmetičkih sredina malih uzoraka.

3.1.5.2 ANALIZA VARIJANCE - ANOVA

Ovaj snažan parametrijski test omogućava nam da jednim ispitivanjem, bez obzira na broj grupa, utvrdimo da li postoji značajna razlika u nekoj varijabli, što je znatno povoljnije nego kad koristimo t-test za svaki od mogućih parova grupa. Najprije se utvrdi da li srednje vrijednosti uzoraka variraju od srednje vrijednosti populacije više nego se očekuje s obzirom na varijance pojedinih slučajeva u odnosu na istu srednju vrijednost. Tek ako se pokaže značajna razlika, primjenjuju se t-testovi za one grupe u kojih su razlike uočljivo najveće. Analiza varijance

upotrebljava se onda kada se želi utvrditi postoje li razlike između nekoliko aritmetičkih sredina, odnosno treba dokazati je li varijabilitet među grupama veći od varijabiliteta unutar grupa. Ako je on statistički značajno veći, onda su to zaista grupe koje ne pripadaju istoj populaciji ili različitim populacijama, ali s jednakom aritmetičkom sredinom.

Analiza varijance se zapravo sastoji u tome da se varijabilitet svih dobivenih rezultata rastavi na dijelove od kojih je sastavljen, tj. na interni varijabilitet unutar svake pojedine grupe rezultata i na varijabilitet između pojedinih grupa. Iz odnosa tih dvaju varijabiliteta može se zaključiti radi li se o grupama koje su među sobom različite (dakle ne pripadaju istoj populaciji) ili su njihove razlike možda samo slučajne, pa sve (grupe) potječu iz iste populacije.

Prema položaju nekog rezultata u masi drugih rezultata i drugih grupa, možemo ustanoviti da se njegovo odstupanje od zajedničke aritmetičke sredine, odnosno njegov varijabilitet, može podijeliti na dva dijela: na odstupanje tog rezultata od vlastite aritmetičke sredine, tj. od aritmetičke sredine grupe kojoj taj rezultat pripada (varijabilitet unutar grupe), te na odstupanje aritmetičke sredine kojoj pripada taj rezultat od zajedničke aritmetičke sredine (varijabilitet među grupama).

Ako odstupanja pojedinih rezultata od totalne aritmetičke sredine kvadriramo (kvadrirane devijacije) i te kvadrate sumiramo, ustanovit ćemo da se ta suma (totalna suma kvadrata SS_{tot}) može rastaviti na dvije „podsume“ kvadrata: na sumu kvadrata unutar grupe (SS_{wg}), te na sumu kvadrata između grupa (SS_{bg}). Spomenuti kvadrati računaju se po formulama:

$$SS_{tot} = \left(\sum X^2 \right) - \left(N_{tot} \cdot \overline{X}_{tot}^2 \right) \quad (3.30)$$

$$SS_{bg} = \left[\sum \left(N_g \overline{X}_g^2 \right) \right] - \left(N_{tot} \overline{X}_{tot}^2 \right) \quad (3.31)$$

$$SS_{wg} = SS_{tot} - SS_{bg} \quad (3.32)$$

Same sume kvadrata unutar i između grupa nisu dovoljne za ocjenjivanje je li veći varijabilitet unutar ili između grupa, jer veličina sume kvadrata ovisi o broju rezultata. Zato je SS neprikladna vrijednost za ocjenu varijabiliteta, nego treba uzeti varijancu σ^2 koja će se kod analize varijance nazivati *srednjim kvadratom* (MS).

Stupnjevi slobode za sumu kvadrata unutar grupa izračunavaju se tako da od ukupnog broja rezultata uzmemo broj grupa (k), dakle

$$df_{wg} = N_{tot} - k \quad (3.33)$$

Stupnjevi slobode za sumu kvadrata između grupa računaju se tako da se od broja grupa oduzme 1, dakle

$$df_{bg} = k - 1 \quad (3.34)$$

Jednako kao što se kod t-vrijednosti iz t-tablice očitava koliko najmanje puta mora razlika biti veća od svoje pogreške, da bi bila značajna na razini značajnosti od 5% ili 1%, jednako tako i kod analize varijanci uz pomoć Snedecorovih F tablica možemo ustanoviti koliko najmanje puta mora varijabilitet među grupama biti veći od varijabiliteta unutar grupa da bi razlika između oba varijabiliteta MS_{wg} i MS_{bg} bila statistički značajna. Računamo po formuli:

$$df_{wg} = N_{tot} F = \frac{MS_{bg}}{MS_{wg}} - k \quad (3.35)$$

F-tablica se čita tako da se stupnjevi slobode brojnika čitaju na gornjem rubu tablice, a stupnjevi slobode nazivnika na njezinom lijevom rubu. Kompletan pregled računa analize varijance obično se postavlja u tablicu analize varijance (Tablica 3. 5.).

Tablica 3. 5. Tablica analize varijance

Izvor varijabiliteta	Suma kvadrata (SS)	Stupnjevi slobode (df)	Srednji kvadrat (MS)	F
Između grupa				
Unutar grupa				
Total				

Za razliku od normalne distribucije, distribucija F-omjera nije simetrična, već je manje ili više asimetrična nadesno: ni jedan F-omjer ne može biti manji od nule, dok neki ekstremni mogu biti i vrlo visoki.

Ako analiza varijance pokaže da možemo smatrati da svi uzorci potječu iz iste populacije (F nije statistički značajan), onda nas ne zanimaju pojedinačne razlike između nekih aritmetičkih sredina. Ako F-test odbaci nul-hipotezu, zanima nas koji se uzorci među sobom statistički značajno razlikuju.

Scheffea-lov postupak za izračunavanje razlika među aritmetičkim sredinama nakon završenog F-testa:

1. Nakon izračunatog F-omjera, za svaki par aritmetičkih sredina, koje želimo usporediti, primijeniti sljedeću formulu:

$$F = \frac{(\bar{X}_a - \bar{X}_b)^2}{MS_{wg} \frac{N_a + N_b}{N_a \cdot N_b}} \quad (3.36)$$

2. Iz F-tablice očitamo granični F uz željenu razinu značajnosti (preporuča se $P=0,01$), za $(k-1)$ i $(N-k)$ stupnjeva slobode.
3. Očitana granična vrijednost F pomnoži se sa $(k-1)$ i tako se dobije nova granična vrijednost F'.
4. Za sve parove aritmetičkih sredina, za koje to želimo, izračunamo F prema (42) i usporedimo ga s F'. Ako je F veći od F', razliku možemo smatrati statistički značajnom.

3.1.6. KORELACIJA

Poznato je da mnoge pojave pokazuju neku manju ili veću međusobnu zavisnost: kada je jedna pojava u porastu, u porastu (ili u padu) je i druga, pa tako do određene granice možemo iz neke vrijednosti u jednoj od tih pojava zaključivati na odgovarajuću vrijednost u drugoj pojavi. Za upoznavanje suštine svake pojave neophodno i upoznavanje međusobne povezanosti između više pojava, tj. više varijabli. Potrebno je upoznati se sa specifičnim postupcima za određivanje nivoa povezanosti, odnosno odnosa među pojavama – *korelacijom*.

Korelacija je termin koji se odnosi na povezanost među varijablama koju je moguće kvantitativno izraziti tj. računskim postupkom se dobije visina stupnja povezanosti ili *koeficijent korelacije*. Npr. ako za varijable odaberemo kvalitetu nastave i ukupan uspjeh učenika, onda porast u kvaliteti nastave prati i porast u uspjehu učenika, te se dobije pozitivni koeficijent korelacije. Kad se želi ispitati povezanost dvije varijable, onda se prije izračunavanja korelacije, prvo rezultati prikažu grafički tzv. dijagramom rasipanja (eng. scatter), pa tek onda, ako je povezanost u cjelini manje-više linearna, izračunava se koeficijent korelacije.

Kada bi povezanost između dvije varijable bila besprijeekorna, svaki ispitanik bi bio u obje varijable na jednakim mjestima, tj. koliko je u jednoj varijabli iznad (ili ispod) prosjeka, upravo toliko morao bi biti i u drugoj varijabli (morale bi biti iste z-vrijednosti). Veličina razlika između sparenih z-vrijednosti ovisi o visini povezanosti između obje varijable: kad je stupanj povezanosti maksimalan, razlike nema, što je povezanost slabija, razlike između z-vrijednosti su veće. Koeficijent korelacije r (Paersonov koeficijent korelacije) računa se po formuli:

$$r = 1 - \frac{1}{2} \frac{\sum (z_x - z_y)^2}{N - 1} \quad (3.37)$$

ili po ekvivalentnoj formuli:

$$r = \frac{\sum (z_x \cdot z_y)}{N - 1} \quad (3.38)$$

Računanje koeficijenta korelacije prema gornjim formulama bili bi dugotrajno u slučajevima kada bi N bio veći i kada aritmetičke sredine i standardne devijacije ne bi bile cijeli brojevi. Skraćeni postupak za izračunavanje koeficijenta r iz rezultat koji nisu grupirani u razrede. Formula glasi:

$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X^2)] \cdot [N \sum Y^2 - (\sum Y^2)]}} \quad (3.39)$$

gdje je $\sum XY$ = suma umnožaka pojedinih parova rezultata, N = broj parova, $\sum X^2$ i $\sum Y^2$ = suma kvadriranih rezultata varijable X i varijable Y .

Da bismo ustanovili je li dobiveni koeficijent korelacije r značajan (tj. razlikuje li se značajno od nule), računamo t prema formuli:

$$t = r \cdot \frac{\sqrt{N-2}}{\sqrt{1-r^2}} \quad (3.40)$$

gdje je N broj parova, a N-2 broj slučajeva slobode.

Jedna od čestih pogrešaka u interpretaciji koeficijenta korelacije sastoji se u tome da se visina koeficijenta korelacije interpretira kao postotak zajedničkih faktora, odnosno ako korelacija iznosi 0,70, smatra se da u obje varijable postoji 70% zajedničkih faktora. To nije točno. Približno možemo odrediti količinu zajedničkih faktora uz pomoć tzv. koeficijenta determinacije, tj. kvadriranjem koeficijenta korelacije.

Koeficijent korelacije može imati vrijednosti između -1 i 1, gdje -1 označava potpunu negativnu korelaciju, 0 označava da korelacija ne postoji, a 1 označava potpunu pozitivnu korelaciju. Ostale vrijednosti iz intervala se nazivaju djelomičnim pozitivnim ili negativnim korelacijama (ovisno o predznaku). Primjer interpretacije koeficijenata:

- do ± 0.20 – neznatna korelacija, gotovo ne postoji
- od ± 0.20 do ± 0.40 – niska korelacija, povezanost je mala
- od ± 0.40 do ± 0.70 – umjerena korelacija, bitna povezanost
- od ± 0.70 do ± 0.90 – visoka korelacija, izrazita povezanost
- preko ± 0.90 – veoma visoka korelacija, veoma uska povezanost

Sama činjenica da između dvije pojave postoji korelacija ne znači da su te dvije pojave uzročno povezane. Uzročna veza između dvije pojave može se definitivno dokazati jedino eksperimentom.

Često se želi dati veća važnost nekim rezultatima nego drugima, jer su neki rezultati „važniji“ od nekih drugih. Da bismo doznali koji su rezultati važniji, koristimo poseban oblik korelacije tzv. *multipla korelacija*. Multipla korelacija je maksimalno mogući koeficijent korelacije između dva ili više prediktora (npr. nekoliko testova) i jednog kriterija (uspjeh u školi). Tu maksimalno visoku korelaciju možemo ostvariti samo ako veću „težinu“ damo važnijim prediktorima. Formula multiple korelacije između kriterija (0) i dva prediktora (1 i 2) glasi:

$$R_{012} = \sqrt{\frac{r_{01}^2 + r_{02}^2 - 2r_{01} \cdot r_{02} \cdot r_{12}}{1 - r_{12}^2}} \quad (3.41)$$

Optimalne „težine“ koje se daju prediktorima nazivaju se „beta-koeficijenti“ i računaju se, za svaki od dva prediktora, po formuli:

$$\begin{aligned} \beta_1 &= \frac{r_{01} - r_{02} \cdot r_{12}}{1 - r_{12}^2} \\ \beta_2 &= \frac{r_{02} - r_{01} \cdot r_{12}}{1 - r_{12}^2} \end{aligned} \quad (3.42)$$

Dobivene beta-koeficijente podijelimo sa standardnom devijacijom odgovarajućeg prediktora te dobiveno pomnožimo s rezultatima prediktora. Dobivene rezultate zbrojimo i dobit ćemo

rezultate koji će s kriterijem imati maksimalno moguću korelaciju. Na taj način možemo rangirati rezultate prema „najispravnijim“ sumama rezultata iz oba prediktora.

Značajnost koeficijenta multiple korelacije može se testirati uz pomoć F-testa.

$$F = \frac{R^2}{1-R^2} \cdot \frac{N-k-1}{k} \quad (3.43)$$

gdje je R = koeficijent multiple korelacije, N = broj opažanja, k = broj prediktora.

Često treba ispitati korelaciju između jedne kontinuirane varijable (npr. bodovi u testu) i jedne dihotomne varijable, tj. takve koja se dijeli u dvije jasno odvojene kategorije (npr. prošao-pao). Koeficijent korelacije ove vrste zove se „point – biserijalni r“, i piše se r_{pb} . Prvo se treba „numerirati“ dihotomna varijabla i to tako da se jednoj karakteristici da jedan, a drugoj karakteristici drugi broj (npr. 0 i 1). Zatim se računa Pearsonov koeficijent korelacije ili se koristi formula:

$$r_{pb} = \frac{N \sum Y_1 - N_1 \sum Y}{\sqrt{N_1 N_0 [N \sum Y^2 - (\sum Y)^2]}} \quad (3.44)$$

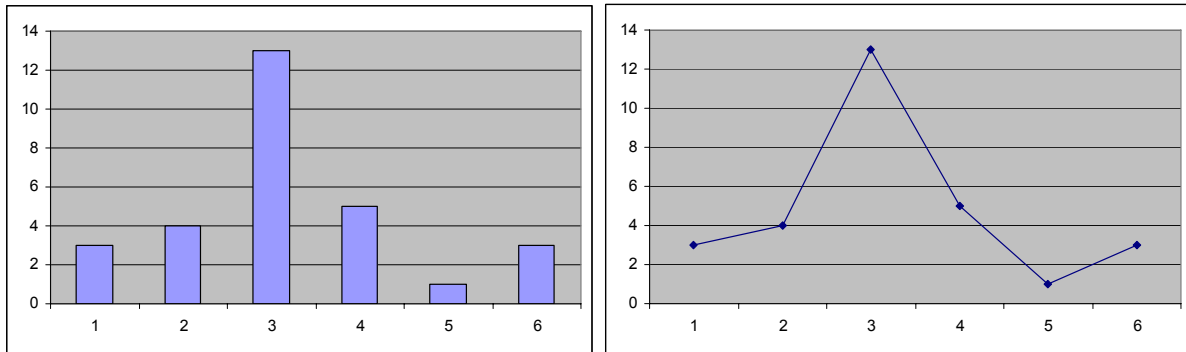
gdje je X = dihotomizirana varijabla, Y = kontinuirana varijabla, $\sum Y_1$ = suma Y vrijednosti vezanih uz vrijednost X varijable označene brojem 1, $\sum Y$ = suma svih Y vrijednosti, $\sum Y^2$ = suma svih kvadriranih Y vrijednosti, N_1 = broj svih opažanja kod kojih je X = 1, N_0 = broj svih opažanja kod kojih je X = 0, N = ukupni broj svih slučajeva (= $N_0 + N_1$).

3.1.7. GRAFIČKO PRIKAZIVANJE REZULTATA

U statistici često bez slike ne možemo dobiti pregled nad rezultatima. Ona nam katkad otkrije neke osnovne karakteristike rezultata. Ako preskočimo grafičko prikazivanje rezultata, izlažemo se dvostrukom riziku. Naime, postoji šansa da uopće nećemo uočiti neku posebnu i neočekivanu karakteristiku rezultata koju je gotovo nemoguće uočiti prilikom rada s brojevima, a isto tako i neke teške računске pogreške mogu ostati neotkrivene (npr. pogrešno stavljen decimalni zarez). U daljnjem će se razmatranju prikazati tek neki oblici grafičkog prikazivanja.

3.1.7.1 HISTOGRAM I POLIGON FREKVENCIJA

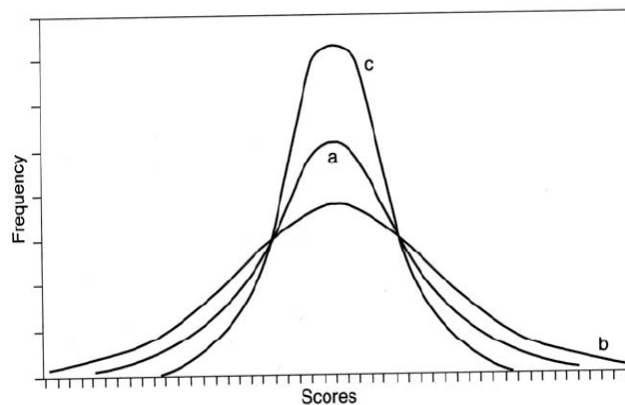
Ako su podaci grupirani u razrede, onda se oni mogu prikazati grafički *poligonom frekvencija* te *histogramima*. Frekvencije se nanose na ordinati tako da se označavaju na apscisi kod razrednih sredina odgovarajućih razreda. Ako na apscisu nanosimo razrede, onda crtamo histogram, a ako nanosimo sredine razreda, onda crtamo poligon frekvencija. Poligoni frekvencija i histogrami poprimaju razne oblike jer su oni grafički odrazi distribucije frekvencija (Slika 3. 2.).



Slika 3. 2. Histogram i poligon frekvencija

Histogram je najtočniji prikaz distribucije frekvencije nekih rezultata, ali se u praksi češće koristi poligon frekvencija. Glavni razlog tomu je činjenica da je poligon obični pregledniji način prikazivanja (krivulja), a osobito je pogodan ako na istoj slici želimo prikazati dvije ili više distribucija kako bi ih se moglo što zornije usporediti.

Poligon frekvencija koji je simetričan i zvonolik se naziva *krivuljom normalne distribucije* ili Gaussovom krivuljom. Osim simetričnih distribucija frekvencija, u statističkoj praksi vrlo često se susreću i distribucije frekvencija bez intenzivnijeg grupiranja oko srednje frekvencije (oblika slova J, slova U, ili drugih oblika). Kod asimetričnih distribucija odstupanja se mogu pojaviti u simetričnosti stranica (vrh poligona, odnosno razred s najvećom distribucijom, je pomaknut od sredine na jednu ili drugu stranu), u strmini stranica (tzv. kurtičnost) te u broju vrhova krivulje (tzv. modalnost: unimodalna, bimodalna, trimodalna distribuciju) (Slika 3. 3.).

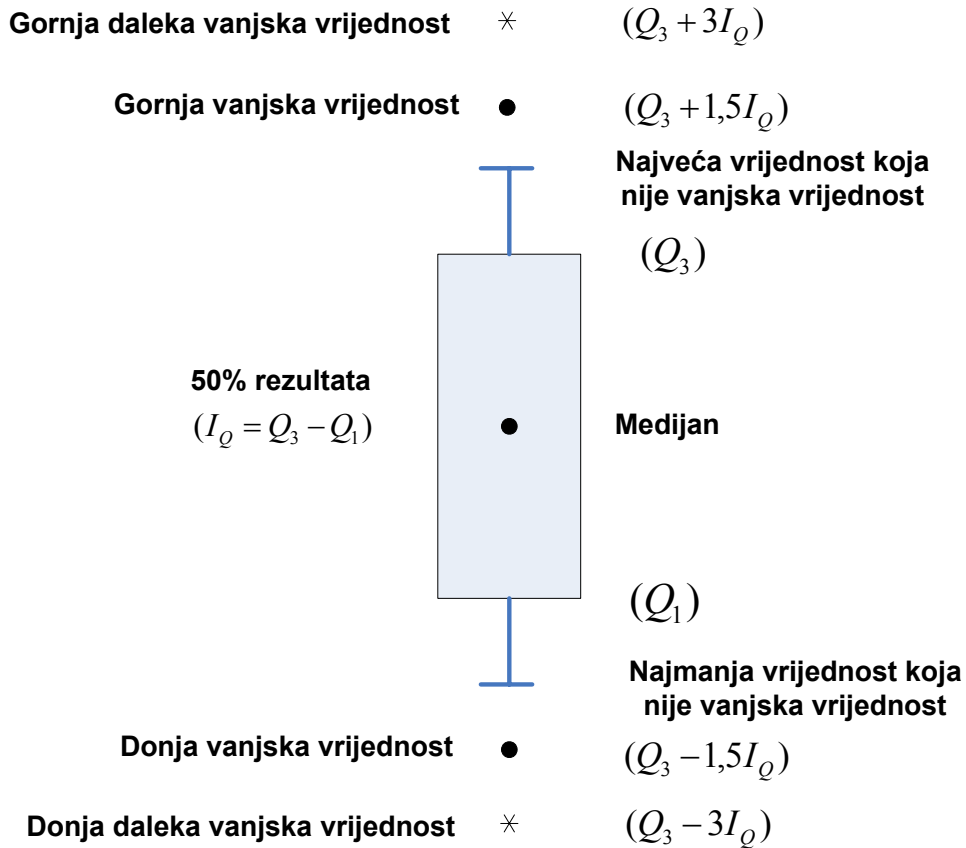


Slika 3. 3. Oblici krivulja distribucije: a) mezokurtična, b) šiljasta, c) spljoštena

Iz histograma i poligona frekvencija izravno se može očitati jedino dominantna vrijednost (mod): to je u histogramu razred s najvišim stupcem, a u poligonu razred iznad kojeg se nalazi vrh krivulje.

3.1.7.2 BOX AND WHISKER PLOT

Box-and-whisker-plot (Slika 3. 1.) dizajnirao je 1970. godine John Tukey. Graf prikazuje položaj medijana, interkvartila, te „outliere“ (vanjske vrijednosti) i ekstreme (daleke vanjske vrijednosti). Prikladan je kada želimo prikazati istu varijablu kod različitih uzoraka.



Slika 3. 4. Box and Whisker plot

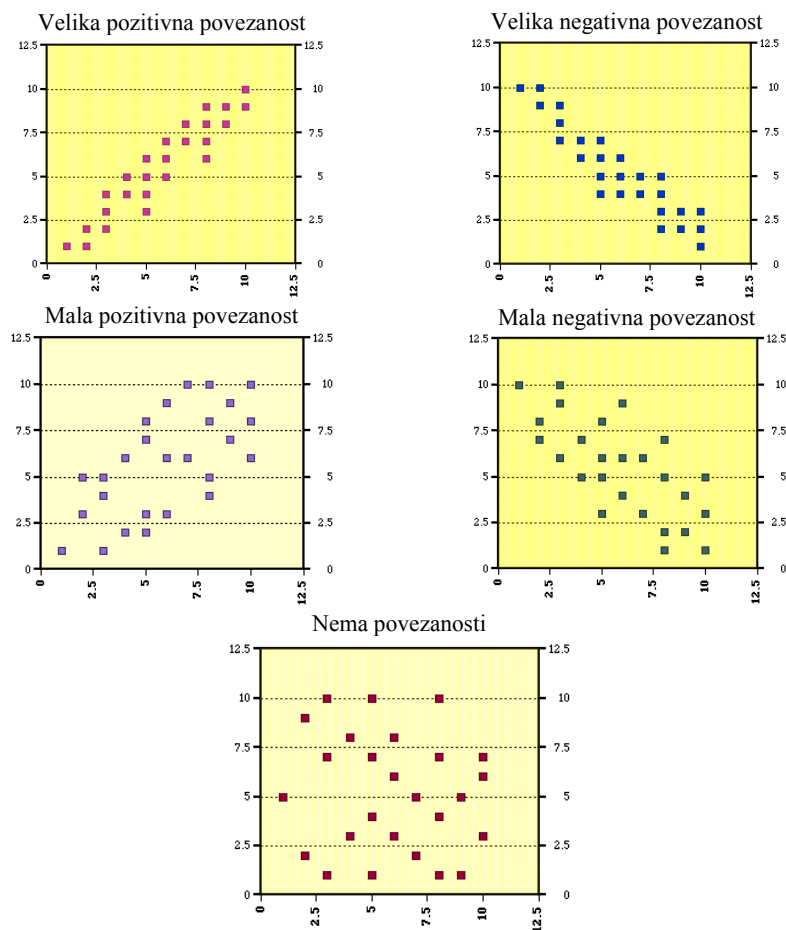
3.1.7.3 DIJAGRAM RASIPANJA

Ponekad želimo vidjeti kako jedna varijabla utječe na drugu varijablu. Dobar način za vizualno prikazivanje utjecaja promjene jedne varijable na drugu je *dijagram rasipanja* (eng. scatter). Njime se lako vide uzročno-posljedične veze između varijabli te se lako vidi smjer (je li pozitivna, negativna, itd.) i jačina te veze. Nezavisna varijabla (ono što možemo samovoljno mijenjati) se nalazi na apscisi, a zavisna (ono što se istraživanjem želi ustanoviti, npr. količina naučenog gradiva) na ordinati. Na sjecištu vrijednosti za obje varijable stavlja se točka.

Struktura podataka prikaza ovim grafom može biti pozitivna, negativna ili neutralna. Pozitivna veza među varijablama prikazana je elipsom točaka koje se idu na gore pokazujući da povećanje nezavisne varijable povećava i zavisnu varijablu. Negativna veza među varijablama prikazana je

elipsom točaka koje se idu na dolje pokazujući da povećanje nezavisne varijable smanjuje zavisnu varijablu. Dijagram s nakupinom točaka za koje je nemoguće utvrditi da li idu na gore ili na dolje pokazuje da nema povezanosti između dvije varijable.

Struktura podataka, bilo pozitivna ili negativna, pokazuje jačinu povezanosti dvije varijable prema gustoći nakupljenih točaka. Što su više točke zgusnute prema nekoj ravnoj liniji, to je veza jača (Slika 3. 5.).



Slika 3. 5. Oblici dijagrama rasipanja

3.2. METODE VREDNOVANJA SUSTAVA E-UČENJA

Informacijska i komunikacijska tehnologija (ICT) prodire u sva područja ljudskog djelovanja pa tako i u sustav obrazovanja. Upravo taj ulazak ICT-a u svijet obrazovanja je rezultirao sveprisutnom pojavom e-učenja, koje predstavlja upravo presjek ta dva svijeta. Svrha vrednovanja sustav e-učenja je ispitivanje na koji način uporaba ICT-a podupire proces učenja i poučavanja. Posebno je važno naglasiti da sva programska podrška koja se koristi u obrazovne svrhe treba biti vrednovana prije korištenja u procesu učenja i poučavanja.

Jedna od definicija vrednovanja je da vrednovanje pruža informacije za donošenje odluka o proizvodu ili procesu [PHIL2002]. Guba i Lincoln (prema [ALEX1994]) definiraju vrednovanje kao proces opisivanja onog što se vrednuje, kao i prosuđivanje o njegovoj vrijednosti. Dobro oblikovano vrednovanje trebalo bi dati dokaz da li je određeni pristup uspješan i da li ima potencijalnu vrijednost za druge [DEMP2004].

Prvi korak u procesu vrednovanja učinka je definiranje pitanja na koje želimo dobiti odgovor, npr. "Koliko sustavi e-učenja utječu na proces učenja i poučavanja?". O izboru pitanja ovisi izbor metodologije vrednovanja jer ne postoji jedinstven model za vrednovanje učinka sustava e-učenja. Zatim se pristupa definiranju nul-hipoteza koje se moraju moći provjeriti da bi ih se moglo potvrditi ili odbiti na temelju određenih uvjeta i rezultata.

Slijedi definiranje metodologije za vrednovanje kojom će se ispitati istinitost nul-hipoteza. Eksperimentalno istraživanje je prikladno za vrednovanje sustava e-učenja jer omogućava ispitivanje odnosa između poučavanja i učenikovih rezultata te dobivanje kvantitativnih mjera značajnosti tih odnosa. Nadalje, pristupa se provođenju istraživanja te analiziranju podataka. U idealnim okolnostima, ako dobiveni rezultati ne potvrde nul-hipotezu, istraživači bi trebali objasniti razloge zbog kojih su dobiveni rezultati takvi kakvi jesu.

Općenito, kvaliteta inteligentnog sustava e-učenja može se odrediti uspoređivanjem tog sustava s tradicionalnim procesom učenja i poučavanja, kao i pojednostavljenom verzijom samog sustava (koju dobijemo tako da uklonimo sve inteligentne komponente), koristeći inicijalna i završna testiranja na stvarnim ili simuliranim učenicima. Simulirani učenici su računalni modeli (simulacije) živog učenika. Prednost korištenja simuliranih učenika je u većoj kontroli postavki eksperimenta, ali oni bi trebali predstavljati tek prvi korak prema potpunom vrednovanju jer su simulirani učenici previše pojednostavljeni naprema živim učenicima. Živi učenici daju realnije ponašanje i smanjuju troškove vrednovanja, ali se mogu manje kontrolirati [DEVE2003].

U ovom poglavlju objašnjavaju se različiti pristupi vrednovanju sustava e-učenja, daje se pregled i klasifikacija metoda vrednovanja, objašnjavaju se postupci i instrumenti prikupljanja podataka, kao i karakteristike tih instrumenata, te se posebno objašnjava eksperiment kao metoda vrednovanja.

3.2.1. MODELI VREDNOVANJA

Postoji velik broj različitih pristupa vrednovanju koji su razvijeni u posljednjih 50 godina i vrednovanje sustava za e-učenje stoga stremi pojednostavljivanju i definiranju jedinstvenog modela vrednovanja. Stoga različiti autori pristupaju klasificiranju modela vrednovanja ističući njihove prednosti i nedostatke.

Alexander i Hedberg [ALEX1994] su prikazali četiri reprezentativna modela: vrednovanje zasnovano na ciljevima, vrednovanje zasnovano na odlukama, vrednovanje zasnovano na vrijednostima, te naturalističko vrednovanje, zajedno s njihovim prednostima i nedostacima (Tablica 3. 6.).

Tablica 3. 6. Modeli vrednovanja (prema [ALEX1994])

Model vrednovanja	Opis	Prednosti	Nedostaci
Zasnovan na ciljevima (eng. Objectives-based)	Vrednovanje kao proces određivanja stupnja ostvarenja obrazovnih ciljeva. Sumativni model.	Ideje su sistematične, točne. Pre-post dizajn. Lagano se primjenjuje.	Ne smatra učenike individuama s različitim sposobnostima, iskustvima i interesima.
Zasnovan na odlukama (eng. Decision based)	Fokusira se na odluke koje se donose tijekom razvoja i poboljšanja koja bi se mogla napraviti.	Obraća pažnju na kritične kontekstualne komponente. Fokusira se na potrebe sudionika.	Teško se provodi uz velika financijska sredstva.
Zasnovan na vrijednostima (eng. Values—based)	Vrednovanje vodi računa ne samo da li su postignuti ciljevi, već i da li su ti ciljevi vrijedni postizanja.	Priznaje važnost neplaniranih rezultata. Omogućava vrednovanje sustava bez potrebe da se poznaju njegovi ciljevi.	Ostavlja neka pitanja neodgovorenima, npr. kako donositi zaključke?
Naturalistički (eng. Naturalistic)	Organizira vrednovanje oko ključnih pitanja sudionika. Koristi kvalitativno prikupljanje podataka kao što su promatranje i razgovori.	Učenici imaju korist od rezultati vrednovanja	Postoji mogućnost da će sudionici identificirati kriterij koji ima malu obrazovnu vrijednost.

Draper [DRAP1996] navodi sljedeće modele vrednovanja:

1. Formativno vrednovanje

Formativno vrednovanje se radi u svrhu mijenjanja sustava da bi on mogao riješiti svaki problem. Vrednovanje mora otkriti postojanje problema, ali i sugerirati koje promjene treba napraviti.

2. Sumativno vrednovanje

Sumativno vrednovanje se radi nakon završetka implementacije i pokazuje da se sustav može koristiti te daje detaljan opis njegovih performansi.

3. Iluminativno vrednovanje

Iluminativno vrednovanje je pojam kojeg su 1972. uveli Parlett i Hamilton (prema [DRAP1996]) da bi označili pristup koji se temelji na etnografskim, a ne na eksperimentalnim metodama. Cilj mu je otkriti koji su faktori važni učenicima u nekoj situaciji.

4. Integrativno vrednovanje

Integrativno vrednovanje se radi u svrhu poboljšanja procesa učenja i poučavanja i to boljom integracijom nastavnih sadržaja u sveukupnu situaciju.

5. Vrednovanje za ispitivanje kvalitete (eng. *quality audit, assessment, or assurance*)
Daje bolju sliku o postignutoj kvaliteti sustava nego samo integrativno vrednovanje.

Iqbal, Oppermann, Patel i Kinshuk [IQBA1999] govore o sljedećim modelima vrednovanja:

1. Interno vrednovanje
Interno vrednovanje se odnosi na vrednovanje neke komponente sustava.
2. Eksterno vrednovanje
Eksterno vrednovanje se odnosi na vrednovanje čitavog sustava.
3. Eksploratorno vrednovanje
Eksploratorno vrednovanje zahtijeva dubinsku analizu sustava u prirodnom okruženju koristeći višestruke izvore podataka.
4. Eksperimentalno vrednovanje
Eksperimentalno vrednovanje koristi eksperimente za sistematično mijenjanje nezavisnih varijabli i istovremeno mjerenje zavisnih varijabli. Nužno je slučajnim odabirom podijeliti sudionike u kontrolnu i eksperimentalnu grupu kao i potreba da te grupe ne budu statistički značajno različite.

Twidale [TWID1993] govori o kontroliranom i neformalnom vrednovanju.

1. Kontrolirano vrednovanje
Kontrolirano vrednovanje pruža najkorisnije informacije o kvaliteti, produktivnosti ili učinkovitosti sustava. Ograničenja kontroliranog vrednovanja su činjenica da su precizni eksperimenti veliki, spori i skupi, kontrolirani eksperiment mjeri samo jednu stvar, osim sveukupnog učinka sustava, bilo bi dobro znati i koje značajke sustava pridonose individualizaciji poučavanja, kombinacija različitih komponenti sustava može dovesti do rezultata različitih od onih kada se komponente koriste pojedinačno, sučelje sustava može imati važan učinak na sam proces učenja i poučavanja. Kada učenik započinje koristiti sustav mora usvojiti područno znanje i naučiti se koristiti sustavom.
2. Neformalno vrednovanje
Neformalno vrednovanje se koristi kada se zna da je sustav nedovršen i formalna procjena sveukupnih performansi je neprikladna. Ono se može primijeniti i na završen sustav. Neformalno vrednovanje opisuje slabosti sustava s preporukama za daljnji razvoj.

Kod neformalnog vrednovanja naglasak je više na dokazivanju negativnog učinka, nego pozitivnog. Dakle, osnovna razlika između neformalnog i kontroliranog vrednovanja je u tome što ako želite pokazati prednosti sustava, onda koristite kontrolirano vrednovanje, a inače neformalno.

Alexander i Hedberg [ALEX1994] su prikazali faze razvoja sustava e-učenja i modele vrednovanja zajedno s njihovom svrhom, pitanjem i metodama, koji su prikladne za pojedinu aktivnost (Tablica 3. 7.). Jedna od navedenih faza razvoja sustava e-učenja je učenje i poučavanje na tom sustavu kroz koju se u kontekstu sumativnog vrednovanja pristupa ispitivanju učinka dotičnog sustava e-učenja.

Tablica 3. 7. Odnos faza razvoja sustava i vrednovanja sustava (prema [ALEX1994])

Aktivnosti	Prikupljanje informacija	Svrha	Pitanja vrednovanja	Metode
Oblikovanje	Analizirati potrebe	Pružiti informacije potrebne za planiranje	1. Koja područja nastavnog sadržaja učeniku zadaju najviše problema? 2. Predstavlja li sustav najbolji način za rješavanje tog problema?	1a. Razgovor s iskusnim učiteljima. 1b. Pregled literature o poučavanju određenog nastavnog sadržaja. 1c. Pregled postignuća učenika. 2. Pregled učenika, drugih provedenih vrednovanja
	Formativno vrednovanje	Objaviti odluke donesene prilikom oblikovanja sustava.	3. Koja bi se obrazovna strategija trebala koristiti? 4. Da li je korisničko sučelje prikladno?	3. Pregled literature da bi se ustanovilo što je poznato o načinu na koji učenici usvajaju određene nastavne sadržaje. 4. Prelaženje svih koraka prototipa sustava s potencijalnim korisnicima i stručnjacima.
Razvoj	Formativno vrednovanje	Objaviti odluke donesene prilikom razvoja sustava.	5. Da li je korisnikova interakcija korisna i efikasna? 6. Što i kako učenici uče koristeći ovaj sustav? 7. U čemu je vrijednost ovog sustava?	5. Promatranje, snimanje, praćenje korisnika, razgovori s korisnicima i stručnjacima. 6. Stimulirano prisjećanje, kritične situacije. 7. Razgovor vršnjaka
Učenje i poučavanje	Sumativno vrednovanje	Odrediti vrijednost sustava u kontekstu njegove uporabe.	8. Koje promjene u saznanjima imaju učenici kao rezultat uporabe sustava? 9. Da li je ovaj sustav prikladan način za ispunjenje definiranih potreba?	8. Inicijalna i završna testiranja/upitnici/ razgovori. 9. Razgovor vršnjaka/stručnjaka
Institucionalizacija	Vrednovanje učinka	Odrediti prijenos učenikovih promjena koje se odnose na razumijevanje uporabe sustava.	10. Da li učenici primjenjuju nova saznanja u drugim kontekstima? 11. Kakav je učinak sustava na druge aspekte nastavnog sadržaja, učenja i poučavanja i organizacije?	10. Problemi za čije je rješavanje potrebno razumijevanje. 11a. Razgovor s sudionicima i ključnim davateljima zahtjeva. 11b. Pregled organizacije dokumenata.
	Vrednovanje održavanja	Određivanje kontekstualne vrijednosti sustava.	12. Da li je sustav valjan u trenutnom kontekstu njegove uporabe?	12. Stručni pregled

3.2.2. METODE VREDNOVANJA

Mužić [MUŽI1977] metode pedagoškog istraživanja, po kriteriju da li se ispituju ili ne uzročne veze među pojavama, dijeli na: *deskriptivne* i *kauzalne*.

Deskriptivna metoda ne obuhvaća samo prikupljanje i sređivanje podataka, nego i uspoređivanje, vrednovanje i interpretaciju podataka. Nakon toga možemo pristupiti razmatranju mogućih uzroka i povezanosti između pojava (kauzalna metoda), te raditi na poboljšanju karakteristika te pojave, odnosno, usavršiti pedagošku praksu.

Da bi se mogla koristiti deskriptivna metoda potrebno je dovoljno precizno odrediti pedagošku pojavu koja će biti predmet deskripcije. Na primjer, ako želimo ispitati znanja učenika, to znači da treba odrediti dob, razred i školu učenika, njihovo zemljopisno određenje i sl., zatim kojih znanja s kojih područja. Treba razlikovati vrijeme kada se vrši deskripcija (vrijeme promatranja) i vrijeme na koje se deskripcija odnosi (kritično razdoblje). Potrebno je definirati postupke koji se

namjeravaju primijeniti (promatranje, razgovori, pismeno ispitivanje, analiza dokumentacije) kao i instrumente opisivanja (ankete, testovi i dr.), ali i na statističku obradu podataka.

Nema pedagoškog istraživanja u kojem se ne bi mogla primijeniti deskriptivna metoda. Postoji mogućnost da se primjeni deskriptivna metoda, a ne i kauzalna, ali nije moguće ikakvo traženje uzročno-posljedičnih veza ako ne raspolažemo s deskripcijom pojava kod kojih namjeravamo ispitati njihovu uzročno-posljedičnu povezanost. Deskripcija obuhvaća uz prikupljanje, obradu i prezentaciju podataka, još i njihovu interpretaciju, izvođenje zaključaka, te ukazivanje na način kako bi se na osnovu tih rezultata moglo usavršiti odgojno-obrazovnu praksu. Sve subjektivne impresije o uzročno-posljedičnoj vezi dobivene kao rezultat deskripcije, predstavljaju hipotezu koja se prihvaća ili odbacuje tek nakon primjene kauzalne metode [MUŽI1977].

Da bi se potpunije upoznale zakonitosti neke pedagoške pojave treba istražiti uzročno-posljedične veze. Znanstvene postupke kojima se ustanovljuje ta veza sistematizirao je Mill još 1885. (Mill, prema [MUŽI1977]) u četiri pravila:

1. Pravilo *slaganja*

„Kada dva ili više slučajeva iste pojave imaju jednu jedinu zajedničku okolnost, onda je ta okolnost uzrok (odnosno posljedica) te pojave.“

2. Pravilo *diferencije*

„Ako su okolnosti u kojima se istraživana pojava pojavljuje, kao i okolnosti u kojima se ne pojavljuje, u svemu jednake osim u jednoj okolnosti, ta jedna okolnost, koja se pojavljuje samo u prvom slučaju, ta okolnost u kojoj se ta dva slučaja jedino razlikuju, jest posljedica, ili uzrok, ili neophodni dio uzroka te pojave.“

3. Pravilo *ostatka*

„Ako se od jedne pojave oduzme onaj dio koji je ranijim indukcijama utvrđen kao posljedica određenih uvjeta, onda je ostatak pojave posljedica ostalih uvjeta.“

4. Pravilo *usporednih promjena*

„Svaka pojava koja se mijenja kad god se neka druga pojava na određeni način mijenja, uzrok je odnosno posljedica te pojave, ili je bilo kako s njome uzročno-posljedično vezana.“

Postoje mnoge metode vrednovanja, ali upravo njihova različitost otežava procjenu koju metodu koristiti u određenom kontekstu, što onemogućava njihovu generalizaciju. U nastavku se daje pregled metoda vrednovanja koje se spominju u literaturi, kao i njihova klasifikacija koja bi trebala omogućiti njihov lakši odabir ovisno o pojavi koju želimo vrednovati.

3.2.2.1 PREGLED METODA VREDNOVANJA

Metode vrednovanja sustava e-učenja uključuju mnoge faktore koje treba promatrati, što čini samo vrednovanje teškom zadaćom. Sljedeće metode koriste istraživači iz područja ekspertnih sustava, obrazovanja te psihologije (prema [IQBA1999]).

1. **Dokaz ispravnosti** (eng. *proof of correctness*) ispituje da li sustav ispunjava željene zahtjeve i ciljeve ili da li postoji korespondencija između njegove strukture i ponašanja i njegovih specifikacija. Ova metoda vrednuje unutrašnje komponente sustava. Ova metoda ne odgovara za programe iz umjetne inteligencije koji rade s analitički neuhvatljivim problemima,

predstavljenim nepotpuno određenim funkcijama (Partridge, prema [MARK1993]), pa prema tome nije dobra ni za vrednovanje sustava e-učenja. Možda se može iskoristiti za komponente ili dijelove sustava koje nisu toliko vezane uz umjetnu inteligenciju.

2. **Usporedbe aditivnog eksperimentalnog dizajna** (eng. *additive experimental design comparisons*) koriste se za vrednovanje velikih komponenti sustava koje se mogu eksperimentalno promijeniti ili izbaciti. Stoga one pripadaju eksperimentalnom istraživanju i koriste se isključivo za interno vrednovanje. Ovaj pristup zahtjeva velik broj učenika i nije isplativ ako komponenta nije od velike važnosti. Prednost aditivnog dizajna je u tome što je moguće manipulirati individualnim aspektima sustava da bi se direktno odredio njihov učinak i važnost.
3. **Dijagnostička točnost** (eng. *diagnostic accuracy*) sustava se može odrediti pomoću procedura koje naglašavaju kvalitetu mikro-teorija koje koristi sustav. Ova metoda pruža kontrolirani pregled modela učenika i pripada eksperimentalnom istraživanju.
4. Direktni utjecaj **povratne informacije/kvalitete poučavanja** (eng. *feedback/ instruction quality*) na učenika može se eksperimentalno procijeniti pomoću sekvencijalnih analitičkih procedura prikladnih za interno vrednovanje. Jedan način za mjerenje ovog utjecaja je izračunati udio određenih akcija u ukupnom broju akcija koje se javljaju unutar određenog okvira koji stremi ciljanoj akciji.
5. **Analiza osjetljivosti** (eng. *sensitivity analysis*) ispituje sustav ili komponentu sustava prema tome na koji način ponašanje odgovara na različite informacije koje sustav prima. To je važno za sustave e-učenja koji bi se trebali prilagođavati različitim karakteristikama individualnih učenika. Osjetljivost sustava na te različite karakteristike bi mogla pokazati treba li sustav dodatno poboljšati. Sustav koji prikazuje slične odgovore na prilično različite informacije nije poželjan. Ako se mogu precizno odrediti mjere osjetljivosti i ako im se može dodijeliti značenje (koja su ponašanja poželjna, kada ih treba pozvati), onda se ova metoda može primijeniti za sumativno vrednovanje, a inače se primjenjuje za formativno vrednovanje [MARK1993]. Ovaj eksperimentalni pristup je prikladan kako za vrednovanje komponenti tako i za vrednovanje cijelog sustava i stoga se može koristiti i za interno i za eksterno vrednovanje.
6. **Eksperimentalno istraživanje** (eng. *experimental research*) je uobičajena metoda u psihologiji i obrazovanju. Omogućava otkrivanje veze između akcija koje se poduzimaju u poučavanju i odgovarajućih rezultata koje učenici postižu u odnosu na te akcije, a također se može odrediti i koliko su takve veze značajne, i stoga je posebno pogodan za ispitivanje učinka procesa učenja i poučavanja. Eksperiment se kod formativnog vrednovanja može koristiti za uspoređivanje učinka nekih značajki sustava. Češće se koristi za sumativno vrednovanje kod kojeg su poželjniji sveukupni zaključci, nego sakupljanje informacija. Prema [MARK1993] postoje različita eksperimentalna istraživanja kao što su eksperimenti s jednom grupom, eksperimenti s kontrolnom grupom te kvazi-eksperimentalna istraživanja. Najprije je potrebno odrediti što se želi istraživati, pa se zatim odrede hipoteze (npr. da će nakon nekog načina poučavanja ili tretmana postojati značajne razlike među grupama). Hipotezu treba biti moguće testirati, te potvrditi ili opovrgnuti na osnovi određenih uvjeta i rezultata istraživanja. Zatim se određuje oblik istraživanja kako bi istraživač mogao ispitati hipotezu. Istraživač nakon provedenog istraživanja analizira podatke. Ako rezultati ne potvrđuju hipotezu, bilo bi dobro da istraživači, ako je moguće, predlože moguća objašnjenja

za svoje rezultate. Eksperimentalno istraživanje je prikladnije za eksterno vrednovanje jer daje sveukupne zaključke.

7. **Vrednovanje proizvoda** (eng. *product evaluation*) je vrednovanje širokog opsega koje za cilj ima ispitati opravdanost daljnjeg razvoja sustava. Obrazovna učinkovitost (eng. *instructional effectiveness*) sustava e-učenja, živih učitelja i tradicionalnih metoda poučavanja mora biti uspoređena na stvarnim podacima uz sudjelovanje velikih grupa učenika. Ova metoda pripada eksperimentalnom istraživanju.
8. **Pregled stručnjaka** (eng. *expert inspection*) se koristi da bi se uvidjelo da li sustav dostiže standardni nivo performansi. Ova vrsta vrednovanja se koristi za razvoj sustava zasnovanih na znanju i prikladna je za interno vrednovanje. Ona se može klasificirati kao eksploratorna jer se ne ispituje nikakva kontrolirana hipoteza.
9. **Nivo dogovora** (eng. *level of agreement*) između stručnjaka područnog znanja se odnosi na verifikaciju, te prikladan za interno vrednovanje. Ovaj pristup koristi stručnjake i korelacijske metode da bi ispitao konzistenciju znanja i uvjerenja te stoga metoda spada u eksploratorno istraživanje.
10. Eksperiment **Čarobnjak iz Oza** (eng. *Wizard of Oz*) koristi čovjeka za simulaciju ponašanja sustava tako da taj čovjek može ispitati aspekte sustava prije same implementacije. Učenik je u interakciji sa sučeljem računala, ali informacija se predaje čovjeku na drugoj lokaciji, koji obrađuje tu informaciju i vraća je natrag putem sučelja. Ova metoda se često koristi za testiranje učinkovitosti sučelja prije razvoja unutarnjih komponenti. Može se koristiti i za testiranje unutarnjih komponenti nedovršenog sustava. U ovom slučaju čovjek bira i priprema podatke koje će koristiti testna komponenta, predaje podatke komponenti, uzima rezultate i obrađuje ih te vraća rezultate korisniku. Iako je ova metoda u suštini eksperimentalna, njoj nedostaje objektivnost konvencionalnog eksperimenta i daje manje uvjerljiviji dokaz uočenog učinka. Ova metoda je prikladna za interno vrednovanje. Teškoća u otkrivanju kako će se čovjek ponašati onemogućava donošenje odluka unaprijed, na način kako bi se parametri trebali proučavati. Ova metoda, dakle, spada u eksploratorno istraživanje.
11. **Metrike performansi** (eng. *performance metrics*) su metode koje koriste kvantitativne rezultate da bi istražile individualne faktore ili značajke sustava, pa stoga spadaju u kategoriju internog vrednovanja. One su jako važne kod eksploratornih postavki kada ima malo podataka da bi se mogli donijeti statistički značajni zaključci. Može doći do pojave lažnih pozitivnih i lažnih negativnih rezultata.
12. Metode **unutrašnjeg vrednovanja** (eng. *internal evaluation*) proučavaju odnos između arhitekture sustava i njegovog ponašanja, te obično uključuju detaljnu eksploratornu analizu akcija i struktura podataka.
13. Metoda vrednovanja **zasnovanog na kriterijima** (eng. *criterion-based*) je metoda slična metodi unutrašnjeg vrednovanja. Najbolje odgovara za ranije faze razvoja sustava, gdje se radi o općenitijim karakteristikama, a ne o preciznim detaljima ili točno određenim aspektima sustava. Preporučuje se odrediti neke općenite smjernice tj. napravi se popis zahtjeva i specifikacija koji se zatim provjeravaju kako bi se ispitali nedostaci. Sustav se smatra uspješnim ako ne pokazuje velike nedostatke unutar okoline koja je namijenjena za njegovu primjenu (Partridge, prema [MARK1993]). Osoba koja vrednuje sustav mora odrediti koji su to "veliki nedostaci", prema tome kako se čini da sustav ispunjava svoje zahtjeve i

specifikacije. Ova metoda se može koristiti za ispitivanje komponente ili cijelog sustava, pa je prikladna za interno i eksterno vrednovanje.

14. **Pilot testiranje** (eng. *pilot testing*) se koristi za ispitivanje ponašanja sustava na neočekivane rezultate do kojih može doći prilikom korištenja sustava. Ova metoda je vrlo važna za otkrivanje problema u sustavu. Ako autori sustava mogu pretpostaviti da će korisnici biti slične stručnosti i iskustva kao i oni, onda mogu izraditi sustav prema svojim potrebama i biti prilično sigurni da će to odgovarati i korisnicima. Međutim, jasno je da se korisnici mogu razlikovati prema svom znanju o tutorskom poučavanju i tutorima, te osnovnim spoznajnim sposobnostima. Stvarni sustav bi se trebao ispitati na stvarnim korisnicima kako bi se otkrili eventualni problemi i neočekivani rezultati. Gagné i Golas (prema [MARK1993]) govore o tri vrste pilot testiranja: jedan-prema-jedan testiranje, testiranje malih grupa, testiranje u stvarnim uvjetima (eng. *field testing*). Kod jedan-prema-jedan testiranja, promatraju se interakcije učenika s materijalom za poučavanje koji se razvija (otkrivaju se neodgovarajuća očekivanja; nejasni smjerovi, pitanja i informacije), a obično se provodi u ranim fazama razvoja. Testiranje malih grupa se obično provodi na malim grupama učenika koji predstavljaju ciljnu populaciju. Izvodi se u kasnijim fazama razvoja, kada se program i njegov sadržaj stabiliziraju. Testiranje u stvarnim uvjetima ispituje upotrebu sustava sa stvarnim predavačima i učenicima u stvarnom okruženju, kada je sustav gotovo završen. Nastoji se otkriti kakvi se problemi događaju kada se sustav uvodi u okolinu gdje će se zapravo koristiti (Hoecker i Elias, prema [MARK1993]).
15. **Certifikacija** (eng. *certification*) se zasniva na metodama za identificiranje od strane kompetentnih živih učitelja. Takva procedura profitira od autoritativne potpore ispravnosti sustava i stoga spada pod eksploratorno istraživanje. Prikladnost pristupa za cjelokupni sustav čini ga prikladnim za eksterno vrednovanje. Ova metoda omogućava dobivanje povratne informacije o prednostima i nedostacima sustava kao dio formativnog vrednovanja i dobivanje ocjena prikladnosti kao dio sumativnog vrednovanja. No, ova metoda nameće pitanje standarda i kriterija za vrednovanje sustava i njegovih komponenti, te točnosti kojom ljudi mogu identificirati učinkovite sustave.
16. **Vanjsko vrednovanje** (eng. *outside assessment*), tj. mišljenja stručnjaka ili velikog broja potencijalnih korisnika sustava, može biti značajno ako postoji dogovor. Kako metoda radi na cijelom sustavu i zasnovana je na proučavanju mišljenja korisnika, ona je prikladna za eksploratorno eksterno vrednovanje.
17. **Postojanje dokaza** (eng. *existence proofs*) je metoda zasniva svoje zaključke na uspješnoj implementaciji sustava i predlaže novu arhitekturu sustava. Kako se metoda odnosi na cijeli sustav, prikladna je za eksterno vrednovanje. Ova metoda je uspješna jedino ako istraživač opiše ciljeve i pretpostavke istraživanja, moguće promjene u dizajnu, te dokumentira moguća iznenađenja i neuspjehe. Korisna je na početku eksploratornog istraživanja za identifikaciju ključnih problema.
18. **Promatranje i kvalitativna klasifikacija fenomena** (eng. *observation and the qualitative classification of phenomena*) ima za cilj identificirati klase fenomena, modele i trendove u interakciji čovjeka i sustava. Ovi podaci se sakupljaju za eksploratorno istraživanje primarno kroz promatranje i to kao dio eksternog vrednovanja.
19. **Strukturne zadaće i kvantitativna klasifikacija fenomena** (eng. *structured tasks and the quantitative classification of phenomena*) je pristup koji se fokusira na skup podataka

ograničavajući ili organizirajući odgovore subjekta unutar eksploratornog istraživanja. Kvantitativni podaci su sakupljeni u planiranim situacijama unutar eksternog vrednovanja, npr. pomoću razgovora ili upitnika. Prednosti ovog pristupa su u tome što veliki i precizni podatci se mogu sakupiti uporabom planiranih zadaća.

20. **Studije usporedbe** (eng. *comparison studies*) su opća klasa metoda koje bilježe sličnosti i razlike između ponašanja ili oblikovanja promatranog sustava i nekog drugog sustava. Ove metode su prikladne za eksterno vrednovanje i primjenjuju se na cjelokupne performanse sustava u eksploratornom istraživanju.

Mark and Greer [MARK1993] spominju uz dokaz ispravnosti, metodu vrednovanja zasnovanog na kriterijima, certifikaciju, analizu osjetljivosti, pilot testiranje, eksperimentalno istraživanje kao tehnike za vrednovanje još i **znanje i ponašanje stručnjaka** (eng. *expert knowledge and behaviour*). Ova metoda se koristi kao eksplicitni standard za vrednovanje sustava (Gaschnig et al., prema [MARK1993]). Stručnjak pregledava sve relevantne aspekte površinskog ponašanja sustava tijekom formativnog vrednovanja. To je moguće samo ako je ponašanje konzistentno i predvidljivo. Ova metoda ima ograničene primjene kod vrednovanja sustava e-učenja jer im je ponašanje složeno i dinamično. Ponašanje stručnjaka se može koristiti za vrednovanje sustava e-učenja. Najpoznatiji primjer takvog vrednovanja je Turingov test (Parry i Hofmeister, prema [MARK1993]). Kako je proces učenja i poučavanja nedovoljno jasan, a živi učitelji i sustav se razlikuju u onome što nude učenicima, Turingov test neće dati pravedne rezultate.

Zanimljivo je pitanje kakav utjecaj ima odabir područnog znanja na rezultate vrednovanja [TWID1993]. Neka područna znanja prikladnija za vrednovanje sustava. Obično je odabir područnog znanja ograničen znanjem tima za razvoj sustava, jer područna znanja kao što su aritmetika, algebra, osnove programiranja i sl. ne zahtijevaju konzultacije sa stručnjacima. Postoje dvije mogućnosti:

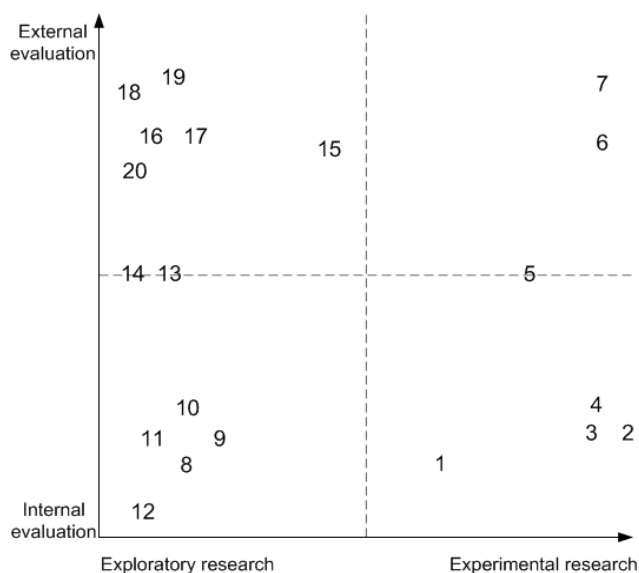
1. Odabrati područno znanje za koje postoji kontinuiran broj učenika koji će se poučavati cijele godine (npr. uvod u obradu teksta, uvod u programiranje,...).
2. Odabrati područno znanje za koje postoji grupa dobrovoljaca koja će se poučavati, a kojima to područno znanje nije obavezno.

3.2.2.2 KLASIFIKACIJA METODA VREDNOVANJA

Složeni sustavi, kao što su sustavi e-učenja, mogu se promatrati u terminima cjelovitih sustava, komponenti sustava i specifičnih značajki. Tehnike koje su prikladne za vrednovanje cjelovitih sustava ne odgovaraju u potpunosti vrednovanju komponenti ili značajki sustava i obrnuto.

S obzirom na različitost metoda vrednovanja, teško je odrediti koja je prikladna u određenom kontekstu. Proučavajući zahtjeve vrednovanja Iqbal, Oppermann, Patel i Kinshuk [IQBA1999] su uočili dva važna pitanja s kojima se susreće osoba koja želi koristiti sustav e-učenja i nakon toga ga vrednovati:

1. Što se vrednuje – cijeli sustav ili jedan njegov dio?
2. Da li ima dovoljno učenika i da li su ispunjeni ostali prikladni uvjeti da bi se provelo vrednovanje zasnovano na eksperimentu?



1. Dokaz ispravnosti 2. Usporedbe aditivnog eksperimentalnog dizajna 3. Dijagnostička točnost 4. Povratne informacije/kvalitete poučavanja 5. Analiza osjetljivosti 6. Eksperimentalno istraživanje 7. Vrednovanje proizvoda 8. Pregled stručnjaka 9. Nivo dogovora 10. Čarobnjak iz Oza 11. Metrike performansi 12. Unutrašnje vrednovanje 13. Vrednovanje zasnovano na kriterijima 14. Pilot testiranje 15. Certifikacija 16. Vanjsko vrednovanje 17. Postojanje dokaza 18. Promatranje i kvalitativna klasifikacija 19. Strukturne zadaće i kvantitativna klasifikacija 20. Studije usporedbe

Slika 3. 6. Klasifikacijski graf metoda vrednovanja [IQBA1999]

Ova dva pitanja se mogu iskoristiti za klasificiranje različitih metoda, tako da se jedna metoda može razlikovati od drugih prema stupnju prikladnosti za eksterno vrednovanje (odnosi se na cijeli sustav), odnosno interno vrednovanje (odnosi se na dio sustava) – prva dimenzija klasifikacije. Nadalje, metoda se može klasificirati prema stupnju prikladnosti za eksploratorno istraživanje, odnosno eksperimentalno istraživanje – druga dimenzija klasifikacije. Metode eksploratornog istraživanja bi se trebale primjenjivati na male uzorke učenika. Metode eksperimentalnog istraživanja se koriste manipulacijom varijabli i zahtijevaju statistički značajne grupe.

Na Sliku 3. 6. je prikazana klasifikacija metoda u odnosu na ove dvije dimenzije. Metode koje se nalaze na donjoj lijevoj strani grafa su prikladne za vrednovanje komponenti sustava i ne zahtijevaju velike uzorke učenika ili rigorozne statističke analize. Metode za vrednovanje komponenti koje zahtijevaju rad s kontrolnom grupom i statističke procedure, nalaze se na gornjoj desnoj strani grafa. Gornji dio grafa prikazuje metode koje vrednuju cjelokupne performanse sustava. Metode na desnoj strani zahtijevaju statističku analizu, dok metode na lijevoj strani ne zahtijevaju.

Metoda vrednovanja ovisi o svrsi istraživanja, tj. prvo, treba usporediti koliko je učinkovit sustav u odnosu na tradicionalni način poučavanja i drugo, treba odrediti značajke sustava koje su važne za vrednovanje sustava. Ako želimo vrednovati učinak sustava e-učenja, onda vrednujemo cijeli sustav, a ne samo jedan njegov dio, pa se radi o eksternom vrednovanju. Usporedba tehnika poučavanja zahtjeva testiranje hipoteze da bi se dobio odgovor na pitanje „Koji način poučavanja je učinkovitiji?“. Ovo dovodi do zaključka da bi se trebao koristiti eksperiment.

3.2.3. POSTUPCI I INSTRUMENTI PRIKUPLJANJA PODATAKA

Pod istraživačkim se *postupcima* podrazumijevaju oblici i specifičnosti samog istraživanja, a pod istraživačkim *instrumentima* onaj alat koji se u tom radu upotrebljava [MUŽI1977]. Postoje instrumenti koji se primjenjuju u raznim postupcima.

Osnovna shema svakog eksperimenta sadrži utvrđivanje inicijalnog stanja, zatim određeno vrijeme unošenja (ili neunošenja) eksperimentalnih faktora, te utvrđivanje završnog stanja. Instrument mjerenja treba primijeniti u vezi s utvrđivanjem inicijalnog i završnog stanja. U završnom se ispitivanju primjenjuje instrument kojim se ispituje ista pojava, instrument koji je iste težine (imaju isti broj bodova), koji je sastavljen istovremeno kad i instrument kojim se utvrđivalo inicijalno stanje, no kod kojeg su zadaci različiti od onog prvog instrumenta. Upotrebljavaju se, dakle, dva instrumenta identična po velikom broju aspekata, pa im se rezultati mogu usporediti, ali su dovoljno različiti da ta sličnost ne može biti parazitarni faktor. Zbog te velike podudarnosti takvi se instrumenti nazivaju *paralelnim formama* istog instrumenta.

Transverzalna istraživanja obuhvaćaju pristup raznim aspektima neke pojave. Da bi se ta složenost mogla obuhvatiti, potrebno je više raznih instrumenata: bilo raznih instrumenata iste vrste, bilo raznih vrsta. Takav skup raznih instrumenata naziva se *baterijom instrumenata*.

Odabir postupaka i instrumenata ovisi o specifičnostima samog problema koji se proučava, o vremenskim i financijskim mogućnostima. U većini istraživanja primjenjuje se više instrumenata: paralelne forme instrumenata, odnosno baterije instrumenata.

3.2.3.1 KARAKTERISTIKE INSTRUMENATA

Instrumenti istraživanja trebaju prikupiti što točnije podatke i trebaju izvršiti funkciju mjerenja. Zato je potrebno da taj instrument ima neke neophodne karakteristike. Testovi se smatraju najrazrađenijim mjernim instrumentima i kod njih se postavlja pitanje svake metrijske karakteristike o kojima inače treba voditi računa pri odabiru instrumenta.

VALJANOST

Valjanost je karakteristika testa koja nam pokazuje da li i koliko test mjeri ono što s njim želimo mjeriti. Ne postoji "opća" valjanost testa, tj. valjanost se može odrediti samo u vezi konkretne namjene za koju se test koristi.

Kriteriji valjanosti:

1. Nastavni sadržaj

Kod instrumenata kojima se mjeri učinak procesa učenja i poučavanja, kriterij valjanosti je slaganje nastavnih ciljeva i nastavnog sadržaja sa sadržajem instrumenta. Ako se nekim instrumentom točno ispituje ono što bi učenik trebao naučiti, onda je taj instrument valjan. Uobičajen postupak određivanja valjanosti instrumenta je usporedna analiza sadržaja testa i nastavnog sadržaja od strane većeg broja stručnjaka (npr. anketiranje suradnika na izradi nastavnog sadržaja ili nastavnika koji su stručnjaci upravo za to

područje). Većinom testova se ne ispituje cijeli sadržaj nastave nego samo jedan dio koji na što bolji način treba predstavljati cjelinu. Da se ta reprezentativnost ostvari potrebno je sistematski odabrati iz svih područja nastavnog sadržaja elemente instrumenta (npr. zadatke u testu). Međutim, kriterij valjanosti ima nedostatak jer se ne može kvantitativno izraziti tj. ne postoji broj kojim bi se moglo izraziti razina slaganja sadržaja testa i nastavnog sadržaja. Tako nije moguće od više testova odabrati onaj koji se najviše slaže sa sadržajem, odnosno onaj kojeg bi trebalo primijeniti.

2. *Subjektivno ocjenjivanje ispitanika*

Instrumenti koji služe za određivanje učenikova napretka u procesu učenja i poučavanja vrše sličnu funkciju kao i uobičajeno ocjenjivanje učenika. Stoga se kao kriterij valjanosti može uzeti slaganje tih dviju veličina (između rezultata u testu i školske ocjene iz određenog gradiva se računa korelacija). Da bi se izbjeglo djelovanje subjektivnog faktora kod nastavnika, kao kriterij valjanosti testa može se uzeti prosječna ocjena koju istim ispitanicima daje veći broj nastavnika istog područnog znanja na osnovu usmenog ispitivanja.

3. *Slaganje s rezultatima koji su postignuti drugim instrumentima*

Kod mjerenja onih pojava kod kojih već postoji valjani mjerni instrument, slaganje novog instrumenta s tim valjanim predstavlja kriterij valjanosti. Stvaranje tog novog instrumenta je u ovoj situaciji opravdano samo ako se novim instrumentom postupak mjerenja pojednostavljuje, skraćuje, pojeftinjuje ili kada je potrebno izraditi paralelne forme istog instrumenta.

4. *Određivanje valjanosti pomoću faktorske analize*

Faktorska analiza je postupak kojim se istražuju veličine koje određuju rezultat u većem broju različitih operacija. Kad je za neki instrument ustanovljena njegova zasićenost pojedinim faktorima, tada se točno zna što se i u kojoj mjeri tim instrumentom ispituje i mjeri.

5. *Prognostička valjanost*

Ovaj kriterij dolazi u obzir kad se na osnovi rezultata u nekom instrumentu želi odrediti vjerojatnost uspjeha u nekom drugom području, ali ne i koliko taj instrument mjeri upravo ono što želimo mjeriti. Visinu valjanosti predstavlja korelacija između uspjeha u tom instrumentu i uspjeha u kasnijoj djelatnosti.

POUZDANOST

Ova karakteristika određuje možemo li se i kojoj mjeri osloniti na rezultat koji se dobije instrumentom, tj. da li instrument točno mjeri, bez obzira što se mjeri. Kriterij određivanja pouzdanosti utvrđuje se stupnjem slaganjem između dviju primjena istog instrumenta, a izražava se koeficijentom korelacije.

Načini određivanja pouzdanosti mjernog instrumenta:

1. *Koeficijent stabilnosti*

Ako se pretpostavlja da je instrument pouzdan, onda je s time povezana i pretpostavka da će dvije primjene istog instrumenta u mjerenju iste veličine dati i isti rezultat.

2. *Koeficijent homogenosti*

Jedan od načina je određivanje *koeficijenta homogenosti testa*. Potrebno je test podijeliti na dva jednaka dijela i zatim ispitati njihovo slaganje. Nakon primjene testa, zadaci se podijele, obično na parne i neparne zadatke. Izračunavanjem koeficijenta korelacije između parnih i neparnih zadataka dobije se pokazatelj pouzdanosti. Nedostatak ovog pristupa je što se polazi od pretpostavke da su parni i neparni zadaci međusobno slični, što ne mora biti točno, pa tako ne možemo biti dovoljno sigurni u dobiveni koeficijent pouzdanosti. Ako brzina rješavanja nema utjecaja na rezultate, tj. ako je vrijeme gotovo neograničeno i nema utjecaja na klasifikaciju rezultata, onda se može primjenjivati ovaj pristup, a u protivnom su potrebne dodatne analize koje se zasnivaju na tabeli analize zadataka ili na primjeni analize varijance.

Kako je svaki zadatak zapravo jedno mjerenje, dulji instrument će imati veću pouzdanost od kraćeg. Za procjenu koliko treba produljiti instrument (novim zadacima iste kvalitete kao što su postojeći) da bi se ostvarila tražena pouzdanost postoji i posebna formula:

$$r_n = \frac{n \cdot r_t}{1 + (n-1) \cdot r_t} \quad (3.45)$$

gdje je r_n = novi koeficijent pouzdanosti, r_t = trenutni koeficijent pouzdanosti, n = broj puta za koliko se instrument namjerava produljiti

$$n = \frac{\text{novi_broj_zadataka}}{\text{trenutni_broj_zadataka}} \quad (3.46)$$

Formula (3.37) naziva se Spearman-Brown-ova formula ili *formula predviđanja* (eng. prophecy formula) zbog toga što se s pomoću nje može procijeniti razlika koja će se pojaviti u koeficijentu pouzdanosti uslijed promjene duljine testa.

Ako se želi odrediti koliko zadataka je potrebno dodati da bi se dobila željena pouzdanost, onda se koristi sljedeća formula:

$$n' = n \cdot \frac{r' \cdot (1-r)}{r \cdot (1-r')} \quad (3.47)$$

gdje je n' = broj zadataka koji je potreban za željenu pouzdanost, n = trenutni broj zadataka, r' = pouzdanost koja se želi postići, r = trenutna pouzdanost

Specifičan slučaj primjene Spearman-Brown-ove formule je upravo pri određivanju pouzdanosti testa na temelju korelacije između dviju njegovih polovina, tj. koeficijenta homogenosti. Time što se izračunava korelacija polovine instrumenta, duljina instrumenta od koje se prilikom izračunavanja polazi se smanjila na polovicu. Želimo li znati pouzdanost za čitav instrument, primjenit će se Spearman-Brown-ova formula uz produljenje za dva puta ($n=2$), pa formula glasi:

$$r_{cijeli} = \frac{2 \cdot r_{pola}}{1 + r_{pola}} \quad (3.48)$$

gdje je r_{cijeli} = koeficijent pouzdanosti cijelog testa, r_{pola} = koeficijent pouzdanosti pola testa

3. Koeficijent ekvivalentnosti

Kod instrumenata kojima se mjeri i brzina rada ispitanika, za određivanje pouzdanosti je potrebno raspolagati paralelnim formama istog instrumenta. Tada se pouzdanost određuje na osnovi slaganja rezultata između tih formi primijenjenih kod istog ispitanika jedne iza druge.

4. Cronbach-ov α koeficijenta

Još jedan način određivanja pouzdanosti je pomoću *Cronbach-ovog α koeficijenta* (eng. Cronbach's Coefficient Alpha) [HEMP2005]. Ovaj način se često koristi za procjenu pouzdanosti jer se s njim 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. 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. Za razliku od prethodnog pristupa gdje se test dijeli na parne i neparne zadatke, Cronbach-ov koeficijent se računa kao aritmetička sredina svih mogućih podjela testa. Ovisi o broju zadataka, a moguće ga je izračunati pomoću statističkih računalnih programa.

Prema [HEMP2005] smatra se da je test pouzdan ako je $\alpha > 0.70$, ali preveliko približavanje 1 može značiti redundanciju tj. nepotrebno ponavljanje istih stvari kroz različite zadatke. Pri analizi rezultata obično se detaljnije promatra stupac tablice koji sadrži α koeficijente. To su koeficijenti koje bi test mogao imati ako se određeni zadatak izbriše.

Visina koeficijenta pouzdanosti

Kao najnižu vrijednost koeficijenta korelacije za svaki instrument Mužić navodi 0.80, a ponekad se traži i 0.85 ili čak 0.90, npr. ako se radi o matematičkim zadacima s računanjem gdje se očekuje veća pouzdanost.

Među faktorima koji djeluju na pouzdanost, važnu ulogu ima jednoznačnost pitanja kao i jednoznačnost uputa koje ispitivač daje ispitanicima. Dvoznačnost, kao i zadaci kod kojih postoji mogućnost da se točan odgovor slučajno pogodi, smanjuju pouzdanost.

OBJEKTIVNOST

Ovdje se radi o razini objektivnosti prilikom ocjenjivanja odgovora ispitanika. Ako zadatak ima samo jedan mogući točan odgovor (npr. zadaci s više ponuđenih odgovora od kojih se bira jedan) onda ocjenjivač ne može biti u nedoumici: odgovor je ili točan ili netočan. Problemi nastaju kod pitanja na koje nije moguće odgovoriti na jedinstven način, ispitanici odgovaraju "svojim

riječima", a ocjenjivač mora subjektivno ocijeniti je li odgovor točan. Događa se i da je pitanje postavljeno na nespretan način ili je previše složeno da bi se na njega mogao dati jednostavan odgovor.

Za određivanje objektivnosti je potrebno usporediti ocjene raznih ocjenjivača. Pri tome se očekuje stupanj korelacije veći od 0.97, osim ako su pitanja prilično složena pa se tada dozvoljavaju i niže vrijednosti, a potrebno je dati i preciznije upute za ocjenjivanje.

OSJETLJIVOST

Test je osjetljiv ako se s njim mogu razlikovati i male veličine onoga što se mjeri. Ako su svi zadaci laki, onda se ne mogu međusobno razlikovati oni najbolji, a ako su svi teški onda postoji mogućnost da slabiji učenici ne riješe gotovo ništa, što naravno ne znači da ništa ne znaju. Pri tome ono što istraživaču izgleda lako (ili teško) ne mora biti lako (ili teško) za učenika. Dakle, subjektivna procjena istraživača nije dovoljna. Potrebna je empirijska verifikacija koja se ostvaruje na temelju njegove probne upotrebe, analize uspjeha u pojedinim zadacima te izmjena koje iz te analize proistječu.

Sama tehnika izrade analize zadataka u testu je jednostavna. Izrađuje se tablica čiji redovi predstavljaju oznake pojedinih ispitanika, a stupci služe za unošenje uspjeha tih ispitanika u pojedinim zadacima. U posljednjem retku nalaze se za svaki pojedini zadatak, ukupan broj ispitanika koji su ga riješili iz čega se računa postotak tih ispitanika. Taj podatak ukazuje na relativnu težinu tog zadatka. Osjetljivost testa garantira manji broj lakih i teških zadataka, a najveći srednje težine. U posljednjem stupcu nalaze se ukupni rezultati koje su postigli pojedini ispitanici. Ti podaci služe za određivanje diskriminativne vrijednosti pojedinih zadataka. Označavanjem najboljih i najslabijih ispitanika u čitavom testu može se ustanoviti slaganje uspjeha u pojedinom zadatku s uspjehom u čitavom testu, tj. njegova diskriminativna vrijednost.

Test se smatra osjetljivijim što je duži. Test koji se sastoji od samo deset zadataka ne može razlikovati više od 11 kategorija ispitanika (jer se rezultati kreću od 0-10), bez obzira koliko ima ispitanika (može ih biti i nekoliko tisuća). Naravno, u ovom slučaju se zadaci označavaju s točan/netočan. U slučaju kada se zadaci ocjenjuju s više bodova, onda se mjere i manje razlike. Međutim, osjetljivost nema smisla ako su pouzdanost i objektivnost testa niske. Konačan kriterij osjetljivosti je distribucija rezultata. Ako je mnogo istih rezultata, onda test ne može biti dovoljno osjetljiv.

DISKRIMINATIVNA VRIJEDNOST ZADATAKA U INSTRUMENTU

Kod instrumenata čiji se rezultat iskazuje zbrojem uspjeha ispitanika u zadacima, pojedini zadaci u instrumentu su više valjani što se više slaže uspjeh učenika na tom zadatku s uspjehom na čitavom instrumentu. Prema tome, takvi zadaci (ako su valjani) 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 instrumentu biti visok, a diskriminativna vrijednost zadatka visoka i pozitivna. Može se dogoditi i da neki zadatak učenici sa slabijim uspjehom rješavaju, a uspješniji učenici ne. U tom slučaju je diskriminativna vrijednost zadatka negativna, a takav zadatak ne bi smio biti dio instrumenta. Kako se tu radi o slaganju uspjeha u čitavom instrumentu sa uspjehom u jednom njegovom

elementu, visina diskriminativne vrijednosti zadataka određuje se izračunavanjem korelacije i to pomoću biserijskog koeficijenta korelacije.

Potrebno je rangirati ispitanike prema uspjehu ostvarenom na testu, te se izdvoji 27% najboljih studenata i 27% najslabijih, a srednji ne ulaze u razmatranje. Broj 27 je odabran zato što je dokazano da je pri takvoj razdiobi koeficijent biserijske korelacije najosjetljiviji. Pri tom bi veličina uzorka trebala biti veća od 370 ispitanika, kako bi u obje grupe bilo bar po 100 ispitanika. Inače koeficijent korelacije treba izračunati na cijelom uzorku. Zatim se za svaki zadatak odredi udio uspješnih ispitanika u grupi najboljih te zatim u grupi najslabijih. Nakon toga se, umjesto računanja korelacije, dobiveni udjeli potraže u tablici u Prilogu E visina korelacije koja predstavlja pokazatelj diskriminativne vrijednosti tog zadatka. U zaglavlju tablice u Prilogu E se pronade postotak koji je najbliži postotku točnih rješenja u danom zadatku kod grupe najboljih ispitanika. Zatim se to isto nađe u prvom stupcu za grupu najslabijih ispitanika. Polje u kojem se nađeni red i stupac sijeku daje približnu vrijednost biserijskog koeficijenta korelacije.

BAŽDARENOST

Svako mjerenje zahtijeva neku mjernu jedinicu kao i skalu na kojoj se rasprostiru. Pri tom je poželjno da razmak između tih mjernih jedinica bude uvijek isti. Kod pedagoškog istraživanja mjerna ljestvica se dobiva na osnovu rezultata ustanovljenih kod samih objekata mjerenja.

Tehnike baždarenja:

1. Baždarenje na bazi norme dobi

Kod ovog baždarenja stavlja se u odnos postignuti nivo znanja s prosjekom uspjeha svih ispitanika određene kronološke dobi pa se dobiva „obrazovna dob“. Ova je tehnika jedna od najstarijih i njena rasprostranjenost nije velika.

2. Baždarenje na temelju postotka ispitanika koji postižu određeni rezultat, tj. na temelju decila i centila

Promotrimo primjer: učenik je u prvom testu postigao rezultat koji otprilike odgovara srednjem rezultatu svih učenika u uzorku (tj. od njega 50 posto ispitanika postiže bolji, a 50 posto slabiji rezultat). U drugom testu je postigao rezultat koji odgovara onom na trećoj četvrtini svih rezultata, počevši od najslabijih prema najboljima (tj. od njega 25 posto ispitanika postiže bolji, a 75 posto slabiji rezultat). Kao što je kod prvog testa određen rezultat koji odgovara onom na 50 posto, te kod drugog testa na 75 posto, tako se određuju rezultati i za ostale postotke. Ako se pritom uzimaju grupe od po deset postotaka, onda se upotrebljavaju *decili*. To su točke u distribuciji koje je dijele na deset dijelova s jednakim brojem jedinica. Prema tome, ako se neki učenik nalazi između šestog i sedmog decila, to znači da se nalazi u grupi rezultata koje ne postiže 60 posto slabijih ispitanika no koji ni ne ulaze u rezultate 30 posto najboljih ispitanika. Ako se ide za finijom razdiobom, pa se uzme u obzir svaki postotak, radi se o *centilima*. To su točke u distribuciji koje je dijele na sto dijelova s jednakim brojem jedinica. Ovom se tehnikom mogu uspoređivati rezultati istog učenika u različitim instrumentima i razni učenici u istim instrumentima.

Bitni nedostaci ove tehnike proistječu iz činjenice da ekstremni decili i centili obuhvaćaju mnogo veći broj različitih rezultata nego oni iz sredine. Zbog toga su područja pojedinih

decila i centila vrlo uska, tako da se individualne razlike u sredini skale prikazuju kao pretjerano velike. Zato slika koju daje ta tehnika nije realna.

3. *Baždarenje na bazi standardiziranog odstupanja*

Osnova ove tehnike je u određivanju mjere za koliko je koji rezultat bolji ili slabiji od aritmetičke sredine rezultata koji se postigao na reprezentativnom uzorku, s time da se kao jedinica pri određivanju tog odstupanja između rezultata i prosjeka uzima standardna devijacija (str.478), tj. mjera raspršenja. Temeljni pojam ove tehnike je tzv. z-vrijednost, koja predstavlja razliku između rezultata koji je postignut mjernim instrumentom i prosjeka uzorka, s time da se ta razlika izražava u jedinicama standardne devijacije. Ovom tehnikom su prevladani osnovni nedostaci tehnike baždarenja na osnovi decila i centila.

4. *Baždarenje na bazi transformacije na temelju normalne distribucije*

Ova tehnika predstavlja usavršavanje prethodne tehnike. Ona je najsloženija, ali je teorijski najopravdanija kod onih pojava kod kojih veličine u uzorku nemaju normalnu distribuciju, ali postoji opravdana osnova za pretpostavku da se pojava ipak normalno distribuira. Pri tom se jedinice skale određuju na temelju zamišljene krivulje normale s aritmetičkom sredinom i standardnom devijacijom koju smo dobili na uzorku, a iz tablice u Prilogu B površina pod normalnom krivuljom uzima se postotak rezultata koji odgovara pojedinoj jedinici skale.

PRAKTIČNOST I EKONOMIČNOST

Pod ovim se karakteristikama podrazumijevaju neka svojstva instrumenta mjerenja koja se ne odnose na samu funkciju mjerenja: jednostavnost primjene, kontrole rezultata i korištenja dobivenih rezultata, te ušteda vremena i financija pri ovim operacijama. Dakle, prednost ima instrument čija primjena ne traje dulje od 45 minuta, čije je ispravljanje kratak i automatiziran posao, kod kojeg se odgovori pišu na posebne listove (da bi se isti primjerci testova mogli ponovno upotrijebiti), itd. Ipak, ne bi se smjelo, ni zbog kakve uštede, dozvoliti nedovoljnu dosljednost instrumenta.

3.2.3.2 VRSTE POSTUPAKA I INSTRUMENATA

RAD NA DOKUMENTACIJI – EVIDENCIJSKI LIST

Pri istraživanju koje polazi od dokumenata s podacima o pedagoškim pojavama koji već postoje, istraživač treba prvo upoznati aspekte istraživačke primjene. Treba povesti računa o tome da kod nekih dokumenata može postojati sumnja o njihovoj originalnosti ili istinitosti. Nakon što su dokumenti pronađeni i nakon što se, po potrebi, izvršila kontrola vjerodostojnosti i istinitosti, iz njih istraživač uzima podatke koji su mu potrebni. Tada istraživač pristupa sređivanju podataka, njihovoj obradi i analizi. Ovaj postupak se najviše koristi kod komparativnih istraživanja, te kao pomoćni postupak kod istraživanja koja su u osnovi orijentirana na sakupljanje i proučavanje podataka.

SISTEMATSKO PROMATRANJE–APARAT I PROTOKOLI ZA SNIMANJE

Sistematsko promatranje predstavlja direktan put upoznavanja pedagoške stvarnosti. Njegova svrha je što točnije evidentirati pojava koja se promatra. Sistematsko promatranje posjeduje sljedeće karakteristike:

1. služi usvajanju novih znanstvenih spoznaja
2. zauzima određeno mjesto u projektu istraživanja
3. vrši se prema unaprijed utvrđenom planu
4. usmjereno je prema određenom objektu promatranja
5. snimanje (prikupljanje podataka) se vrši simultano
6. snimak (rezultat snimanja) treba biti pristupačan kvantitativnoj obradi
7. ispitivač je stručno osposobljen za taj posao
8. pouzdanost, objektivnost i valjanost snimke mogu se kontrolirati

Prema načinu promatranja sredstva promatranja dijele se na:

1. promatranje pomoću tehničkih pomagala
2. promatranje od strane čovjeka koji izrađuje protokol promatranja (zapisnik o snimanju)

RAZGOVOR – PROTOKOLI RAZGOVORA

Razgovor služi planskom izazivanju pokazatelja osobnosti osobe s kojom razgovaramo s ciljem otkrivanja novih spoznaja, te uzročno-posljedičnih veza među pedagoškim pojavama, odnosno postoji težnja za znanstvenom generalizacijom.

Istraživački razgovor se prema svom sadržaju i toku dijeli na:

1. *vezani* - Pitanja unaprijed određena. Ovaj razgovor je obično analitički, tj. njime raščlanjujemo pojavu koju želimo proučiti i zato se treba pridržavati predviđenog toka razgovora.
2. *slobodni* - To je više diskusija u kojoj spontano ispitanik iznosi svoje poglede.

Drugi kriterij diobe istraživačkog razgovora je vezan za osobe s kojima se razgovara:

1. *direktni* – ispitanik je osoba o kojoj se traže podaci
2. *indirektni* – ispitanici daju podatke o osobi koja je predmet našeg interesa

Prema broju ispitanika razlikuju se:

1. *individualni* – razgovor s jednom osobom
2. *grupni* – simultani razgovor s većim brojem osoba

Kod svakog razgovora ispitaniku treba protumačiti važnost upravo njegovih odgovora. Ako je riječ o osjetljivim temama, korisno je obećati anonimnost koje se treba onda i pridržavati. Korisno je uvjeriti ispitanika da se ne radi o ispitu i da nema dobrih i loših odgovora. Isto tako, ispitanik treba steći dojam da ispitivaču razgovor ne predstavlja teškoću ili napor, te da vlada sadržajem razgovora. Svojim nastupom ispitivač treba stvoriti ugodnu atmosferu, biti ugodan, pristojan, pristupačan i taktičan. Ispitivač ne iznosi svoje mišljenje jer bi time mogao djelovati na usmjerenost ispitanikovih izjava.

Protokol (zapisnik) je rezultat razgovora, tj. dokument čiji je sadržaj izvor podataka koji su dobiveni razgovorom i koji postaju predmet daljnje obrade (ekvivalent protokolu snimanja kod sistematskog promatranja).

ANKETIRANJE – ANKETNI UPITNICI

Anketa je postupak kojim se ispitanicima postavljaju pitanja u vezi s činjenicama od znanstvenog interesa, a koje su poznate ispitanicima, ili pitanja u vezi s mišljenjima ispitanika. Ispitanici na njih odgovaraju pismeno. Anketa je vremenski mnogo ekonomičnija od razgovora jer je moguće istovremeno anketirati više osoba.

Instrument ankete u kojoj se pitanja postavljaju i na njih se odgovara pismenim putem naziva se anketni upitnik. U anketnom upitniku nema procjenjivanja pomoću skale sudova i ne ocjenjuje se uspjeh. U potpuni tok anketiranja ulazi i određivanje uzorka, obrada podataka i njihova interpretacija.

Prema sadržaju anketnog upitnika razlikuje se traženje činjeničnih podataka od traženja podataka o stavovima, mišljenjima i sl. Prema tipu anketnog pitanja može se upotrijebiti otvoreni (ispitanik ispisuje odgovor) i zatvoreni tip (ispitanik odabire odgovore). U svezi diskrecije o osobnosti ispitanika razlikuju se anonimni anketni upitnici o onih u kojima se pojavljuje ime ispitanika.

Predug anketni upitnik vrlo demotivira. Stoga je njegova kratkoća bitan preduvjet uspjeha ankete. Kratkoća ne ovisi samo o broju pitanja, već i o složenosti pitanja i tehnici odgovaranja. Općenito bi trebalo nastojati da anketa ne zahtjeva više od desetak minuta ispitanikova vremena.

TESTIRANJE – TESTOVI

Pri nastojanju da se odredi u kojoj su mjeri učenici usvojili neko gradivo, polazi se od direktnog ispitivanja tog gradiva kod samih učenika. Direktno ispitivanje se vrši najčešće testiranjem. Test je standardizirani postupak kojim se izaziva određena aktivnost čiji se učinak mjeri i vrednuje međusobnim uspoređivanjem individualnih rezultata.

Vrste testova prema osobinama ispitanika koje se njima ispituju:

1. *testovi znanja* - Ovim se testovima nastoji utvrditi da li je i u kojoj mjeri ispitanik usvojio određena znanja, vještine i navike. U eksperimentalnim istraživanjima oni su jedan od instrumenata određivanja inicijalnog stanja te osnovni instrument pri određivanju završnog stanja.
2. *testovi sposobnosti* - Ovim se testovima nastoje utvrditi osobine koje su preduvjet za uspjeh na određenom području aktivnosti. Oni se upotrebljavaju kao pokazatelji inicijalnog stanja u eksperimentima.
3. *testovi osobnosti u užem smislu* - Ovim se testovima mjere složenije osobine: interesi, stavovi, karakteristike temperamenta, karaktera i sl.

Vrste testova prema načinu rješavanja zadataka:

1. *pismeni testovi* - Omogućavaju kolektivnu primjenu, a samim time i veliku uštedu u vremenu.
2. *usmeni testovi* - Poprimaju značajke ekstremno strogo vezanog razgovora.
3. *testovi čina* - Omogućavaju učeniku da pokaže da nešto zna napraviti, a ne da zna samo kazati kako bi to trebalo napraviti.

Shema toka konstrukcije testa:

1. Treba sastaviti puno više zadataka nego što će ih ući u test, kao bi se omogućio izbor. Treba sastaviti zadatke koji će obuhvatiti čitavo gradivo (odnosno reprezentativan uzorak tog gradiva), što je preduvjet za valjanost testa. Svaki zadatak treba biti onog tipa koji će

onemogućiti djelovanje drugih faktora na ispitanikovo rješavanje zadatka. Tipovi zadataka pismenog testa: dosjećanja, nadopunjavanja, dvočlanog izbora, višestrukog izbora, uspoređivanja, sređivanja.

2. Prva eliminacija zadataka za koje se već na osnovu subjektivnog dojma može uočiti da ne odgovaraju postavljenim zahtjevima.
3. Od zadataka koji su ostali izrađuje se test s još uvijek većim brojem zadataka od predviđenog konačnog broja i empirijskim se putem ustanovljava koje još zadatke treba eliminirati (provodi se test na ograničenom broju ispitanika).
4. Dobiveni se rezultati analiziraju. Pri tom se vrši analiza svakog zadatka prema relativnoj težini i prema diskriminativnoj vrijednosti. Vršer se pokusna izračunavanja ostalih metrijskih karakteristika. Na taj se način dolazi do posljednjeg izbora zadataka koji će sačinjavati konačni test.
5. Konačni test se primjenjuje na reprezentativnom uzorku ispitanika. Na temelju dobivenih podataka vrši se baždarenje testa i određuju se konačne metrijske karakteristike.

POSTUPCI PROCJENJIVANJA I PROSUDIVANJA – SKALE SUDOVA

Ovi se postupci koriste tamo gdje podatke treba prikupiti na temelju procjene nečega. Pri tom se putem raznih varijanti skala određuje postojanje ili nepostojanje, odnosno stupanj postojanja pojave koju se procjenjuje. Instrument koji se pri tom koristi naziva se *skalom sudova*. Skala sudova je, dakle, instrument kojim se prikupljaju sudovi (mišljenja, procjene) ispitanika o osobinama ili postupcima određenih osoba, stvari ili pojava. Svojim vanjskim oblikom skale sudova su slične anketnim upitnicima.

Po vanjskom obliku skale sudova dijelimo na:

1. *deskriptivne* – sastoji se od niza predloženih tvrdnji
2. *grafičke* – sastoji se od vodoravne crte ispod koje su upisane samo ekstremne osobine, a ponekad i srednja

Ponekad se kod skala sudova unose i brojčane vrijednosti. Pri statističkoj obradi svaki sud dobiva svoju brojčanu vrijednost, bez obzira na činjenicu da ona najčešće nije upisana u samu skalu.

Varijante skale sudova:

1. *Alternativna skala* - predstavlja varijantu deskriptivne skale u kojoj umjesto većeg broja predloženih tvrdnji nalazimo samo dvije: afirmativnu i negativnu.
2. „*Odredi tko*“ *skala* (eng. „*guess who*“ *technique*) – ispitanik treba navesti ime jedne ili više osoba na koje bi se predložena karakteristika mogla odnositi.
3. *Skala rangova* (eng. *rank-order scaling*) – ispitanik rangira, tj. određuje redosljed od maksimalnog do minimalnog inteziteta pojave.
4. *Postupak pomoću uspoređivanja parova* (eng. *paired comparisons*) – ispitanik izvršava samo elementarni postupak od kojeg se sastoji svako rangiranje, tj. uspoređuje dvije po dvije osobe, stvari ili pojave.
5. *Postupak uspoređivanja s prethodno određenim rangom* (eng. *man-to-man technique*) – od ispitanika se najprije traži da „odredi tko“ postupkom imenuje po jednog tipičnog

predstavnik za svaku od predloženih karakteristika. Nakon toga treba za svaku od ostalih osoba za koje daje svoj sud odrediti između koja dva tipična predstavnika dolazi.

6. *Skalogram* - ova skala je rezultat nastojanja stvaranja jednodimenzionalnih skala, tj da svaka viša kategorija obuhvaća sve kategorije koje su od najniže, tj. da posjeduje kumulativno svojstvo. Naziva se i Guttmanova skala.
7. *Q-postupak* (eng. *q-sort technique*) – ispitaniku se predloži niz tvrdnji koje se odnose na razne karakteristike osobe, s tim da ih redom posloži na razne načine. Najprije da odabere one karakteristike za koje smatra da se odnose upravo na njega, zatim da odabere one karakteristike koje bi želio imati, nakon toga da odabere karakteristike koje odgovaraju onima kakvim ga smatra okolina.
8. *Skala proizvoda* – služi za procjenjivanje određenih proizvoda pomoću njihove usporedbe s tipičnim predstavnicima pojedinih kvalitetnih kategorija.
9. *Sociogram* – predstavlja grafički oblik iznošenja rezultata „odredi tko“ skale u vezi sa socijalnom prihvatljivošću članova određene grupe.

3.2.4. EKSPERIMENT KAO METODA VREDNOVANJA

Eksperiment predstavlja metodološki način pristupa istraživanja pedagoške pojave kojim se nastoje uočiti uzročno-posljedične veze među pojavama, služeći se pri tom raznim postupcima prikupljanja podataka, kao što je promatranje, testiranje, razgovor i sl. [MUŽI1977]. Osnovna je karakteristika pedagoškog eksperimenta plansko ispitivanje učinka procesa učenja i poučavanja.

Cook i Campbell (prema [GARS2006]) dijele eksperiment kao metodu vrednovanja na:

1. Pravi eksperiment
Karakterizira ga postojanje kontrolne i eksperimentalne grupe koje su ekvivalentne i dobivene slučajnim odabirom ispitanika. Na kontrolnu grupu ili ne djeluje eksperimentalni faktor ili na nju djeluje neki drugi faktor provjerenog učinka. Tu spadaju osnovni i složeni modeli eksperimentalnih postupaka o kojima će više riječi biti kasnije.
2. Kvazi-eksperiment
Karakterizira ga nemogućnost raspodjele ispitanika u grupe slučajnim odabirom i grupe ne moraju biti ekvivalentne. Također se ne mora koristiti inicijalni test ili se može koristiti više inicijalnih testova (eksperimenti s isprekidanim vremenskim slijedom, eng. *interrupted time series*).

3.2.4.1 ZAVISNA I NEZAVISNA VARIJABLA

U eksperimentu se uvijek susrećemo s *postupkom* i *učinkom tog postupka*, tj. nekom pojavom (znanjem, vještinom, stavovima učenika) koja je u većoj ili manjoj mjeri zavisna od postupka koji je trebao izazvati tu pojavu. Pojava koja predstavlja postupak, kao i pojava koja predstavlja njen učinak može se mijenjati, te stoga ih nazivamo *varijablama*. Varijabla je, dakle, svaka

karakteristika bilo koje pedagoške pojave kod koje promatramo kvantitativne ili kvalitativne razlike koje se u njoj javljaju (npr. razni oblici procesa učenja i poučavanja, nivoi znanja učenika, dob ili spol učenika, itd.).

Eksperimentom zapravo utvrđujemo u kojoj mjeri zavisi druga varijabla (posljedica) od prve (utjecaj). Stoga se u eksperimentu razlikuju *nezavisne* varijable (one koje djeluju, koje predstavljaju uzroke) i *zavisne* (one na koje se djeluje, koje predstavljaju posljedice). S obzirom na odnos nezavisnih i zavisnih varijabli u eksperimentu, najlakše bi bilo doći do zaključka o djelovanju određene nezavisne varijable na određenu zavisnu varijablu kad bi se na neki način eliminiralo ili neutraliziralo djelovanje ostalih uzroka, tj. nezavisnih varijabli.

Time se zapravo teži k ostvarenju drugog Mill-ovog pravila ostatka. Drugo pravilo se može ilustrirati na sljedeći način: u dvije jednake grupe učenika (po „kvaliteti“ učenika i učitelja, uvjetima rada itd.) primjenjuju se različiti nastavni postupci; stoga se razlika u uspjehu između tih dviju grupa može pripisati razlikama u postupcima. Ova ilustracija predstavlja pojednostavljenu sliku pedagoškog eksperimenta kod kojeg se eliminacijom ili neutralizacijom ostalih nezavisnih varijabli nastoji stvoriti situacija u kojoj se u dva slučaja (npr. dva različita nivoa znanja) razlikuje samo jedna nezavisna varijabla (npr. procesi učenja i poučavanja). Kako se tu nastoji pratiti djelovanje samo jedne varijable, ovo pravilo se naziva i *zakon jedne varijable*. Ta se jedna varijabla (npr. proces učenja i poučavanja) naziva *eksperimentalnom varijablom*, a svaki se od oblika u kojima se ta varijabla pojavljuje (npr. različiti procesi učenja i poučavanja) naziva *eksperimentalnim faktorom*. One varijable koje se nastoje eliminirati ili neutralizirati (npr. kvaliteta učenika, učitelja, uvjeti rada, itd.), nazivaju se *parazitarnim varijablama*, a oblici u kojima se oni pojavljuju *parazitarnim faktorima*. Potpuno, idealno ostvarenje zakona jedne varijable nije moguće zbog složenosti i međusobne povezanosti u prirodi i društvu.

Kontrolu u eksperimentu čine postupci kojima se detektira, pa zatim, eliminira ili/i stabilizira utjecaj parazitarnih faktora u eksperimentu. Osnovni cilj eksperimenta jest utvrditi da li postoji, koliki je i kakav je utjecaj određene nezavisne varijable na određenu zavisnu varijablu, pa je zbog toga kontrolna parazitarnih faktora jedan od najvažnijih postupaka u planiranju i provođenju eksperimenta.

Nezavisna varijabla	Zavisna varijabla
uzrok je zavisne varijable	posljedica je nezavisne varijable
U STVARNOJ SITUACIJI	
Više nezavisnih varijabli A, B, C, D, E, itd.	djeluje na zavisnu varijablu Z
PRI OSTVARENJU ZAKONA JEDNE VARIJABLE	
A - djeluje kao eksperimentalna varijabla B, C, D, E, itd. - predstavljaju parazitarnu varijable, pa im se djelovanje nastoji eliminirati ili neutralizirati	Z

Slika 3. 7. Shema odnosa zavisne i nezavisne varijable

3.2.4.2 VRSTE POGREŠAKA

Djelovanje varijabli koje su na Slika 3. 7. označene s B, C, D, itd. treba eliminirati ili neutralizirati ili kontrolirati njihovo djelovanje pomoću složenijih statističkih postupaka. Ako se to ne postigne, djelovanje tih parazitarnih varijabli dovest će do pogrešnog rezultata eksperimenta, rezultat će biti pristran.

Sistematizacija pogrešaka koje nastaju djelovanjem parazitarnih varijabli prema njihovim izvorima služi da bi se dobio pregled odakle one potječu i u kojim se sve oblicima mogu pojavljivati, što omogućava bolju kontrolu nad njima.

Izvori pogrešaka koje utječu na rezultat eksperimenta:

Tip pogreške „S“ - Izvor pogrešaka leži u subjektima s kojima se vrši eksperiment (npr. učenici, učitelji, i sl.) njihovim izborom u grupe u koje će se uvoditi eksperimentalni faktori može doći do toga da se te grupe razlikuju prema:

- predznanju učenika (eventualne se razlike mogu ustanoviti prethodnim ispitivanjem njihovog nivoa obrazovanja testom)
- općoj mentalnoj sposobnosti učenika (može se ustanoviti testovima opće mentalne sposobnosti)
- introvertiranosti odnosno ekstrovertiranosti učenika (može se ustanoviti adekvatnim testom osobnosti)
- spolu učenika
- dobi učenika
- socijalnom porijeklu učenika

Tip pogreške „G“ - Izvor pogrešaka leži u razlikama do kojih dolazi između grupa u koje se uvode eksperimentalni faktori, a koje se shvaćaju kao jedinstvene cjeline. Ove pogreške djeluju u čitavoj grupi bez obzira na razlike između ispitanicima.

U vezi s učiteljem:

- stručnost
- metoda priprema
- dob
- spol
- način pristupa učenicima
- mišljenje učitelja o vrijednosti eksperimentalnog faktora koji on uvodi

U vezi s prostorijama:

- zračnost
- osvjetljenje
- broj učenika u prostoriji
- veličina

U vezi s vremenom unošenja faktora:

- doba dana kad se nastava vrši
- redni sat
- dan u tjednu

U vezi s nekim drugim okolnostima:

- specifične karakteristike grupa

- nehotične pogreške u organizaciji samog eksperimenta
- razlike u materijalu koji služi u eksperimentu

Tip pogreške „R“ - Izvor pogrešaka leži u razlikama do kojih dolazi prilikom ponavljanja eksperimenta.

Mužić u [MUŽI1977] naglašava da se ove pogreške nužno uvijek javljaju u svakom eksperimentu. Težiti se može samo za njihovim smanjivanjem, a za one koje se ne mogu eliminirati, a pojavljuju se slučajno, treba ostvariti takva situacija da se prilikom generalizacije može statističkim putem voditi računa o vjerojatnosti njihove pojave.

Načini smanjivanja pogrešaka:

„S“ – ako u eksperimentu radimo, na primjer, s dvije grupe od pedeset slučajno odabranih učenika, mnogo je veća vjerojatnost da će se eventualne razlike između ispitanika tih dviju grupa međusobno neutralizirati, nego ako se eksperiment vrši s dvije grupe od deset učenika.

„G“ – koliko god povećamo broj učenika unutar pojedine grupe, djelovanje G pogreške neće se smanjiti. Na primjer, razlike u djelovanju osobnosti učitelja neće se smanjiti ako u grupi umjesto s deset radimo s pedeset učenika. Da nju eliminiramo, korisno je povećati broj grupa. Ako se eksperiment vrši s dvadesetak grupa, razlike u osobnostima učitelja vjerojatno će se u velikoj mjeri neutralizirati.

„R“ – povećanjem broja grupa ne smanjuje se mogućnost R pogreške, već treba eksperiment primijeniti i na grupe van granica nekog specifičnog sociološkog okruženja.

Pri izvedbi eksperimenta treba pridati najveću pažnju što potpunijoj eliminaciji pogrešaka koje nastaju djelovanjem parazitarnih faktora.

3.2.4.3 OSNOVNI MODELI EKSPERIMENTALNIH POSTUPAKA

Osnovnim se modelima prikazuju osnovne mogućnosti eksperimentiranja s obzirom na broj i raspored grupa s kojima se vrši eksperiment. Kriterij na osnovu kojeg se dijele osnovni modeli eksperimentalnih postupaka je broj i raspored grupa u kojima djeluju (odnosno ne djeluju) pojedini eksperimentalni faktori. Razlikujemo slijedeće osnovne modele [MUŽI1977]:

1. *Eksperiment s jednom grupom* – jedan ili više faktora se uvode uvijek u istu grupu ispitanika
2. *Eksperiment s paralelnim grupama* – postoje dvije ili više grupa, od kojih je svaka nosilac svog eksperimentalnog faktora (odnosno nije nosilac nijednog faktora)
3. *Eksperiment s rotacijom faktora* – faktori se rotiraju među grupama, tj. ista grupa redom postaje nosilac raznih faktora, odnosno, svaki postupak kojeg ispituje ulazi kao eksperimentalni faktor sad u jednu, sad u drugu grupu.

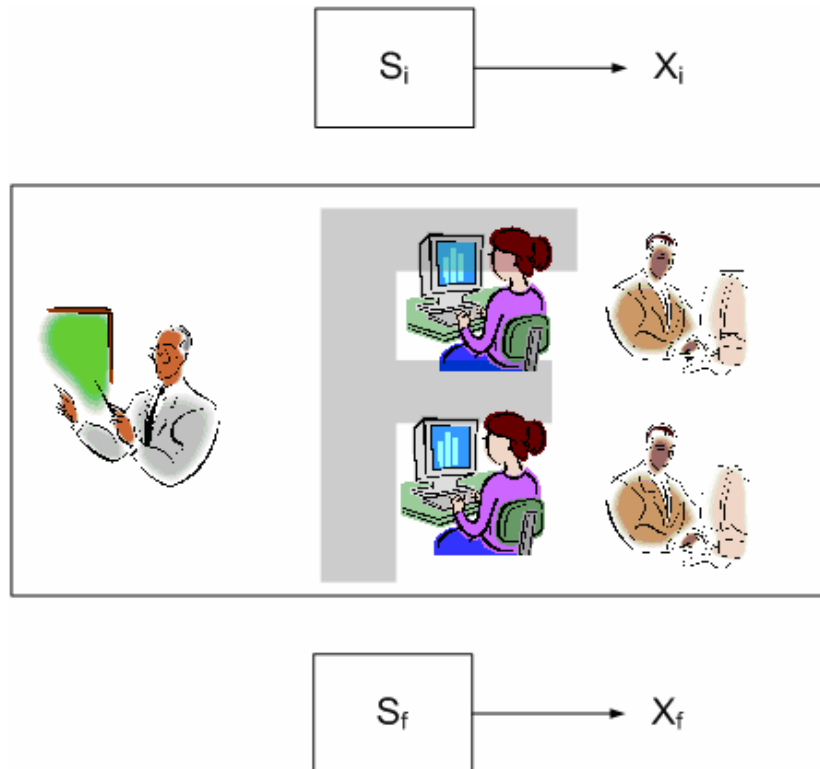
EKSPERIMENT S JEDNOM GRUPOM

Bitna karakteristika pedagoškog eksperimenta je, kao što je to već prije utvrđeno, utvrđivanje učinka određenog procesa učenja i poučavanja. Najprije treba upoznati početno stanje neke pojave, zatim se ostvaruje proces učenja i poučavanja, a na kraju se utvrđuje njegov učinak.

INICIJALNO STANJE – DJELOVANJE FAKTORA – ZAVRŠNO STANJE

Primjer: želimo ustanoviti koliki je napredak učenika prilikom usvajanja nekog područnog znanja pomoću sustava e-učenja.

Najprije se izvrši ispitivanje učenika u vezi s eventualnim predznanjem iz područnog znanja koje će se pomoću sustava e-učenja prenijeti na učenike. To treba unaprijed znati da bi se mogao dobiti „čisti“ učinak procesa učenja i poučavanja. Ovim se ispitivanjem doznaje inicijalno stanje (oznaka S_i). To inicijalno stanje varira od učenika do učenika, pa je potrebno odrediti aritmetičku sredinu inicijalnog stanja svih učenika (oznaka \bar{X}_i). Zatim se ostvaruje usvajanje novog područnog znanja pomoću sustava e-učenja. Sustav e-učenja je eksperimentalni faktor čiji učinak želimo ispitati. (oznaka F). Na kraju se vrši ispitivanje učenika da se ustanovi koliko su naučili prilikom korištenja sustava e-učenja. Na taj se način dobiva završno stanje (oznaka S_f). Ovdje također treba odrediti aritmetičku sredinu završnih stanja svih učenika koji su obuhvaćeni eksperimentom (oznaka \bar{X}_f) (Slika 3. 8.).



Slika 3. 8. Jednostavan model eksperimenta

Da se ustanovi čisti učinak djelovanja eksperimentalnog faktora, potrebno je od završnog stanja usvojenog područnog znanja oduzeti ono što su učenici od toga znali prije eksperimenta. Dakle, prosječna efikasnost eksperimentalnog faktora je:

$$\bar{X}_F = \bar{X}_f - \bar{X}_i \quad (3.49)$$

Izraz \bar{X}_F predstavlja veličinu napretka koji u prosjeku otpada na jednog učenika uslijed djelovanja eksperimentalnog faktora F.

Ovim se eksperimentom provjerava učinak nekog postupka samog po sebi, tj. bez ikakve usporedbe s kojim drugim postupkom. Prema tome takav jednostavan oblik eksperimenta s jednom grupom odgovara na pitanje: „Koliki je napredak učenika, primijenimo li neki određeni postupak?“, no ne daje odgovore na pitanje: „Da li je ovaj ili onaj postupak učinkovitiji?“

Osnovno je područje primjene takvog eksperimenta u nastojanju učitelja da ustanovi učinak svog rada. Ovakav individualni eksperiment ne daje rezultate na osnovu kojih se mogu vršiti generalizacije. On služi za verifikaciju da li neki postupak, čiji je učinak općenito već dokazana, odgovara i individualnim osobinama određenog učitelja.

Želimo li eksperimentom s jednom grupom usporediti učinke više postupaka, treba ih uvesti kao eksperimentalne faktore, tj. trebat će s istom grupom primijeniti razne postupke. Biti će onoliko ciklusa eksperimenta, koliko ima faktora.

Promotrimo to na prethodnom primjeru: usporedimo učinak rada sa sustavom e-učenja s učinkom tradicionalne nastave. Ista grupa ne može simultano raditi sa sustavom e-učenja i imati tradicionalnu nastavu. Za usporedbu su stoga potrebne dvije nastavne jedinice i dva ciklusa. U prvom ciklusu bi u grupi učenika djelovao prvi faktor (npr. sustav e-učenja), a u drugom bi u istoj grupi djelovao drugi faktor (npr. tradicionalna nastava). Unutar svakog od tih ciklusa odvija se ranije prikazani tok:

Prvi ciklus: INICIJALNO STANJE – DJELOVANJE PRVOG FAKTORA – ZAVRŠNO STANJE nakon toga
Drugi ciklus: INICIJALNO STANJE – DJELOVANJE DRUGOG FAKTORA – ZAVRŠNO STANJE

Da bi i put ovakvog eksperimenta označili shematski (Slika 3. 9.), potrebno je uvesti sljedeće oznake:

S_{i_1} - inicijalno stanje prije uvođenja prvog faktora

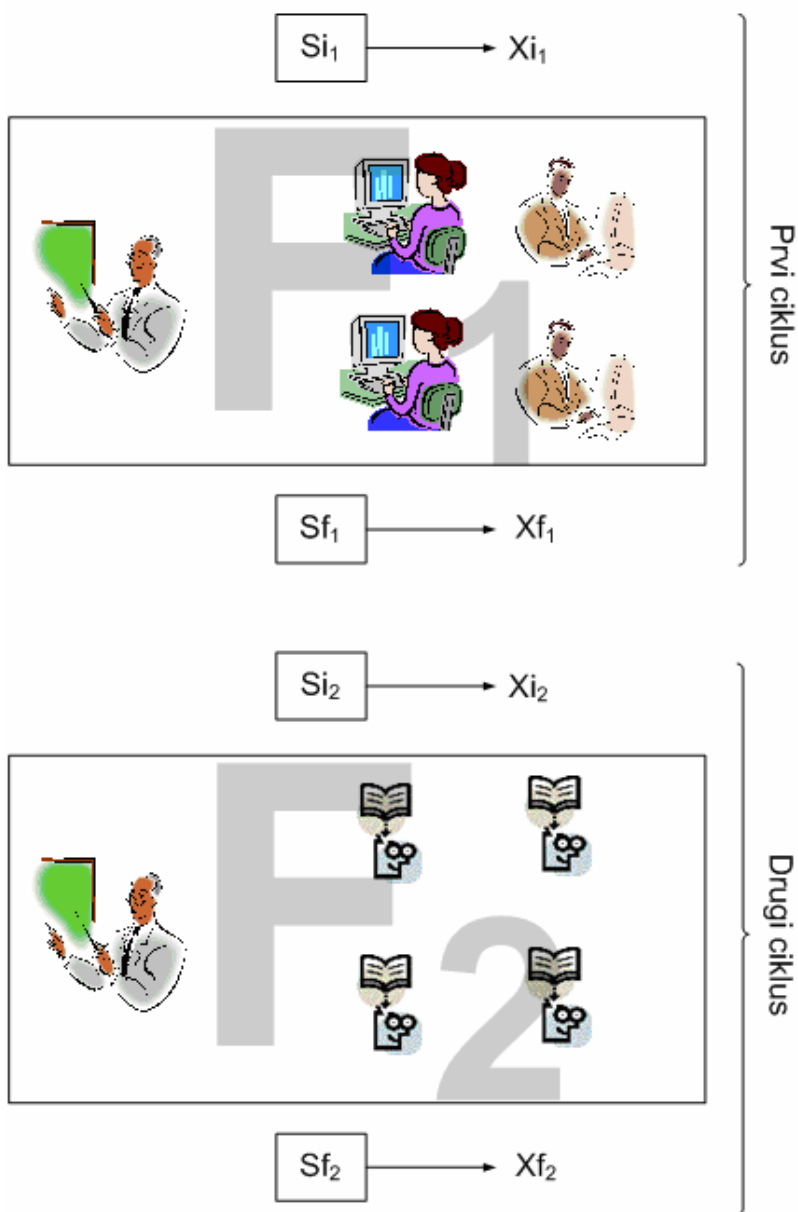
\bar{X}_{i_1} - aritmetička sredina pojedinačnih inicijalnih stanja prije uvođenja prvog faktora

F_1 - djelatnost kojom se uvodi prvi faktor

S_{f_1} - završno stanje nakon uvođenja prvog faktora

\bar{X}_{f_1} - aritmetička sredina pojedinačnih inicijalnih stanja nakon uvođenja prvog faktora

Analogni su simboli za drugi faktor tijekom drugog ciklusa.



Slika 3. 9. Eksperiment s jednom grupom

Usporedba učinka tih dvaju faktora vrši se pomoću usporedbe njihovih „čistih“ učinaka, tj. oduzimanjem razlike u prosječnom napretku pri primjeni prvog i pri primjeni drugog faktora. Dakle, razlika D je

$$D_{\bar{X}} = \bar{X}_{F_1} - \bar{X}_{F_2} = (\bar{X}_{f_1} - \bar{X}_{i_1}) - (\bar{X}_{f_2} - \bar{X}_{i_2}) \quad (3.50)$$

Iz činjenice da se u jednokratnoj usporedbi dvaju faktora pojavila diferencija u korist jednog od njih ne može se izvesti zaključak da je učinak jednog faktora općenito veći. Stoga, ako se radi o eksperimentu kojemu je cilj generalizacija, treba upotrijebiti metodu uzoraka.

U eksperimentu s jednom grupom dolazi do izražaja nemogućnost objektivne usporedbe eksperimentalnih faktora i sljedećih nedostataka, koji se teško mogu izbjeći ako postoji više od dva ciklusa:

1. U eksperimentu s jednom grupom neophodno je da se eksperimentalni faktori uspoređuju na različitom područnom znanju. Da bi se omogućila usporedba učinka faktora, područno znanje u jednom i drugom ciklusu mora biti različito i u što većoj mjeri nepovezano. Stupanj usvojenosti područnog znanja prvog ciklusa ne bi smio djelovati na lakše usvajanje područnog znanja iz drugog ciklusa. To je veoma teško postići.
2. Čest je nedostatak i nejednaka težina područnog znanja iz pojedinih ciklusa. Podjednaka težina područnog znanja bitan je uvjet za usporedbu učinaka eksperimentalnih faktora.
3. Da se omogući usporedba učinaka raznih eksperimentalnih faktora, treba raspolagati mjernim instrumentima kojima se može utvrditi napredak s primjenom jednog faktora i izvršiti međusobna usporedba rezultata tih ciklusa.
4. Pouzdanost usporedbe između napretka u pojedinim ciklusima može poremetiti i činjenica da se ispitanici i bez obzira na djelovanje eksperimentalnih faktora stalno razvijaju i mijenjaju. Stoga se eksperiment s jednom grupom može primijeniti samo ako će trajati do mjesec dana.

Iz navedenog se vidi da se model eksperimenta s jednom grupom za uspoređivanje više eksperimentalnih faktora može u znanstvenom radu teško uspješno primijeniti. Naime, primjena metode uzoraka je neizbježna ako se želi doći do pouzdanog zaključka o mogućnosti generalizacije, zahtijeva znatan utrošak rada i financijskih sredstava. Stoga je neracionalno da se metoda uzoraka poveže s relativno primitivnim i nepouzdanim modelom eksperimenta s jednom grupom.

Ipak, ovaj model ima svoje mjesto jer je važno oruđe nastavnika u njegovom osobnom nastojanju da pronađe one postupke kojima postiže najbolje rezultate i koji najbolje odgovaraju njemu i njegovim učenicima.

EKSPERIMENT S PARALELNIM GRUPAMA

Želimo li izbjeći slijedno uvođenje eksperimentalnih faktora i probleme koje ta sukcesivnost donosi, te uvesti istovremeno te faktore čije učinke želimo usporediti, potrebno je umjesto jedne uzeti više grupa: za svaki faktor po jednu. Na taj način se dobiva model eksperimenta s paralelnim grupama. Tako je put provođenja takvog eksperimenta s dva faktora sljedeći:

GRUPA A: INICIJALNO STANJE – DJELOVANJE PRVOG FAKTORA – ZAVRŠNO STANJE
istovremeno
GRUPA B: INICIJALNO STANJE – DJELOVANJE DRUGOG FAKTORA – ZAVRŠNO STANJE

Dok se kod eksperimenta s jednom grupom radi o slijednim ciklusima koji se ostvaruju s istim učenicima, kod eksperimenta s paralelnim grupama radi se o simultanom radu s različitim grupama učenika.

Da bi put ovakvog eksperimenta označili shematski (Slika 3. 10.), potrebno je uvesti sljedeće oznake:

S_{iA} - inicijalno stanje prije uvođenja faktora u grupu A

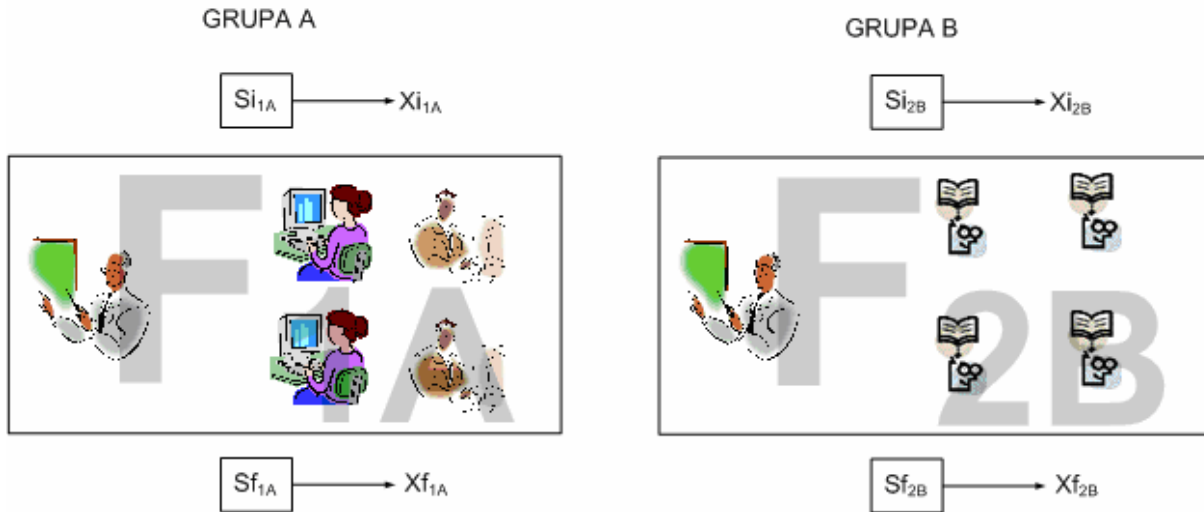
\bar{X}_{iA} - aritmetička sredina pojedinačnih inicijalnih stanja prije uvođenja faktora u grupu A

F_A - djelatnost kojom se uvodi faktor u grupu A

S_{fA} - završno stanje nakon uvođenja faktora u grupu A

\bar{X}_{fA} - aritmetička sredina pojedinačnih inicijalnih stanja nakon uvođenja faktora u grupu A

Analogni su simboli za faktor koji se simultano uvodi u grupu B.



Slika 3. 10. Eksperiment s paralelnim grupama

Usporedba učinka tih dvaju faktora vršit će se pomoću usporedbe njihovih „čistih“ učinaka, tj. oduzimanjem razlike u prosječnom napretku pri primjeni prvog te pri primjeni drugog faktora. Dakle, razlika D je

$$D_{\bar{X}} = \bar{X}_{fA} - \bar{X}_{fB} = (\bar{X}_{fA} - \bar{X}_{iA}) - (\bar{X}_{fB} - \bar{X}_{iB}) \quad (3.51)$$

Kad se radi o eksperimentu na osnovu kojeg se želi izvršiti generalizacija, trebat će primijeniti metodu uzorka.

U gornjoj se shemi uspoređuju dva faktora, pa se eksperiment vrši s dvije grupe učenika. Ako se želi usporediti više faktora, treba raditi s isto toliko grupa. U svakoj se grupi, na temelju razlike između završnog i inicijalnog stanja, određuje učinak eksperimentalnog faktora i tako će se ustanoviti koji je eksperimentalni faktor najučinkovitiji. Kad se statističkim postupcima određuje u kojoj se mjeri mogu ovi rezultati generalizirati, polazi se od analize varijance. Njom se ispituje

da li uopće postoji među pojedinim eksperimentalnim faktorima tolika razlika da se, s dovoljno visokim postotkom vjerojatnosti, ne može pripisati djelovanju slučaja, a zatim se ispituje razlika između svakog od eksperimentalnih faktora.

Ponekad se kod eksperimenata s dvije paralelne grupe uspoređuje učinak nekog novijeg, modernijeg nastavnog postupka s nekim starijim, uobičajenijim (npr. sustav e-učenja i grupni rad). Tada se grupa u koju se uvodi novi postupak naziva *eksperimentalna*, a ona u koju se uvodi uobičajeni postupak *kontrolna* grupa. Različiti nazivi ovih grupa već unaprijed sugeriraju zaključak o raznoj vrijednosti faktora, što se protivi samoj biti eksperimenta kao objektivnog ispitivanja postavljene hipoteze. Naziv „kontrolna grupa“ ima opravdanje jedino kao sinonim za „nul-grupa“. U tu se grupu zbog kontrole ne uvodi nijedan faktor (npr. sustav e-učenja – eksperimentalna grupa, tradicionalna nastava – kontrolna grupa). Pri usporedbi dvaju faktora koji se uvode u dvije grupe, a uz postojanje treće, kontrolne grupe, put će biti sljedeći:

$$\text{istovremeno} \left\{ \begin{array}{l} \text{prva grupa} : S_{i1} \rightarrow F_1 \rightarrow S_{f1} \\ \text{druga grupa} : S_{i2} \rightarrow F_2 \rightarrow S_{f2} \\ \text{treća grupa} : S_{i0} \rightarrow \rightarrow \rightarrow S_{f0} \end{array} \right.$$

Slika 3. 11. Eksperiment s tri paralelne grupe

Najvažniji problem u vezi s ovim modelom je da li i kako stvoriti podjednake, ekvivalentne grupe da bismo, u skladu sa zakonom jedne varijable, mogli dobiti što čistiju razliku u učinak eksperimentalnih faktora. Izjednačavanje grupa se vrši izjednačavanjem onih varijabli koje se odnose na osobnost ispitanika, tj. koje predstavljaju izvore „S“ pogreški. Ono se obično vrši na jedan od sljedeća dva načina:

1. izjednačavanjem karakteristike distribucija tih varijabli među grupama. Na primjer, kod rezultata testova nastoji se izjednačiti mjera centralne tendencije (obično aritmetička sredina) i mjera disperzije kao pokazatelj homogenosti podataka (obično standardna devijacija).
2. ako su samo dvije grupe u eksperimentu, vrši se još i određivanje parova ispitanika, odnosno uzima se po jedan učenik iz svake grupe kod kojih su podaci u vezi s pojedinim varijablama što sličniji. Na taj se način može povisiti preciznost eksperimenta koristeći pozitivnu korelaciju između kontrolne i eksperimentalne varijable.

Učenici previše odskaču iz neke distribucije ili koji se ne mogu spariti, izbacuju se iz eksperimenta.

Prednosti modela eksperimenta s paralelnim grupama u odnosu na model eksperimenta s jednom grupom:

1. Nema opasnosti da raniji faktor djeluje na kasniji, jer se oba uvode simultano u različitim grupama.
2. Nema opasnosti djelovanja razlike područnog znanja jer je ista.
3. Nema opasnosti djelovanja razlike u mjernim instrumentima jer su isti.

Specifične poteškoće vezane za ovaj model uzrokovane su težnjom da se kontrolira i izjednači što veći broj varijabli:

1. Velik je broj varijabli čije inicijalno stanje treba kontrolirati pa je teško utvrditi vjernu sliku inicijalnog stanja u grupama.

2. Ostvarivanje ekvivalentnosti često zahtijeva smanjenje broja ispitanika čiji se rezultati mogu u obradi uzeti u obzir što dovodi do sniženja preciznosti eksperimenta.
3. Postoji opasnost da se varijable koje smo u početku podvrgli uspješnoj kontroli za vrijeme samog eksperimenta bitno promijene (npr. bolest učenika u jednoj grupi).

Usprkos poteškoćama, eksperiment s paralelnim grupama se upotrebljava češće od drugih modela. Spomenute poteškoće nastoji se izbjeći složenijim modelima.

EKSPERIMENT S ROTACIJOM FAKTORA

Ovaj model predstavlja kombinaciju eksperimenta s jednom grupom i eksperimenta s paralelnim grupama. Osnovna mu je svrha da se izbjegniju neke specifične teškoće jednog i drugog modela. U eksperimentu s jednom grupom faktori se uvode slijedno, pa imamo dva ili više ciklusa. U eksperimentu s paralelnim grupama uvode se simultano, pa postoje dvije ili više paralelnih grupa. Kod eksperimenta s rotacijom faktora imamo dvije ili više paralelnih grupa u kojima se faktori, rotacijom, još i ciklički izmjenjuju.

Tok ovakvog eksperimenta uz pretpostavku dvaju faktora je sljedeći:

Prvi ciklus:
GRUPA A: INICIJALNO STANJE – DJELOVANJE PRVOG FAKTORA – ZAVRŠNO STANJE
istovremeno
GRUPA B: INICIJALNO STANJE – DJELOVANJE DRUGOG FAKTORA – ZAVRŠNO STANJE
nakon toga
Drugi ciklus:
GRUPA A: INICIJALNO STANJE – DJELOVANJE DRUGOG FAKTORA – ZAVRŠNO STANJE
istovremeno
GRUPA B: INICIJALNO STANJE – DJELOVANJE PRVOG FAKTORA – ZAVRŠNO STANJE

I ovaj se slijed može prikazati shemom (Slika 3. 12.). Uvedimo sljedeće oznake:

$S_{i,A}$ - inicijalno stanje prije uvođenja prvog faktora u grupu A

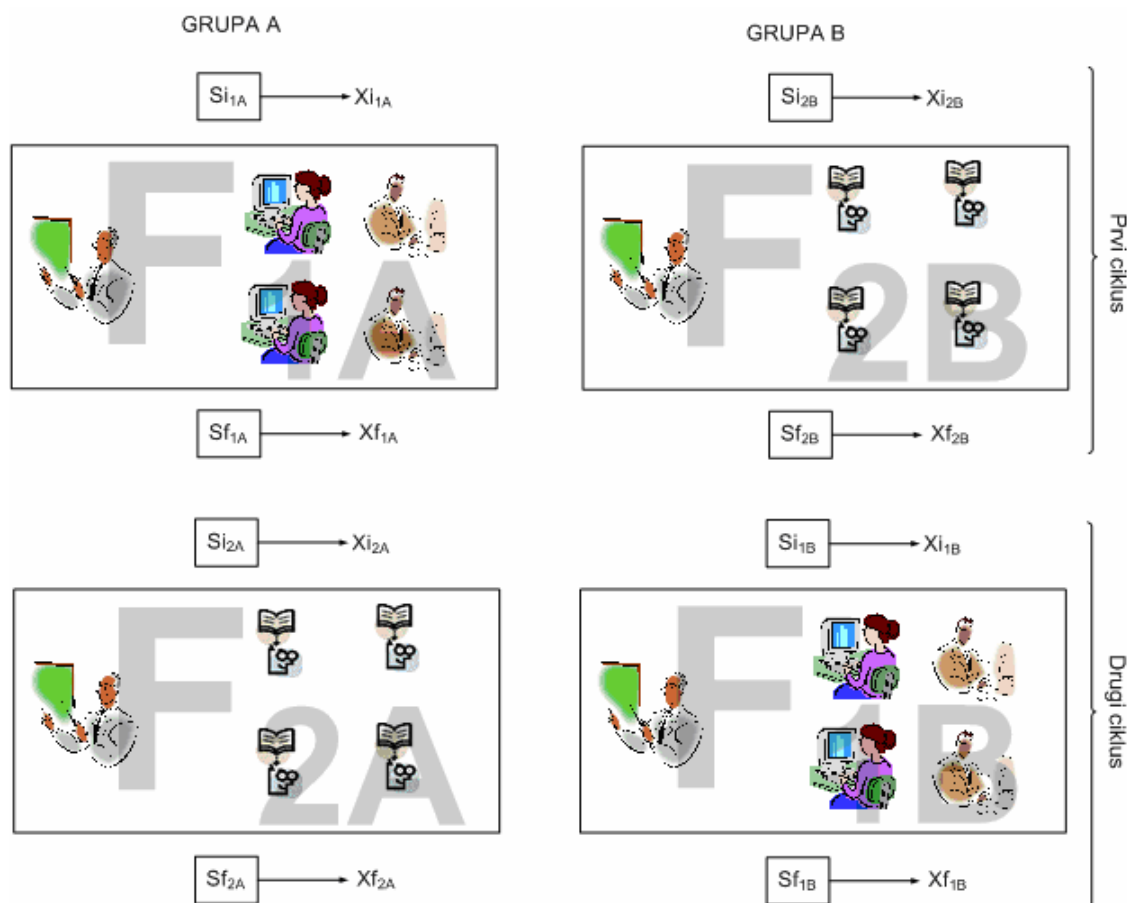
$\bar{X}_{i,A}$ - aritmetička sredina pojedinačnih inicijalnih stanja prije uvođenja prvog faktora u grupu A

$F_{1,A}$ - djelatnost kojom se uvodi prvi faktor u grupu A

$S_{f_1,A}$ - završno stanje nakon uvođenja prvog faktora u grupu A

$\bar{X}_{f_1,A}$ - aritmetička sredina pojedinačnih inicijalnih stanja nakon uvođenja prvog faktora u grupu A

Analogni su simboli za drugi ciklus grupe A i za faktore koji se rotirajući uvode u grupu B.



Slika 3. 12. Eksperiment s rotacijom faktora

Razlika u učinku dva faktora određuje se posebno za svaki ciklus, a način njenog određivanja je jednak onome u eksperimentu s paralelnim grupama. Dakle,

$$\text{za prvi ciklus: } D_{\bar{X}_{prvi}} = \bar{X}_{F_1A} - \bar{X}_{F_2B} \quad (3.52)$$

$$\text{za drugi ciklus: } D_{\bar{X}_{drugi}} = \bar{X}_{F_1B} - \bar{X}_{F_2A} \quad (3.53)$$

Mogla bi se odrediti i ukupna razlika u efikasnosti u oba ciklusa pomoću algebarskog zbroja razlika u svakom ciklusu (uzimajući u obzir predznake). Kako se u svakom ciklusu radi o drugom područnom znanju i u svakom se ciklusu učinak mjeri drugim mjernim instrumentom, takvo bi zbrajanje obično predstavljalo zbrajanje dviju različitih stvari.

Usporedimo li shemu prikaza eksperimenta s rotacijom faktora sa shemom eksperimenta s jednom grupom te sa shemom eksperimenta s paralelnim grupama, vidi se da je eksperiment s rotacijom faktora sinteza ovih drugih dvaju modela. Promatrajući shemu u okomitom pravcu imamo dva eksperimenta s jednom grupom, a promatramo li je u vodoravnom pravcu imamo dva eksperimenta s paralelnim grupama.

Kod modela s rotacijom faktora nastoje se riješiti problem druga dva modela, tj. problem različitih područnih znanja i problem ekvivalencije grupa. Ako se pomoću rotacije faktori uvode pri usvajanju dva različita područna znanja i ako se pri tom u oba slučaja pokaže isti faktor učinkovitiji, možemo zaključiti o prednosti tog faktora bez obzira na razliku između ta dva područna znanja. Slično tome, ako isti faktor pokaže prednost i kad je, u prvom ciklusu, primijenjen u prvoj grupi, i kad je, u drugom ciklusu, primijenjen u drugoj grupi, opravdano je zaključiti da je taj faktor bolji bez obzira na eventualnu neekvivalentnost grupa. Iz toga proizlazi da eksperiment s rotacijom faktora daje jedan rezultat jedino onda kad isti faktor pokaže svoju prednost. Zato se razlika u učinku izračunava posebno za svaki ciklus, a ne određuje se ukupna razlika.

Pozitivna karakteristika ovog modela je to što on omogućuje, u slučaju kad se u svim ciklusima pokaže stalna i značajna prednost jednog faktora, jedan zaključak o većem njegovom učinku čak i u slučaju da se nisu mogli u dovoljnoj mjeri eliminirati izvori pogrešaka, kao što su međusobni utjecaj područnih znanja pojedinih ciklusa i nepostojanje inicijalne ekvivalentnosti grupa.

Negativne karakteristike ovog modela se očituju u tome što se slučaj kad se u svim ciklusima pokaže stalna i značajna prednost jednog faktora često neće pojaviti, a ne može se unaprijed ni odrediti da li će se pojaviti. Stoga ipak treba težiti za ekvivalentnošću. Osim toga, provođenje eksperimenta s rotacijom faktora obično je povezano s mnogim organizacijskim poteškoćama. Zbog svega navedenog se ovaj model rijetko upotrebljava jer složeniji eksperimentalni modeli, koje ćemo u nastavku opisati, bolje rješavaju pitanje djelovanja izvora pogrešaka.

3.2.4.4 SLOŽENI MODELI EKSPERIMENTALNIH POSTUPAKA

Složeniji modeli nastoje u što većoj mjeri eliminirati izvore pogrešaka. Ono se očituje u raznim načinima uvođenja eksperimentalnih faktora pri proširenju eksperimenta na veći broj grupa, u nastojanju da se kontrolom istovremenog djelovanja više nezavisnih varijabli prijede klasično poimanje zakona jedne varijable i da se model uskladi sa složenošću pedagoških pojava [MUŽI1977].

Tehnika statističke obrade podataka koja se u ovim modelima upotrebljava zasniva se na *analizi varijance*, te ponekad i *analizi kovarijance*. Ti se postupci primjenjuju da bi se uklonila potreba za stvaranjem ekvivalentnih grupa. Bitna značajka analize kovarijance je u tome da rezultati inicijalnog mjerenja, umjesto da služe stvaranju ekvivalentnih grupa, djeluju na rezultate završnog mjerenja, i to tako da se u postupku analize varijance završnih rezultata vrše izvjesna prilagođavanja s obzirom na inicijalne rezultate. Pojam kovarijance vezan je uz pojam regresije i korelacije. U analizi kovarijance kod spomenutog prilagođavanja polazi se od kovarijance između varijable inicijalnih i varijable završnih rezultata.

PONOVLJENI SLUČAJNI IZBOR GRUPA (ENG. RANDOM REPLICATION DESIGN)

U njemu se zapravo radi o ponavljanju osnovnog modela eksperimenta s paralelnim grupama u većem broju istovrsnih ustanova (na uzorku dobivenom na osnovu zahtjeva metode uzoraka), kako bi se omogućila generalizacija dobivenih rezultata. Pri tome se ne misli na slijedno ponavljanje, već istovremeno izvršavanje eksperimenta u više ustanova.

Osnovna je karakteristika ovog modela da se u svakoj ustanovi uvode svi eksperimentalni faktori. Pri tom se u svako paralelno odjeljenje uvodi po jedan faktor, pri čemu je taj odabir slučajan.

Ovaj je model osobito pogodan kad postoji mogućnost da se u većem broju ustanova s paralelnim odjeljenjima ispita i uspoređi učinak dvaju ili više eksperimentalnih faktora. Ta se odjeljenja uvode u eksperiment kao zatvorene cjeline (eksperimentalne odnosno kontrolne grupe) te se u njima primjenjuju ili ne primjenjuju pojedini eksperimentalni faktori.

Činjenica slučajnog određivanja faktora u pojedine eksperimentalne grupe omogućuje da se odredi vjerojatnost da je jedan eksperimentalni faktor učinkovitiji od drugog, tj. da se ispita hipoteza o razlici učinka tih faktora. Što je veći broj ponavljanja, to će biti pouzdaniji rezultati. Bilo bi dobro da je i izbor ustanova slučajan (što je ponekad teško ostvariti).

I u ovom se modelu nastoji postići što veća ekvivalencija u najvažnijim varijablama. Na taj se način povećava preciznost eksperimenta. Međutim, ukoliko je broj ustanova u kojima se ponavlja eksperiment dovoljno velik, zadovoljavajuća se preciznost eksperimenta može, uz primjenu adekvatnih statističkih postupaka, postići i bez prethodnog izjednačavanja grupa.

U SVAKU SE USTANOVU UVODI SAMO JEDAN FAKTOR (ENG. GROUPS-WITHIN-TREATMENTS DESIGN)

On se koristi ondje gdje u istoj ustanovi nije moguće unositi više različitih eksperimentalnih faktora. Tada takva situacija zahtjeva specifičan model.

Ponekad se u modelu koji je prethodno prikazan javlja okolnost koja može predstavljati ozbiljan izvor pogreške. To je mogućnost razmjene informacija između pripadnika paralelnih grupa u koje se uvode različiti eksperimentalni faktori. Ovakva se pogreška obično ne javlja kod eksperimenata u kojima se u istom danu vrši i inicijalno testiranje i unošenje eksperimentalnih faktora i završno testiranje. Ondje gdje je opravdana pretpostavka da bi takva pogreška mogla bitno utjecati na rezultat eksperimenta, prikladna je primjena modela kojemu je osnovna karakteristika da se u jednu ustanovu uvijek uvodi samo jedan eksperimentalni faktor.

Tada se u jednak broj istovrsnih ustanova uvodi svaki eksperimentalni faktor. Statističkom se analizom ustanovljuje da li je i u kojoj mjeri opći rezultat jednog eksperimentalnog faktora u svim ustanovama u kojima je on primijenjen viši od rezultata ostalih eksperimentalnih faktora. Određivanje ustanova u koje će se uvesti određeni eksperimentalni faktor vrši se slučajnim izborom, a ne prema postojanju boljih preduvjeta za neki eksperimentalni faktor.

Ovaj model zahtijeva, za isti stupanj preciznosti, veći broj ustanova u kojima se eksperiment provodi nego što to zahtijevaju raniji modeli. To je potrebno zbog toga što se u svaku ustanovu uvodi samo jedan eksperimentalni faktor čiji se učinak određuje, ali i zbog toga što su razlike između istovrsnih ustanova mnogo veće nego razlike unutar odjeljenja iste ustanove.

ISTOVREMENO PRAĆENJE VIŠE VARIJABLI U ISTOM EKSPERIMENTU (FAKTORSKI MODEL, ENG. FACTORIAL DESIGN)

Taj je model specijalno namijenjen slučaju kad se, prelazeći doslovni smisao zakona jedne varijable, u istom eksperimentu prati djelovanje većeg broja eksperimentalnih varijabli. Na taj se način izbjegava i potreba za ekvivalentnošću grupa u tim varijablama.

Zahvaljujući razvoju suvremene statistike, eksperimentalno istraživanje se obogatilo mogućnošću istovremenog, jednim eksperimentom, praćenja djelovanja faktora različitih varijabli na istu pojavu, kao i međusobnu interakciju pojedinih varijabli, što daje potpuniju sliku nego modeli u kojima se strogo nastoji ostvariti zakon jedne varijable.

U svim se grupama, čiji je broj određen umnoškom faktora varijabli, primjenjuje inicijalni test, zatim se uvode eksperimentalni faktori i na kraju primjenjuje završni test, kao i u osnovnom modelu s paralelnim grupama. Nakon toga se složenom statističkom analizom utvrđuje visina vjerojatnosti da su pojedini eksperimentalni faktori učinkovitiji ili manje učinkoviti bez obzira na djelovanje drugih varijabli. Nakon toga se utvrđuje u kolikoj mjeri je njihov učinak povezana s utjecajem drugih varijabli. Isto tako se ispituje u kojoj se mjeri pojavljuje interakcija tih ostalih varijabli.

LATINSKI I LATINSKO-GRČKI KVADRAT (ENG. LATIN SQUARE DESIGN, GRAECO-LATIN SQUARE DESIGN)

Ovi modeli predstavljaju specifične primjere pojednostavnjenja prethodnog modela.

Broj grupa u faktorskom modelu eksperimenta je jako velik jer se radi o umnošku broja eksperimentalnih faktora svake varijable. Glomaznost takvih eksperimenata predstavlja ozbiljnu prepreku njihovom ostvarenju. Međutim, ako se radi o varijablama kod kojih nas zanima samo osnovno djelovanje svake od njih, a ne i njihove interakcije, te imamo razloga za pretpostavku da tih interakcija uopće nema, moguće je faktorski model eksperimenta bitno pojednostavniti pomoću modela latinskog kvadrata.

Ime modela *latinskog kvadrata* potječe iz stare matematičke zagonetke kojom je trebalo odrediti na koliko načina se mogu rasporediti slova latinske abecede u kvadratnoj tablici, a da se pri tom svako slovo pojavljuje u svakom stupcu i u svakom retku samo jedanput (vidi Slika 3. 13.).

A	B	C
B	C	A
C	A	B

Slika 3. 13. Latinski 3x3 kvadrat

Preuzimajući bitni uvjet te stare zagonetke, eksperimentalni se model latinskog kvadrata, za primjer 4 eksperimentalna faktora u 3 varijable, zasniva na shemi u kojoj se u svakom retku i

svakom stupcu, uz kombinaciju prve i druge varijable, javlja samo jedanput svaki faktor iz treće varijable (vidi Tablicu 3.8.). Na taj se način svaki faktor iz svake varijable susreće samo jedanput sa svakim faktorom iz ostalih dviju varijabli, a takvom se uravnoteženošću postiže pouzdanija procjena učinka svakog pojedinog faktora u svakoj od varijabli. Na ovaj način se dobila velika ušteda: umjesto $4 \times 4 \times 4 = 64$ grupe (3 varijable po 4 faktora), potrebno je samo $4 \times 4 = 16$ grupa (4 retka i četiri stupca) (Tablica 3. 8.).

Tablica 3. 8. Latinski 3x3 kvadrat

	S ₁	S ₂	S ₃	S ₄
R ₁	T ₁	T ₂	T ₃	T ₄
R ₂	T ₂	T ₃	T ₄	T ₁
R ₃	T ₃	T ₄	T ₁	T ₂
R ₄	T ₄	T ₁	T ₂	T ₃

Kad postoje razlozi za pretpostavku naglašene interakcije između varijabli, a osobito ako je broj faktora u varijablama malen (npr. u 2x2 modelu), čime se smanjuje mogućnost međusobne neutralizacije djelovanja interakcije, ne može se primijeniti model latinskog kvadrata.

U latinskom kvadratu radi se o tri varijable. Ukoliko se po istom principu želi dodati i četvrta varijabla, dobiva se *grčko-latinski kvadrat*. Sastoji se u tome da se različite veličine četvrte varijable (u modelu predstavljene grčkim slovima) unose u polja kvadrata tako da se uz svako latinsko slovo pojavljuju različita grčka slova, svako samo jedanput (vidi Slika 3. 14.).

A α	B β	C γ	D δ
B γ	A δ	D α	C β
C γ	D γ	A β	B α
D β	C α	B δ	A γ

Slika 3. 14. Grčko-latinski 4x4 kvadrat

Za primjer eksperimenta 4 varijable s po 4 faktora imali bismo opet samo 16 grupa prema sljedećoj shemi (Tablica 3. 9.):

Tablica 3. 9. Grčko-latinski 4x4 kvadrat

	S ₁	S ₂	S ₃	S ₄
R ₁	T ₁ P ₁	T ₂ P ₂	T ₃ P ₃	T ₄ P ₄
R ₂	T ₂ P ₂	T ₃ P ₄	T ₄ P ₁	T ₁ P ₂
R ₃	T ₃ P ₂	T ₄ P ₃	T ₁ P ₂	T ₂ P ₂
R ₄	T ₄ P ₂	T ₁ P ₁	T ₂ P ₄	T ₃ P ₃

3.3. VELIČINA UČINKA

Dok testovi značajnosti govore o vjerojatnosti kojom se rezultati na testovima eksperimentalne grupe razlikuju od slučajnih rezultata, veličina učinka (eng. effect size) govore o relativnoj jačini eksperimentalnog faktora. Veličina učinka omogućava prijelaz s jednostavnog pitanja istraživanja „Da li je nešto dobro ili ne?“ na složenije pitanje „Koliko je nešto dobro?“.

Veličina učinka kvantificira veličinu razlike između dvije grupe te se stoga smatra stvarnom mjerom značajnosti te razlike. Veličina učinka je standardni način za iskazivanje rezultata različitih eksperimenata pomoću uniformne ljestvice učinkovitosti, čime ona postaje ulazni parametar postupka nazvanog *meta-analiza*, te omogućava usporedbu jačine eksperimentalnih faktora. Veličina učinka može biti pozitivna ili negativna. Pozitivna je kada je eksperimentalna grupa bolja, a negativna je kada je kontrolna grupa bolja.

Glavni cilj sustava e-učenja je učenje, a glavni način vrednovanja učinka sustava e-učenja je ispitivanje da li učenici učinkovito uče uz pomoć tog sustava (prema [MARK1993]). Kriteriji koji se koriste za vrednovanje utjecaja sustava e-učenja na obrazovanje uključuju: postignuće (eng. *achievement*) i afekt (eng. *affect*). Postignuće se odnosi na stjecanje, razumijevanje, performanse, zadržavanje i prijenos učenikovog znanja i vještina (Haertel i Calfee, prema [MARK1993]). Afektivne mjere se odnose na stavove i emocije koje mogu utjecati na način kako učenici koriste sustave e-učenja i uče od njih (Malone, prema [MARK1993]).

3.3.1. MJERE POSTIGNUĆA

Znanje se može podijeliti na kategorije: referencijalno (eng. *referential*), činjenično (eng. *factual*) i proceduralno (eng. *procedural*) [MARK1993]. Znanje o simbolima i značenjima se naziva *referencijalno* znanje ili semantičko znanje. *Činjenično* znanje je znanje o objektima i vezama između objekata unutar svijeta. *Proceduralno* znanje je o tome kako nešto raditi. Dijeli se na eksplicitno i implicitno. Eksplicitno proceduralno znanje se može opisati u terminima algoritama, pravila ili procedura (Mayer, prema [MARK1993]), te se naziva i intelektualna vještina i deklarativno znanje. Implicitno proceduralno znanje je znanje o tome kako nešto raditi što se ne može opisati riječima. Mayer to naziva vještinama (prema [MARK1993]).

Bloom [BLOO1956] je napravio taksonomiju ciljeva obrazovanja, općih kategorija rezultata učenja koje se mogu povezati s pokazateljima performansi. Određuje šest glavnih tipova ciljeva koje naziva znanje, razumijevanje, primjena, analiza, sinteza i vrednovanje. Bloomovo znanje se naziva *prisjećanje* (eng. *recall*) označava mogućnost učenika da se prisjeti nekog termina, procedure ili činjenice, ali se ne radi o razumijevanju i primjeni naučenog. *Razumijevanje* (eng. *comprehension*) označava da učenik može do određenog stupnja koristiti naučeni materijal, što je

dovoljno za definicije i izvlačenje izravnih zaključaka. Učenik počinje upotrebljavati znanje u konkretnim situacijama na razini *primjene* (eng. *application*). *Analiza* označava da učenik može odrediti ideje, ali i ispitati njihove međusobne veze. *Sinteza* označava da učenik može organizirati znanje za generiranje novih ideja. *Vrednovanje* uključuje mogućnost procjene vrijednosti znanja. Spomenute kategorije znanja i razine razumijevanja mogu služiti kao kriteriji pri određivanju mjera koje će se koristiti za vrednovanje postignuća. Naime, ove kategorije znanja i rezultata učenja čine taksonomiju tipova ponašanja koja pokazuju da li je učenik naučio ono što ga se poučavalo.

Drugi kriteriji koji se mogu koristiti pri vrednovanju obrazovnog utjecaja uključuju prijenos, zadržavanje, vrijeme učenja i brzinu završavanja. Prva tri kriterija se mogu procijeniti eksperimentalno, a brzine završavanja se promatraju tijekom isprobavanja sustava.

Prijenos (eng. *transfer*) je sposobnost primjene informacija ili vještina naučenih u jednom kontekstu u novi i obično nepoznati kontekst, gdje je to znanje također važno (Mayer, prema [MARK1993]). Prijenos podrazumijeva generalizaciju ili apstraktno razumijevanje materijala. Važno je da kad učenik uči nešto npr. u simuliranoj okolini da to može primijeniti i u stvarnoj okolini. Učenika se testira u situaciji u kojoj uči i u novoj situaciji kada to znanje treba primijeniti (Johnson, prema [MARK1993]).

Zadržavanje (eng. *retention*) je sposobnost zadržavanja naučenog tijekom vremena (Mayer, prema [MARK1993]). Smatra se da je učenik bolje naučio neko znanje što se duže može toga prisjetiti. To bi također bio dokaz prednosti nekog pristupa poučavanju. Ispituje se na način da se mjeri isto znanje tijekom vremena. Npr. dva testa iz matematike, koji naravno nisu isti, ali se mogu usporediti, se daju učenicima u dva različita vremenska razdoblja (možda na početku i kraju polugodišta), kako bi se vidjelo je li se njihovo znanje promijenilo tijekom vremena, neovisno o poučavanju.

Vrijeme učenja (eng. *time*) je vrijeme potrebno da bi se nešto naučilo [MARK1993]. Ponekad se vrijeme učenja na različitim sustavima uspoređuje kako bi se pokazalo je li jedan sustav bolji u odnosu na drugi. Vrijeme učenja se mora gledati i u kontekstu postignuća i motivacije.

Brzina završavanja (eng. *completion rate*) ili *odustajanje* (eng. *drop-out*) se promatra prilikom testiranja sustava u stvarnim uvjetima kada učenici mogu prestati ili nastaviti koristiti program za poučavanje (Hoecker i Elias, prema [MARK1993]). Npr. ako je neki sustav težak za upotrebu onda može i obeshrabriti učenike, te ih tako navesti da odustanu. Čak i ako učenici koji koriste takav sustav postižu bolje rezultate, a većina učenika ga smatra nepoželjnim i teškim za upotrebu, onda ipak nije dobar izbor za poučavanje. Također, može se umjesto pojma "odustajanje" reći da učenici osjećaju *otpor* (eng. *attrition*) (Phipps i Merisotis, Terry, prema [WISH2003]).

3.3.2. MJERE AFEKTA

Pri vrednovanju učinka sustava e-učenja treba uzeti u obzir i motivaciju ili želju učenika da budu aktivni i uključeni u poučavanje. Treba razmišljati o interesu učenika za uključivanjem u aktivnosti poučavanja, interesu za postizanje nekog uspjeha ili željenog cilja i o stavovima prema obrazovanju, računalima ili određenim programima (Thomas, prema [MARK1993]). Uobičajeni

način vrednovanja je da se zatraži od učenika da ocjene koliko im odgovaraju određene aktivnosti (Moore, prema [MARK1993]). Također se uspoređuje i vrijeme provedeno na aktivnostima koje su vezane i koje nisu vezane uz zadatak.

Afektivne mjere se mogu koristiti da bi se vidjelo što učenici misle o sustavu, ali one nužno ne odražavaju postignuće. Mogu se koristiti za određivanje problema, ali nisu dovoljno pouzdane da bi mogli biti sigurni da će postojeći problemi uistinu i biti uočeni. Prema [MARK1993] mogu se koristiti kao prijedlog o prihvaćanju ili neprihvaćanju ITS-a. Mogu se koristiti kao dodatne mjere prilikom vrednovanja postignuća, ali nisu uvjerljiv dokaz o učinku ako se koriste same.

3.3.3. VRSTE VELIČINA UČINKA

Veličina učinka se može izračunati uz pomoć različitih formula i pristupa i njena vrijednost se može zbog toga uvelike razlikovati. Postoje četiri vrste veličina učinka: razlika aritmetičkih sredina (eng. mean difference), standardna razlika aritmetičkih sredina (eng. standardized mean difference) ili Glass-ova Δ , Cohen-ov d , Hedges-ov g . Za utvrđivanje razlika između grupa u eksperimentalnom istraživanju Mohammad [MOHA1998] preporuča uporabu standardne razlike aritmetičkih sredina.

3.3.3.1 RAZLIKA ARITMETIČKIH SREDINA

Najjednostavnija mjera učinka nekog eksperimentalnog faktora je razlika aritmetičkih sredina. Ona je najprikladnija ako u istraživanju postoje samo dvije grupe [MAXW2004]. Računa se po formuli:

$$\Delta = \bar{X}_e - \bar{X}_c \quad (3.54)$$

gdje su \bar{X}_e i \bar{X}_c aritmetičke sredine redom eksperimentalne i kontrolne grupe.

3.3.3.2 STANDARDNA RAZLIKA ARITMETIČKIH SREDINA – GLASS Δ

Da bi se mogle uspoređivati veličine učinaka iz različitih istraživanja, Glass je 1976. godine predložio uvođenje standardizirane veličine učinka koja se naziva Glass-ova Δ (prema [MAXW2004]). Rješenje je pronađeno u analogiji s izračunom z -vrijednosti: dijeljenjem sa standardnom devijacijom. Standardna razlika aritmetičkih sredina računa se tako što se razlika između aritmetičkih sredina eksperimentalne i kontrolne grupe podijeli sa standardnom devijacijom kontrolne grupe. Koristimo sljedeću formulu (prema [MOHA1998]):

$$\Delta = \frac{\bar{X}_e - \bar{X}_c}{\sigma_c} \quad (3.55)$$

gdje su \bar{X}_e i \bar{X}_c aritmetičke redom eksperimentalne i kontrolne grupe te σ_c standardna devijacija kontrolne grupe. Albacete i VanLehn [ALBA2000] navode u svom istraživanju da su kod ovakvog izračunavanja veličine učinka \bar{X}_e i \bar{X}_c aritmetičke razlika rezultata inicijalnog i završnog testa redom eksperimentalne i kontrolne grupe te σ_c standardna devijacija razlika rezultata inicijalnog i završnog testa kontrolne grupe

Glass, McGaw i Smith [GLAS1981] smatraju da eksperimentalni faktor može utjecati na varijaciju zavisne varijable kao i na njenu aritmetičku sredinu, pa stoga preporučuju uporabu samo standardne devijacije kontrolne grupe.

Fletcher-Flinn i Gravatt (prema [MOHA1998]) su naveli da se veličina učinka može izračunati i pomoću veličina učinka inicijalnog i završnog testa koristeći formulu:

$$\Delta = \Delta(\text{završni test}) - \Delta(\text{inicijalni test}) \quad (3.56)$$

te pomoću t-vrijednosti:

$$\Delta = t \sqrt{\frac{1}{N_e} + \frac{1}{N_c}} \quad (3.57)$$

Hedges (prema [LIPS2001]) smatra da veličina učinka izračunata kao standardna razlika aritmetičkih sredina pokazuje povećanu pristranost kada se radi s malim uzorcima (pogotovo s uzorcima manjima od 20). Stoga je on definirao formulu za „nepristranu“ veličinu učinka:

$$\sigma' = \left(1 - \frac{3}{4N - 9}\right) \cdot \sigma \quad (3.58)$$

gdje je σ „pristrana“ veličina učinka, a N ukupna veličina uzorka.

3.3.3.3 COHEN d

Ova veličina učinka ima dvije prednosti naspram ostalih [THAL2002]. Prvo, njena popularnost gotovo je čini standardnom. Drugo, Cohen je definirao granice vrijednosti veličine učinka, kod istraživanja s dvije grupe, i smatra da su vrijednosti oko 0.2 su male, oko 0.5 umjerene, a oko 0.8 velike [COHE1969]. To omogućava usporedbu veličine učinka eksperimentalnog faktora s poznatim referentnim vrijednostima. Cohen-ova veličina učinka se računa po formuli:

$$d = \frac{\bar{X}_e - \bar{X}_c}{\sigma_{\text{zajednicka}}} \quad (3.59)$$

gdje je $\sigma_{zajednicka}$ zajednička standardna devijacija *populacija* eksperimentalne i kontrolne grupe. Cohen-ova veličina učinka se može izračunati i po sljedećim formulama [THAL2002]:

$$d = t \sqrt{\left(\frac{N_e + N_c}{N_e N_c} \right) \left(\frac{N_e + N_c}{N_e + N_c - 2} \right)} \quad (3.60)$$

gdje je t vrijednost t-testa,

$$d = \sqrt{F \left(\frac{N_e + N_c}{N_e N_c} \right) \left(\frac{N_e + N_c}{N_e + N_c - 2} \right)} \quad (3.61)$$

gdje je F vrijednost F-testa.

3.3.3.4 HEDGES g

Hedges-ov g je veličina učinka koja se računa po formuli:

$$g = \frac{\bar{X}_e - \bar{X}_c}{\sigma_{zajednicka}} \quad (3.62)$$

gdje je $\sigma_{zajednicka}$ zajednička standardna devijacija eksperimentalne i kontrolne grupe.

Prema [PETZ1997] ako su i kontrolna i eksperimentalna grupa jednako velike, nije ni potrebno testirati F-testom da li im se varijance statistički značajno razlikuju ili ne, jer će se – ako se one razlikuju – pogreška u računu biti neznatna. Ako se te grupe razlikuju u veličini, onda treba ispitati statističku značajnost razlike između aritmetičkih sredina malih uzoraka. Zajedničku standardnu devijaciju smijemo izračunati samo onda ako se standardne devijacije samih grupa zaista značajno ne razlikuju, što se ispituje pomoću tzv. F-testa dijeleći varijance grupa po formuli:

$$F = \frac{\text{veća } \sigma^2}{\text{manja } \sigma^2} \quad (3.63)$$

Ako su grupe jednako velike ili ako im se varijance ne razlikuju značajno, pristupamo izračunavanju *zajedničke standardne devijacije* (eng. pooled) koja se izračunava prema formuli:

$$\sigma_{zajednicka} = \sqrt{\frac{(N_e - 1)\sigma_e^2 + (N_c - 1)\sigma_c^2}{N_e + N_c}} \quad (3.64)$$

gdje su σ_e i σ_c standardne devijacije redom eksperimentalne i kontrolne grupe, a N_e i N_c veličine redom eksperimentalne i kontrolne grupe. Zajednička standardna devijacija se koristi u slučaju

da je kontrolna grupa mala pa ne može predstavljati standardnu devijaciju populacije ili u slučaju da nema „prave“ kontrolne grupe.

Zajednička standardna devijacija se može izračunati i po formulama:

$$\sigma_{zajednicka} = s_{\bar{x}_e - \bar{x}_c} \sqrt{N} \quad (3.65)$$

gdje je $s_{\bar{x}_e - \bar{x}_c}$ standardna pogreška između dviju aritmetičkih sredina

$$\sigma_{zajednicka} = \sqrt{\frac{\sum (X - \bar{X}_c)^2 + \sum (X - \bar{X}_e)^2}{(N_c - 1) + (N_e - 1)}} \quad (3.66)$$

Cohen-ova veličina učinka se može izračunati i po sljedećim formulama [BACK2000]:

$$g = t \sqrt{\frac{N_e + N_c}{N_e N_c}} \quad (3.67)$$

$$g = \frac{2t}{\sqrt{N}} \quad (3.68)$$

gdje je t vrijednost t-testa.

3.3.4. INTERVALI POUZDANOSTI

Jasno je da, ako je veličina učinka izračunata na temelju velikog uzorka, onda je ona točnija od one koja je izračunata na temelju malog uzorka [COE2000]. Ova *granica pogreške* (eng. margin of error) se može kvantificirati koristeći *intervale pouzdanosti* (eng. confidence interval).

Intervali pouzdanosti daju istu informaciju kao i testovi značajnosti: interval pouzdanosti od 95% je ekvivalentan razini značajnosti od 5%. Da bi izračunali interval pouzdanosti od 95%, pretpostavimo da je dobivena vrijednost veličine učinka istinita. Za svakih 100 novih jednakobrojnih uzoraka, na 95 od njih bi se dobila vrijednost veličine učinka unutar intervala pouzdanosti od 95%. Ako ovaj interval uključuje nulu, onda možemo reći da rezultat nije statistički značajan. Ako ovaj interval ne sadrži nulu, onda kažemo da je statistički značajan na razini 5%.

Formulu za izračunavanje intervala pouzdanosti dali su Hedges i Olkin [HEDG1985]. Ako je veličina učinka Δ , onda je njena standardna devijacija:

$$\sigma_{\Delta} = \sqrt{\frac{N_e + N_c}{N_e N_c} + \frac{\Delta^2}{2(N_e + N_c)}} \quad (3.69)$$

Interval pouzdanosti od 95% za veličinu učinka Δ je:

$$\text{od } \Delta - 1.96 \cdot \sigma_{\Delta} \text{ do } \Delta + 1.96 \cdot \sigma_{\Delta} \quad (3.70)$$

3.3.5. INTERPRETACIJA VELIČINE UČINKA

Prvi način interpretacije je da u izjednačavanju vrijednosti veličine učinka i vrijednosti standardne devijacije za koju je eksperimentalna grupa bolja od kontrolne. Npr. veličina učinka 0.8 znači da je eksperimentalna grupa bolja od kontrolne grupe za 0.8 standardne devijacije.

Pošto je veličina učinka je ekvivalentna z-vrijednosti standardne normalne distribucije [COE2000], može se kazati da veličina učinka 0.8 znači da je uspjeh prosječne osobe u eksperimentalnoj grupi bolji od uspjeha 79% osoba iz kontrolne grupe (Tablica 3.10.), što odgovara površini ispod normalne krivulje za z-vrijednost 0.8. U jeziku percentila to bi značilo da je aritmetička sredina eksperimentalne grupe na 79. percentilu kontrolne grupe, odnosno da se rezultati prosječne osobe iz eksperimentalne grupe nalaze na 79-om percentilu svih rezultata na završnom testu, dok se rezultati prosječne osobe iz kontrolne grupe nalaze na 50-om percentilu.

Treći način interpretiranja veličine učinka je u vjerojatnosti pogađanja iz koje je grupe neka osoba na temelju poznavanja njenog uspjeha (Tablica 3. 10.). Ako je veličina učinka 0 (to znači da su eksperimentalna i kontrolna grupa jednake), onda je vjerojatnost pogađanja 0.50. Ako je razlika između eksperimentalne i kontrolne grupe 0.3, vjerojatnost pogađanja se povećava na 0.56. kod veličine učinka 1, vjerojatnost je 0.69, što predstavlja dvije trećine šanse za točnim pogotkom.

Četvrti način interpretacije veličine učinka su predložili McGraw i Wong (prema [BACK2000]) koristeći vjerojatnost da će slučajno odabrani rezultat iz eksperimentalne grupe biti veći od slučajno odabranog rezultata iz kontrolne grupe. Na primjer, ako se razlike u visinama između mladića i djevojaka razlikuju za veličinu učinka 2, to je, prema Tablica 3. 10., 0.92, tj. u 92 od 100 parova mladića i djevojaka mladić će biti viši od djevojke.

Tablica 3. 10. Interpretacija veličine učinka (promijenjeno prema [COE2000] i [BACK2000])

Veličina učinka d	Postotak osoba iz kontrolne grupe čiji je uspjeh ispod uspjeha prosječne osobe iz eksperimentalne grupe (percentili)	Vjerojatnost pogadanja iz koje je grupe neka osoba na temelju poznavanja njenog uspjeha	Postotak nepreklapanja distribucija grupa	Vjerojatnost da će osoba iz eksperimentalne grupe imati veći rezultat od osobe iz kontrolne grupe (ako su obje slučajno odabrane)
0.0	50%	0.50	0%	0.50
0.1	54%	0.52	7.7%	0.53
0.2	58%	0.54	14.7%	0.56
0.3	62%	0.56	21.3%	0.58
0.4	66%	0.58	27.4%	0.61
0.5	69%	0.60	33.0%	0.64
0.6	73%	0.62	38.2%	0.66
0.7	76%	0.64	43.0%	0.69
0.8	79%	0.66	47.4%	0.71
0.9	82%	0.67	51.6%	0.74
1.0	84%	0.69	55.4%	0.76
1.2	88%	0.73	62.2%	0.80
1.4	92%	0.76	68.1%	0.84
1.6	95%	0.79	73.1%	0.87
1.8	96%	0.82	77.4%	0.90
2.0	98%	0.84	81.1%	0.92

Peti način interpretacije veličine učinka je u terminima postotka nepreklapanja rezultata eksperimentalne grupe s rezultatima kontrolne grupe [BACK2000]. Veličina učinka 0 znači da se distribucija rezultata eksperimentalne grupe potpuno preklapa s distribucijom rezultata kontrolne grupe, tj. ima 0% nepreklapanja. Veličina učinka 0.8 označava nepreklapanje od 47.4% dviju distribucija.

Šesti način interpretiranja veličine učinka je uspoređivanjem s veličinama učinaka koje su nam poznate. Cohen [COHE1969] smatra da su vrijednosti veličine učinka oko 0.2 male što ilustrira primjerom da razlike između visina djevojčica od 15 i 16 godina odgovaraju ovoj veličini učinka. Veličina učinka od 0.5 se smatra umjerenom i dovoljno velikom da se može vidjeti golim okom. Ona odgovara razlici između visina djevojčica od 14 i 18 godina. Veličina učinka od 0.8 se smatra velikom i odgovara razlici između visina djevojčica od 13 i 18 godina.

Glass i ostali (prema [COE2000]) smatraju da se učinak nekog eksperimentalnog faktora može interpretirati samo u odnosu na ostale eksperimentalne faktore kojima se želi postići isti učinak. Oni smatraju da veličina učinka 1 odgovara razlici jedne godine školovanja u postignućima na testu učenika osnovne škole.

Interpretacija veličine učinka može se olakšati ako se promotre primjeri provedenih istraživanja. U Tablica 3. 11. navedeni su primjeri samo za ilustraciju uporabe veličine učinka.

Tablica 3. 11. Neke veličine učinka dobivene kroz istraživanja [COE2000]

Eksperimentalni faktor	Rezultat	Veličina učinka	Istraživanje
Smanjenje veličine razreda s 23 na 15 učenika	Rezultati testa iz čitanja	0.30	Finn and Achilles (1990) 'Some questions and answers about class size' AERJ?
	Rezultati testa iz matematike	0.32	
Mali razredi (<30) vs veliki razredi	Stavovi učenika	0.47	Smith, M.L. and Glass, G.V. (1980) 'Meta-analysis of research on class size and its relationship to attitudes and instruction.' American Educational Research Journal, 17, 419-433.
	Stavovi učitelja	1.03	
Otvoreni (djetetu orijentirani) razredi vs tradicionalni razredi	Postignuća učenika	-0.06	Giaconia, R.M. and Hedges, L.V. (1982) 'Identifying features of effective open education.' Review of Educational Research, 52, 579-602.
	Odnos učenika prema školi	0.17	
Vježbanje	Bodovi na testu	0.32	Kulik, J.A., Kulik, C.C. and Bangert, R.L. (1984) 'Effects of practice on aptitude and achievement test scores.' American Education Research Journal, 21, 435-447.
Povratna informacija o performansama učenika	Postignuća učenika	0.70	Fuchs, L.S. and Fuchs, D. (1986) 'Effects of systematic formative evaluation: a meta-analysis.' Exceptional Children, 53, 199-208.
Poučavanje od strane vršnjaka	Postignuća učenika	0.40	Cohen, P.A., Kulik, J.A. and Kulik, C.C. (1982) 'Educational outcomes of tutoring: a meta-analysis of findings.' American Educational Research Journal, 19, 237-248.
	Postignuća poučavatelja	0.33	
Individualizirano poučavanje	Postignuća učenika	0.10	Bangert, R.L., Kulik, J.A. and Kulik, C.C. (1983) 'Individualised systems of instruction in secondary schools.' Review of Educational Research, 53, 143-158.
Computer assisted instruction (CAI)	Postignuća učenika (sva istraživanja)	0.24	Fletcher-Flinn, C.M., & Gravatt, B. (1995). The efficacy of computer- assisted instruction (CAI): A Meta-analysis. Journal of Educational Computing Research, 12(3), 219-242.
	Postignuća učenika (dobro kontrolirana istraživanja)	0.02	

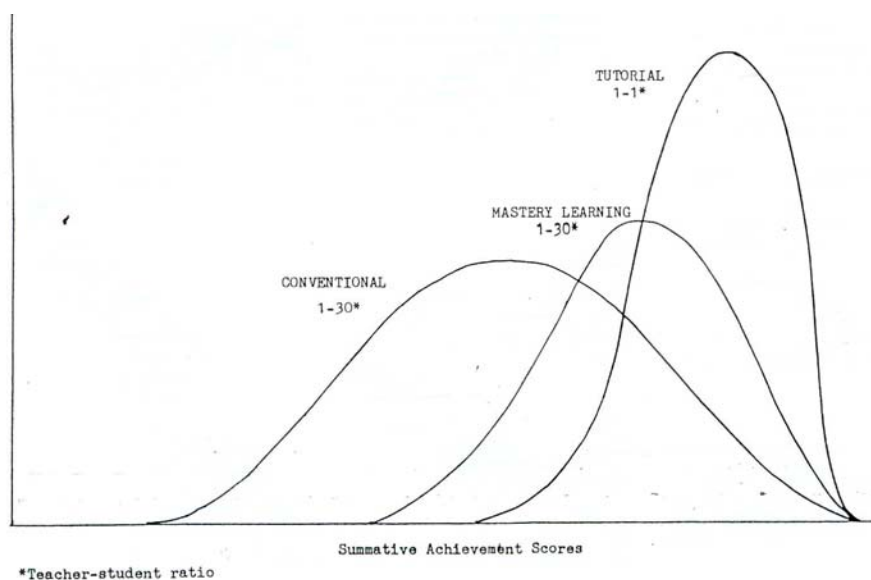
3.3.6. 2-SIGMA PROBLEM

Jedno od najpoznatijih istraživanja u području vrednovanja učinka procesa učenja i poučavanja je definitivno Bloom-ovo istraživanje o razlikama između tradicionalnog učenja (eng. conventional learning), učenja s provjeravanjem (eng. mastery learning) i tutorskog učenja (eng. tutoring learning) [BLOO1984]:

- Kod tradicionalnog učenja su se nastavni sadržaji za zadano područno znanje savladavali u grupi od 30 studenata pod vodstvom jednog nastavnika. Testovi za provjeru znanja obavljali su se periodički radi ocjenjivanja postignuća studenata.
- Kod učenja s provjeravanjem su se nastavni sadržaji za zadano područno znanje savladavali u grupi od 30 studenata pod vodstvom jednog nastavnika obično istim onim kao i kod tradicionalnog učenja. Međutim, bitna se razlika očituje u tome da su se testovi koristili radi stjecanja povratne informacije za nastavnika radi popravljivanja tijekom izlaganja nastavnih sadržaja, a osim toga isti testovi su služili za ocjenjivanje postignuća kod studenata.
- Kod tutorskog učenja studenti savladavaju nove nastavne sadržaje uz pomoć osobnog nastavnika (tutora) i to na način da na svakog studenta ide jedan tutor (ili jedan tutor ide na

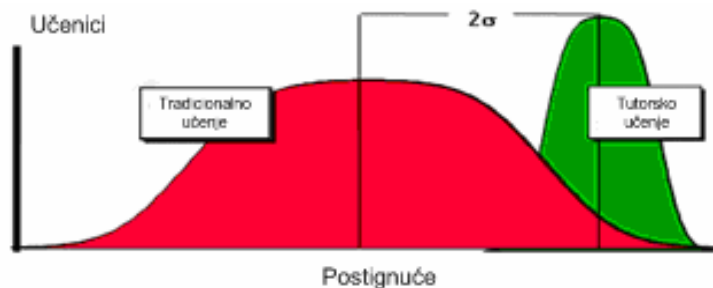
dva ili tri studenta). Ova vrsta nastave popraćena je povremenim testovima, povratno-korektivnim procedurama i paralelnim testovima kao kod učenja s provjeravanjem. Važno je istaknuti da se pokazalo kako je potreba za korektivnim procedurama u tutorskom sustavu veoma mala.

Studenti su bili po slučajnom uzorku raspoređeni u neku od ove tri grupe, ali su imali slična prethodna postignuća i sličnu razinu znanja koja je vezana za danu materiju. Vrijeme koje je utrošeno za nastavu jednako je u svim trima grupama, osim za korektivni postupak u grupi učenja s provjeravanjem i grupi s tutorskim učenjem. Bloom je u svom istraživanju ustvrdio prosječan student u tutorskoj grupi za oko dvije standardne devijacije bolji od prosječnog studenta u tradicionalnoj grupi. Prosječan student u grupi učenja s provjeravanjem je za oko jednu standardnu devijaciju bolji od prosječnog studenta u tradicionalnoj grupi (Slika 3. 15.).



Slika 3. 15. Distribucija postignuća učenika kod tradicionalnog učenja, učenja s provjeravanjem i tutorskog učenja [BLOO1984]

Dakle, učenici koji su bili individualno poučavani od strane učitelja, pokazuju veličinu učinka 2.0 kada se uspoređuju s učenicima koji su bili tradicionalno poučavani (Slika 3. 16.). Dobivenu razliku od dvije standardne devijacije Bloom je nazvao 2-sigma razlika, jer se obično standardna devijacija označava grčkim slovom sigma. Drugim riječima, rezultati učenika koji su bili individualno poučavani su u prosjeku bili veći za dvije standardne devijacije, tj. individualna nastava pomiče učenika s 50. percentila na 98. percentil na ljestvici rezultata standardiziranog testa.



Slika 3. 16. 2-sigma razlika

Bloom je smatrao da je ova razlika od dvije standardne devijacije ideal prema kojem proces učenja i poučavanja treba stremiti, te je postizanje ove razlike u procesu učenja i poučavanja nazvao 2-sigma problem. Stoga su se mnogi istraživači uhvatili ukoštac s ovim problemom ne bi li pronašli način kako iskoristiti informacijsku i komunikacijsku tehnologiju u svrhu postizanja učinka individualizirane nastave pod praktičnijim i realnijim uvjetima nego što je jedan na jedan učenje, koje je preskupo da bi se primijenilo na veći broj studenata. U Tablici 3. 12. prikazani su faktori koji utječu na proces učenja i poučavanja te pripadne im veličine učinka koje su dobivene kroz istraživanja.

Tablica 3. 12. Učinak nekih faktora na poučavanje [BLOO1984]

Veličina učinka	Faktori
2.00	Tutorsko poučavanje (eng. tutorial instruction)
1.20	Pojačanje (eng. reinforcement)
1.00	Korektivna povratna veza (eng. corrective feedback)
1.00	Znakovi i objašnjenja (eng. cues and explanations)
1.00	Sudjelovanje učenika u učionici (eng. student classroom participation)
1.00	Učeničko vrijeme za rješavanje zadataka (eng. student time-on-task)
1.00	Poboljšane vještine čitanja i učenja (eng. improved reading/study skills)
0.80	Suradno učenje (eng. cooperative learning)
0.80	Domaća zadaća ocijenjena (eng. homework graded)
0.60	Moral razreda (eng. classroom morale)
0.60	Provjera početnih spoznajnih preduvjeta (eng. verifying initial cognitive prerequisites)
0.50	Posredovanje u kućnom okruženju (eng. home environment intervention)
0.40	Pomoć od strane drugih učenika (eng. peer and cross-age remedial tutoring)
0.30	Domaća zadaća (eng. homework assigned)
0.30	Pitanja višeg reda (eng. higher order questions)
0.30	Novi nastavni program za znanost i matematiku (eng. new science and math curricula)
0.30	Očekivanje učitelja (eng. teacher expectancy)
0.25	Socijalno-ekonomski status (eng. socioeconomic status)
0.20	Utjecaj grupe vršnjaka (eng. peer group influence)
0.20	Napredno organiziranje (eng. advance organizers)

3.3.7. META-ANALIZA

Jedna od glavnih prednosti uporabe veličine učinka je u tome što se mogu koristiti različite veličine učinka dobivene kroz ponavljanja istog eksperimenta u svrhu dobivanja najbolje procjene veličine učinka. Ovaj proces uspoređivanja kvalitativnih rezultata eksperimenata i

izračunavanje jedne vrijednosti veličine učinka, naziva se *meta-analiza*. Razvio ju je statističar Glass [GLAS1981] koji ju smatra analizom analiza i danas se koristi u svim granama znanosti.

Meta-analize mogu puno više od pukog izračunavanja prosječne veličine učinka. Ako neka istraživanja o određenom eksperimentalnom faktoru imaju velike veličine učinka, a neka istraživanja male, bilo bi loše samo ih kombinirati i kazati da je prosječna veličina učinka osrednja. Mnogo je korisnije proučiti originalna istraživanja da bi se uočile razlike između onih s velikom veličinom učinka i onih s malom veličinom učinka, te pokušati uočiti koji parazitarni faktori su utjecali na tu razliku. Najbolje meta-analize, stoga uključuju traženje uzročno-posljedičnih veza između veličine učinka i karakteristike eksperimentalnog faktora, konteksta i organizacije istraživanja [COE2000]. Najveća opasnost kod izrade meta-analiza leži u kombiniranju neusporedivih rezultata, što može dovesti do potpuno besmislenih zaključaka.

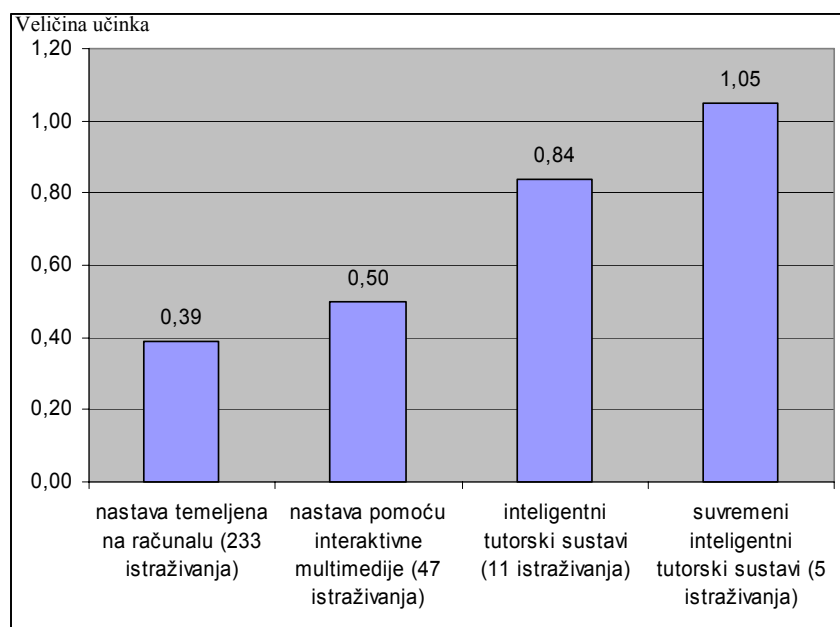
Pravi izazov s kojim se susreću istraživači kod izrade meta-analiza je pronalaženje i izračunavanje veličina učinka koje nisu uvijek eksplicitno naglašene u istraživanjima. Rezultati eksperimenata su obično prezentirani na način da je kasnije meta-analitičarima teško pronaći veličine učinka. Stoga se od njih zahtjeva poznavanje osnova statističke analize.

3.3.7.1 META-ANALIZE O UČINKU TEHNOLOGIJOM POTPOMOGNUTOG UČENJA I POUČAVANJA

Fletcher je u [FLET2003] prikupio i sistematizirao veći broj istraživanja koja se bave postignućima učenika u učenju i poučavanju uz pomoć informacijske i komunikacijske tehnologije. Tehnologijom potpomognuto učenje i poučavanje (eng. technology-assisted instruction) obuhvaća, prema [FLET2003], nastavu temeljenu na računalima (eng. computer-based instruction CBI), nastavu pomoću interaktivne multimedije (eng. interactive multimedia), inteligentne tutorske sustave (eng. intelligent tutoring systems), mrežne simulacije tutoriala (eng. networked tutorial simulation) te nastavu temeljenu na Internetu (eng. web-based instruction). Fletcher je postavio pitanje da li je tehnologijom poduprto učenje i poučavanje poboljšanje, odnosno, da li omogućava stvaranje okruženje procesa učenja i poučavanja koje je imalo bolje od postojećeg.

Vinsonhaler i Bass su 1972. godine (prema [FLET2003]) proučili 34 objavljena vrednovanja procesa učenja i poučavanja zasnovanog na računalima koja su obuhvatila više od 10 000 učenika. Ustanovili su da postoje pozitivni ili jednaki rezultati u 30 od 34 istraživanja i da je medijan razlika postignuća između CBI učenika 40%. Kulik je (prema [FLET2003]) proveo mnogo istraživanja vezana za učenje i poučavanje zasnovano na tehnologiji. Proveo je meta-analizu kojom je ustvrdio veličinu učinka 0.35 takvog pristupa nastavi.

Fletcher je sažeo rezultate svojeg proučavanja meta-analiza u jedan grafički prikaz (Slika 3. 17.). Proučio je 233 istraživanja o učinku nastave temeljene na računalima koja je uključivala računalne prezentacije s tekstom, slikama, jednostavnim animacijama, kao i malu razinu individualizacije. Pokazalo se da nastava temeljena na računalima ima veličinu učinka 0.39. Nastava pomoću interaktivne multimedije uključuje zvuk, složenije animacije i video, te je pokazala veličinu učinka 0.50. Inteligentni tutorski sustavi nastoje oponašati učenje i poučavanje jedan-na-jedan i pokazuju veličinu učinka od 0.84. Suvremeni inteligentni tutorski sustavi pokazuju obećavajućih 1.05.



Slika 3. 17. Veličine učinka tehnologijom potpomognutog učenja i poučavanja [FLET2003]

3.4. PRIMJERI PROVEDENIH ISTRAŽIVANJA O VREDNOVANJU UČINKA SUSTAVA E-UČENJA

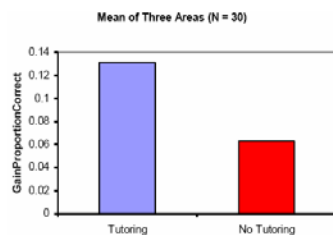
U ovom poglavlju prikazani su primjeri istraživanja o učinku sustava e-učenja. Iz same strukture tih istraživanja vidi se da oni ne koriste potpuno istu metodologiju vrednovanja, ali se u većini slučajeva radi o eksperimentalnom istraživanju s kontrolnom grupom. Navedeni odabir vrednovanih sustava je napravljen prema dostupnosti rezultata o provedenom vrednovanju učinka te prema njihovoj orijentaciji ka inteligentnim tutorskim sustavima.

3.4.1. SPOKEN CONVERSATIONAL TUTOR - SCoT

Spoken Conversational Tutor SCoT [PONB2004] je sustav koji omogućuje studentima diskutiranje prirodnim jezikom uporabom sučelja za govor. Verzija sustava SCoT-DC je primijenjena na područno znanje o kontroliranju šteta na brodovima.

Trideset studenata s engleskog govornog područja sudjelovalo su u eksperimentu koji je proveden u svrhu testiranja hipoteze da SCoT-DC pomaže studentima naučiti kontrolu štete. Studenti su slučajnim odabirom podijeljeni u tri grupe. Sve grupe su prošle iste sesije učenja na sustavu, samo različitim redosljedom. Učenje je mjereno na dva načina. Prvo, opće znanje je ispitano inicijalnim i završnim testom s pitanjima višestrukog izbora. Drugo, kvantitativne mjere performansi su dobivene na temelju scenarija učenja na sustavu.

Svi studenti su prošli kroz interaktivni multimedijalni uvod (30-40 minuta) da bi se upoznali sa sustavom i osnovama kontrole štete te da bi izvježbali uporabu sučelja za prepoznavanje govora. Nakon uvoda, studenti su pisali inicijalni test s 20 pitanja višestrukog izbora, a zatim su pristupili 10-minutnom vježbanju na sustavu. Nakon toga su uslijedile četiri sesije učenja na sustavu (svaka po 15 minuta). Na kraju su studenti pisali završni test s 20 pitanja višestrukog izbora. Studenti su pokazali napredovanje i kod pisanih testova i kod rada na sustavu. Svaki student je postigao više bodova na završnom testu nego na inicijalnom testu, a aritmetička sredina završnog testa (84%) je značajno veća (t-test: $p = 0.000$) nego aritmetička sredina inicijalnog testa (67%) (Slika 3.18.).



Slika 3. 18. Grafički prikaz rezultata eksperimenta [PONB2004]

3.4.2. AutoTutor

AutoTutor [JACK2004] je računalni program za poučavanje studenata računalnoj pismenosti ili konceptualnoj fizici simulirajući strategije poučavanja živog učitelja.

Studenti koji su sudjelovali u eksperimentu su bili studenti koji su upisali uvodne kolegije iz računalne pismenosti (70 studenata) i iz fizike (24 studenta). Za sudjelovanje u eksperimentu bili su nagrađeni dodatnim bodovima. Studenti su pisali inicijalni test, radili na sustavu, te zatim pisali završni test. Inicijalni i završni test su se sastojali od pitanja višestrukog izbora. Razmjeri točnih odgovora na pitanja inicijalnog i završnog testa su korišteni kao mjera studentova znanja. Izračunata je normalizirana razlika bodova kao mjera rezultata studenta po formuli:

$$\frac{\text{završni test} - \text{inicijalni test}}{\sigma_{\text{inicijalni}}} \quad (3.71)$$

Izračunata je još jedna mjera razlika u znanju, nazvana Estes bodovi, po formuli:

$$\frac{\text{završni test} - \text{inicijalni test}}{1 - \text{inicijalni test}} \quad (3.72)$$

Mjere studentovih rezultata se prikazuju na tri načina: rezultati inicijalnog testa, normalizirana razlika, Estes bodovi. Izračunate su korelacije između mjera studentovih rezultata i razmjera dijaloških akcija (Tablica 3. 13.). Studenti koji su dobili više „pumps“ i „hints“ (student kontrolira izgradnju znanja) naučili su više od onih koji su dobili više „prompts“ i „assertions“ (sustav kontrolira izgradnju znanja).

Tablica 3. 13. Korelacije razmjera studentova znanja i dijaloških akcija

	Pump	Hint	Prompt	Assertion
Računalna pismenost				
završni test	0.188	0.466**	-0.165	-0.433**
normalizirana razlika	0.151	0.200	-0.120	-0.198
Estes bodovi	0.131	0.141	-0.117	-0.136
Fizika				
inicijalni test	0.527*	0.235	-0.166	-0.418*
normalizirana razlika	0.043	0.103	-0.096	-0.089
Estes bodovi	0.026	0.373*	-0.322	-0.290

* $p < .10$, ** $p < .05$

3.4.3. EQUATION TUTOR - E-tutor

Equation tutor (E-tutor) [RAZZ2004] je inteligentni tutorski sustav za rješavanje jednadžbi koji kombinira spoznajni model područnog znanja s modelom poučavanja temeljenog na dijalogu.

Ekperiment je proveden tijekom ljetnih praznika s 15 učenika srednje škole koji su slušali predmet Algebra I na pripremama za upis na fakultet. Osam učenika je koristili eksperimentalni E-tutor, a sedam učenika je koristilo kontrolni Cognitive Tutor (Koedinger prema [RAZZ2004]).

Ekperiment je trajao dva dana po 1.5 sati. Učenici su prvo pisali 20-minutni inicijalni test. Zatim je uslijedila 5-minutna demonstracija oba sustava svim učenicima. Nadalje, učenici su pristupili rješavanju što većeg broja zadataka u prvom dijelu eksperimenta. U drugom dijelu eksperimenta, učenici su pristupili rješavanju preostalih zadataka u sustavima. U eksperimentu se nije ograničavalo vrijeme, već broj zadataka koje je učenik mogao riješiti. Na kraju eksperimenta, učenici su pisali 20-minutni završni test.

Rezultati četiri učenika su isključena iz analize rezultata. Tri učenika koji su imali savršen rezultat na inicijalnom testu se nisu uzela u obzir jer se smatralo da su oni već usvojili područno gradivo. Također se nije uzelo u obzir i jednog učenika koji je preskočio četiri zadatka kod inicijalnog testa jer se smatralo da se to dogodilo slučajno. Analiza rezultata je provedena nad vrlo malim uzorkom od 6 učenika u eksperimentalnoj grupi i 5 učenika u kontrolnoj grupi. Zbog malog uzorka nije se mogla postići statistička značajnost. Ipak se pristupilo izračunavanju veličine učinka i zaključeno je da su učenici koji su koristili E-tutor postigli bolje rezultate od učenika koji su koristili kontrolni sustav za 0.4 standardne devijacije.

3.4.4. REUSABLE EDUCATIONAL DESIGN ENVIRONMENT AND ENGINEERING METHODOLOGY - REDEEM

REDEEM (Reusable Educational Design Environment and Engineering Methodology) [AINS2002] je autorsko okruženje koje omogućava učitelju izradu jednostavnih inteligentnih tutorskih sustava definirajući kako se koja grupa učenika treba najbolje poučavati.

U eksperimentu je sudjelovalo 86 učenika srednje škole između 14 i 15 godina. Oni su podijeljeni u 5 nejednako brojnih kategorija: 8 (A), 30 (B), 23 (C), 15 (D), 10 (E).

Napravljeno je 60 pitanja višestrukog izbora od kojih je 30 bilo iz predmeta Genetika I, a 30 iz predmeta Genetika II. Te dvije grupe od po 30 pitanja su dalje podijeljene na 10 pitanja za REDEEM, 10 pitanja za površinske transformacije i 10 pitanja koja nisu za REDEEM. Postojale su dvije verzije testova tako da je pola učenika odgovarala prvo na pitanja iz Genetike I, a zatim na pitanja iz Genetike II, a druga polovina učenika obrnuto. Inicijalni i završni test su bili jednaki i nosili su maksimalno 30 bodova. Svaki učenik je prisustvovao na minimalno dvije, a maksimalno pet sesija učenja o genetici, koje su trajale od 30 do 90 minuta. Rezultati 12 učenika nisu se uzeli u obzir kod analize jer nisu pisali inicijalni ili završni test. Analiza rezultata testova je pokazala da se znanje učenika sveukupno poboljšalo, ali stupanj poboljšanja nije uzrokovan vrstom okruženja učenja ili pripadnošću grupi. Iako su učenici poboljšali svoje znanje genetike,

REDEEM nije imao statistički značajan učinak na rezultate učenja. Povećanje razlike u bodovima između inicijalnog i završnog testa je bilo isto neovisno da li su učili pomoću CBT-a ili REDEEM-a. REDEEM poboljšava učenje za 0.2 sigma u usporedbi s CBT.

Tablica 3. 14. Rezultati inicijalnog i završnog testiranja

	REDEEM				CBT			
	Genetika I (n = 40)		Genetika II (n = 34)		Genetika I (n = 34)		Genetika II (n = 40)	
	\bar{X}	σ	\bar{X}	σ	\bar{X}	σ	\bar{X}	σ
Inicijalni test	11.00	3.37	12.68	3.94	11.71	3.01	11.83	3.69
Završni test	14.38	4.83	15.47	4.63	14.18	3.75	14.30	3.69

3.4.5. LEARNING BY ASKING - LBA

Learning by asking (LBA) [ZHAN2006] je sustav e-učenja temeljen na multimediji koji integrira multimedijalni nastavni materijal uključujući i interaktivni video, prezentacije i bilješke. LBA sustav pruža visok nivo interakcije omogućavajući studentima direktan pristup dijelovima videa.

Provedeno je istraživanje kojim se htjelo ispitati da li interaktivni video povećava studentovo razumijevanje i povećava učinak učenja. Zavisna varijabla je bilo učenje mjereno rezultatima testova. Svaki od 138 studenata, koji su sudjelovali u eksperimentu, je slučajnim odabirom svrstan u jednu od četiri grupe kojima su slučajnim odabirom pridružena okruženja učenja (Tablica 3. 15.).

Tablica 3. 15. Grupe u eksperimentu

Grupe	Okruženje učenja	Veličina grupe
1	LBA sustav + interaktivni video	35
2	LBA sustav + neinteraktivni video	35
3	LBA sustav bez videa (samo prezentacije i bilješke)	34
4	Tradicionalna nastava	34

Studenti u grupama 1, 2 i 3 imali su isti scenarij:

1. Uvod: objašnjeni su ciljevi i procedure eksperimenta
 2. Inicijalni test: pitanja višestrukog izbora i da-ne pitanja
 3. Učenici koji su trebali koristiti LBA sustav prisustvovali su 5-minutnoj demonstraciji sustava, te su 5 minuta sami radili na sustavu.
 4. Učenici su pristupili učenju na sustavu u trajanju od 50 minuta, od kojih je na video otpalo 29 minuta. Pogledali su i prezentaciju od 20 prikaza. Učenici tradicionalne nastave su slušali predavanja u trajanju 50 minuta.
 5. Završni test
- Testovi su bodovani na ljestvici 0-50 bodova.

U Tablica 3. 16. te u Tablica 3. 17. su prikazani statistički obrađeni rezultati eksperimenta.

Tablica 3. 16. Analiza rezultata po grupama (razlika bodova između završnog i inicijalnog testa)

Grupe	Aritmetičke sredine	Standardne devijacije
LBA sustav + interaktivni video (1)	34.1	8.87
LBA sustav + neinteraktivni video (2)	27.7	8.85
LBA sustav bez videa (3)	26.7	10.02
Tradicionalna nastava (4)	23.7	8.79

Tablica 3. 17. Analiza rezultata između grupa (razlika bodova između završnog i inicijalnog testa)

Grupe	2	3	4
1	6.49 (0.005)**	7.41 (0.001)**	10.47 (0.00)**
2		0.92 (0.967)	3.98 (0.184)
3			3.06 (0.417)

** Razlika aritmetičkih sredina je značajna na razini 0.01

Nije bilo značajne statističke razlike u razlici bodova između završnog i inicijalnog testa između Grupe 2 i Grupe 3, što implicira da interaktivni video pomaže boljem razumijevanju.

3.4.6. CONCEPTUAL HELPER

Conceptual Helper [ALBA2000] je inteligentni tutorski sustav za učenje i poučavanje fizike pomoću domaćih zadataka orijentiranih na rješavanje kvalitativnih problema koji zahtijevaju primjenu konceptualnog znanja. Ovaj sustav temelji svoj model učenika na modelu traganja uz primjenu probabilističkih procjena za ispravljanje grešaka.

Četrdeset i dva studenta koji su upisali kolegij Uvod u mehaniku su slučajnim odabirom raspodijeljeni u kontrolnu i eksperimentalnu grupu. Obje grupe su pisale inicijalni test na papiru koji se sastojao od 29 kvalitativnih problema. Zatim su studenti rješavali neke probleme u sustavu Andes (u trajanju od dva školska sata) i primali odgovarajuću povratnu informaciju ovisno o grupi kojoj su pripadali. Studentima kontrolne grupe su se njihovi ulazni podaci bojali crveno ili zeleno ovisno o točnosti. U slučaju netočnih ulaznih podataka, student je mogao tražiti pomoć, u obliku kratkih uputa, birajući opciju iz izbornika. Ako student zatraži još pomoći, sustav mu kaže točan odgovor. Studentima eksperimentalne grupe su se ulazni podaci također bojali crveni ili zeleno, kao i kod kontrolne grupe, ali u slučaju netočnog ulaznog podatka, Conceptual Helper im je sam nudio pomoć. Na kraju eksperimenta, svi su studenti pisali završni test.

Prvo je trebalo provjeriti da li su početne sposobnosti studenata u grupama kao cjelini ekvivalentne. Aritmetička sredina inicijalnog testa je 33.7 sa standardnom devijacijom 7.47. Aritmetička sredina završnog testa je 31.36 sa standardnom devijacijom 8.14. Između kontrolne i eksperimentalne grupe nije pronađena statistički značajna razlika ($t(40)=0.965$, $p=0.34$). Zatim se pristupilo usporedbi razlika rezultata inicijalnog i završnog testa. Aritmetička sredina kontrolne grupe je 4.12 sa standardnom devijacijom 5.33. Aritmetička sredina eksperimentalne grupe je 7.47 sa standardnom devijacijom 5.03. Pronađena je statistički značajna razlika ($t(40)=2.094$, $p=0.043$) koja znači da je način na koji Conceptual Helper pruža pomoć studentu imala pozitivan utjecaj na njegovo razumijevanje. Izračunata je veličina učinka po formuli (3.53) i dobiven je rezultat:

$$\Delta = \frac{7.47 - 4.12}{5.33} = 0.63 \quad (3.73)$$

3.4.7. SQL- Tutor

SQL-Tutor [MITR1999] je inteligentni tutorski sustav koji pomaže studentima formulirati upite u Structured Query Language. SQL-Tutor podrazumijeva da studenti poznaju osnovne koncepte o bazama podataka, relacijski model te osnove SQL jezika.

Eksperiment je proveden sa 46 studenata zadnje godine studija na Sveučilištu u Canterbury u Novom Zelandu. Oni su poslušali 6 predavanja o SQL-u i potrošili su najmanje 8 sati na vježbanje definiranja upita. Eksperimentalnu grupu je činilo 20 studenata koji su radili sa sustavom SQL-Tutor dva sata. Kontrolnu grupu je činilo 26 studenata.

Test znanja je proveden dva tjedna nakon rada sa sustavom. Test se sastojao od 6 problema koje je trebalo riješiti za 90 minuta. Test je bodovan na ljestvici 0-100. aritmetička sredina rezultata eksperimentalne grupe je 82.7, a kontrolne grupe 71.2. razlika u aritmetičkim sredinama je statistički značajna ($t= 2.68$, $p = 0.01$). Zajednička standardna devijacija je 15.4. Izračunata je veličina učinka po formuli (3.60) i dobivena vrijednost je 0.75.

3.4.8. NORMIT

NORMIT [MITR2005] je inteligentni tutorski sustav koji poučava normalizaciju podataka. Normalizacija podataka je proceduralna zadaća: student prolazi kroz niz koraka da bi analizirao kvalitetu baze podataka.

Proveden je eksperiment u svrhu utvrđivanja da li samoobjašnjavanje (eng. self-explanation) ima pozitivan učinak kako na proceduralno, tako i na konceptualno znanje. Prije eksperimenta studenti su slušali četiri predavanja o normalizaciji podataka. Studenti u kontrolnoj grupi su koristili osnovnu verziju sustava, a eksperimentalna grupa je koristila NORMIT-SE, verziju sustava koja podržava samoobjašnjavanje. U eksperimentu je sudjelovalo 49 studenata i mogli su koristiti sustav kad god su htjeli i koliko dugo su htjeli. Na početku eksperimenta studenti su pisali inicijalni test (maksimalni broj bodova 4). Na kraju eksperimenta su pisali završni test (maksimalni broj bodova 29) koji nije potpuno komparabilan inicijalnom pa se rezultati ova dva testa ne mogu uspoređivati. Rezultati eksperimenta su prikazani u Tablica 3. 18.

Tablica 3. 18. Aritmetičke sredine rezultata eksperimenta (u zgradama su standardne devijacije)

	NORMIT	NORMIT-SE
Broj studenata	27	22
Broj pristupa sustavu	2.9 (1.95)	2.4 (1.7)
Vrijeme (minute)	231 (202)	188 (167)
Broj problema	16.7 (11.2)	11.9 (10.4)
Rješeni problemi (%)	81.9 (22.5)	80.4 (16.2)
Inicijalni test (%)	55.6 (26.2)	64.77 (26.3)
Završni test (%)	51.3 (15.4)	53.61 (22.3)

Rezultati inicijalnog testa su pokazali da su grupe ekvivalentne. Nije bilo statistički značajne razlike između grupa po broju pristupa sustavu. Razlika između aritmetičkih sredina broja problema dviju grupa je statistički značajna ($p=0.067$). obje grupe su bile jednako uspješne kod rješavanja problema. Nije bilo statistički značajne razlike između rezultata završnog testa.

3.4.9. SYSTEM DYNAMICS MODEL

Vrednovanje učinka System Dynamics [PFAH2004] modela za učenje i poučavanje o upravljanju projektiranjem programske podrške provodi se pomoću kontroliranog eksperimenta s inicijalnim i završnim testiranjem te kontrolnom i eksperimentalnom grupom.

Provedena su tri jednaka eksperimenta. U svakom eksperimentu studenti su slučajnim odabirom podijeljeni u kontrolnu i eksperimentalnu grupu. U prvom eksperimentu je sudjelovalo 9 studenata (5 u eksperimentalnoj grupi, 4 u kontrolnoj grupi), u drugom 12 (6 u eksperimentalnoj grupi, 6 u kontrolnoj grupi), a u trećem 13 studenata (7 u eksperimentalnoj grupi, 6 u kontrolnoj grupi). Organizacija eksperimenata je prikazana u Tablica 3. 19.

Tablica 3. 19. Organizacija eksperimenta

	Experiment	Replication 1	Replication 2
Uvod	5 min	5 min	5 min
Osobne karakteristike	5 min	5 min	5 min
Inicijalni test	30 min	35 min	35 min
Uvod u metode učenja	5 min	5 min	5 min
Slučajni odabir studenata u grupe	5 min	5 min	5 min
Učenje i poučavanje	45 min	80 min	80 min
Završni test	30 min	35 min	35 min

Kontrolna grupa je svoje učenje i poučavanje realizirala na temelju COCOMO (Constructive Cost Model) modela (Boehm, prema [PFAH2004]), dok je eksperimentalna grupa radila po SD modelu. U eksperimentima su definirane četiri zavisne varijable: Y.1 – interes, Y.2 – znanje, Y.3 – razumijevanje jednostavnog, Y.4 – razumijevanje složenog. Podaci o zavisnim varijablama su sakupljeni pomoću inicijalnih i završnih testiranja. Rezultati eksperimenata s aritmetičkom sredinom, medijanom i standardnom devijacijom za svaku grupu i svaku zavisnu varijablu, su prikazani u Tablica 3. 20.

Tablica 3. 20. Rezultati eksperimenta za svaku grupu i svaku zavisnu varijablu (1)

	Inicijalni test				Završni test				Razlika bodova			
	Y.1	Y.2	Y.3	Y.4	Y.1	Y.2	Y.3	Y.4	Y.1	Y.2	Y.3	Y.4
E: originalni eksperiment												
Grupa A (5)												
Mean	0.69	0.56	0.31	0.37	0.79	0.84	0.66	0.43	0.10	0.28	0.34	0.07
Stdev.	0.18	0.30	0.26	0.25	0.19	0.17	0.13	0.32	0.09	0.36	0.28	0.19
Grupa B (4)												
Mean	0.81	0.50	0.43	0.33	0.79	0.60	0.82	0.46	- 0.03	0.10	0.39	0.13
Stdev.	0.13	0.26	0.31	0.24	0.19	0.16	0.07	0.37	0.09	0.35	0.38	0.34
R1: prvo ponavljanje eksperimenta												
Grupa A (6)												
Mean	0.83	0.57	0.41	0.44	0.85	0.97	0.67	0.44	0.03	0.40	0.26	0.00
Stdev.	0.14	0.23	0.11	0.23	0.15	0.08	0.27	0.09	0.07	0.28	0.25	0.24
Grupa B (6)												
Mean	0.70	0.47	0.33	0.33	0.78	0.43	0.74	0.33	0.08	- 0.03	0.41	0.00
Stdev.	0.18	0.21	0.17	0.30	0.21	0.15	0.17	0.24	0.13	0.32	0.19	0.45
R2: drugo ponavljanje eksperimenta												
Grupa A (7)												
Mean	0.84	0.63	0.41	0.62	0.87	0.83	0.61	0.52	0.03	0.20	0.20	- 0.10
Stdev.	0.09	0.15	0.27	0.30	0.10	0.23	0.24	0.20	0.12	0.22	0.35	0.19
Grupa B (6)												
Mean	0.82	0.60	0.43	0.44	0.91	0.60	0.55	0.53	0.09	0.00	0.12	0.09
Stdev.	0.18	0.33	0.13	0.30	0.11	0.42	0.11	0.19	0.11	0.18	0.19	0.21

Izračunate veličine učinka, stupnjevi slobode, t-vrijednost i p-vrijednost za eksperimentalnu grupu su prikazani u Tablica 3. 21.

Tablica 3. 21. Rezultati eksperimenta za svaku grupu i svaku zavisnu varijablu (2)

Varijabla / Eksperiment	Δ	df	t-vrijednost	p-vrijednost
Varijabla Y.1				
E	1.07	4	2.39	0.04
R1	0.36	5	0.89	0.21
R2	0.23	6	0.62	0.28
Varijabla Y.2				
E	0.77	4	1.72	0.08
R1	1.41	5	3.46	0.01
R2	0.91	6	2.24	0.04
Varijabla Y.3				
E	1.23	4	2.75	0.03
R1	1.06	5	2.61	0.02
R2	0.59	6	1.55	0.09
Varijabla Y.4				
E	0.35	4	0.78	0.24
R1	0.00	5	0.00	0.50
R2	20.50	6	21.33	0.88

U prethodnom poglavlju opisani su statistički instrumenti potrebni za analizu podataka, postupci i instrumenti prikupljanja podataka, metode vrednovanja sustava e-učenja, te eksperiment kao najčešće upotrebljavana metoda vrednovanja. Objasnjeno je što je to veličina učinka i kako se može izračunati te njena primjena kod meta-analiza. Na kraju je ukratko objašnjeno nekoliko primjera vrednovanja sustava e-učenja.

S obzirom na rastuće interese za razvoj sustava koji podupiru proces e-učenja i svekoliku prisutnost e-učenja u svijetu obrazovanja, potrebno je osigurati valjano i učinkovito okruženje u kojem bi se e-učenje odvijalo. Danas se skoro na svakom fakultetu na kojem postoji odjel za računarstvo, pristupa oblikovanju i implementaciji sustava e-učenja. Iako se ti sustavi razvijaju u svrhu potpore nastavi, mali je broj vrednovanih sustava. Naime, vrednovanje sustava, a pogotovo vrednovanje učinka sustava, zahtijeva dugotrajan angažman kako nastavnika koji te sustave koriste u nastavi, tako i studenata koji prisustvuju toj nastavi.

Upravo zbog činjenice da sustavi e-učenja zbog svoje prisutnosti u procesu učenja i poučavanja, utječu na taj isti proces, kao i na postignuća učenika, smatramo da se u nastavi mogu koristiti samo oni sustavi čija je učinkovitost provjerena. Pošto je proces vrednovanja učinka dugotrajan, a izbor metoda za vrednovanje velik, odlučili smo pristupiti definiranju vlastite metodologije za vrednovanje učinka sustava e-učenja.

4. METODOLOGIJA ZA VREDNOVANJE UČINKA PROCESA UČENJA I POUČAVANJA U SUSTAVIMA E-UČENJA

U ovom poglavlju se, na temelju spoznaja iz literature, pristupa definiranju vlastite metodologije za vrednovanje učinka procesa učenja i poučavanja u sustavima e-učenja. U tu svrhu potrebno je odrediti vrstu veličine učinka koja će se računati, zatim definirati strukturu eksperimenta, postaviti nul-hipoteze, pokazati kako se provodi ispitivanje statističke značajnosti, te kako analizirati i interpretirati rezultate. Na kraju će se pokazati kako se definiranom metodologijom pristupilo vrednovanju učinka konkretnog sustava koji pripada klasi inteligentnih tutorskih sustava, a samim time i sustavima e-učenja.

Naša metodologija predstavlja sintetički pristup vrednovanju učinka kojim se povezuju eksperiment kao najčešće korištena metoda vrednovanja i ideja o provjeravanju učinka sustava u što većem broju međustanja. Ostale metodologije vrednovanja učinka sustava e-učenja promatraju učinak sustava samo na kraju eksperimenta, čime gube uvid u utjecaj sustava na postignuće učenika tijekom samog procesa učenja i poučavanja. Strog pristup definiranju metodologije za vrednovanje je potreban jer smatramo da će uniformnost pristupa eliminirati sve vanjske utjecaje koji bi mogli dovesti do krivih pokazatelja o učinku sustava.

4.1. PROCES VREDNOVANJA

Eksperimentalno istraživanje je prikladno za vrednovanje sustava e-učenja jer omogućava ispitivanje odnosa između poučavanja i učenikovih rezultata te dobivanje kvantitativnih mjera značajnosti tih odnosa. Ono omogućava otkrivanje veze između akcija koje se poduzimaju u poučavanju i odgovarajućih rezultata koje učenici postižu u odnosu na te akcije, a također se može odrediti i koliko su takve veze značajne, i stoga je posebno pogodan za ispitivanje učinka procesa učenja i poučavanja.

Eksperimentom utvrđujemo u kojoj mjeri zavisi druga varijabla (posljedica) od prve (utjecaj). U našem pristupu zavisna varijabla je znanje učenika, a nezavisna varijabla je sustav e-učenja čiji učinak želimo ispitati. Osnovni cilj eksperimenta jest utvrditi da li postoji, koliki je i kakav je utjecaj određene nezavisne varijable na određenu zavisnu varijablu.

4.1.1. OBLIKOVANJE EKSPERIMENTA

U svrhu našeg istraživanja definirali smo jednu varijantu eksperimentalnog istraživanja s paralelnim grupama [MUŽI1977]. Kod eksperimenta s paralelnim grupama postoje dvije grupe ispitanika, od kojih je svaka nosilac svog eksperimentalnog faktora - kontrolna grupa se poučava na tradicionalan način, a eksperimentalna grupa se poučava uz pomoć sustava e-učenja.

U točki 3.1.7.3. je objašnjen osnovni model eksperimenta s paralelnim grupama kod kojeg se pomoću inicijalnog i završnog testa mjeri inicijalno i završno znanje ispitanika. Kod modela eksperimentalnog istraživanja s paralelnim grupama kojeg ovdje opisujemo, osim inicijalnog i završnog stanja, ispituje se znanje ispitanika i u međustanjima uporabom testova provjere. Stoga smo ovakav model nazivali *eksperiment s ekvivalentnim paralelnim grupama uz parcijalno ispitivanje stanja*.

4.1.1.1 IZRADA I PROVEDBA INICIJALNOG TESTA

Inicijalni test se koristi zbog toga što u eksperimentu sudjeluju ispitanici s različitim vještinama i predznanjima. Moramo utvrditi početno stanje njihovog poznavanja i razumijevanja određenog područnog znanja da bi mogli kvantificirati veličinu njegove promjene. Inicijalni test omogućava prikupljanje informacija o postojanju statistički značajnih razlika u predznanju ispitanika.

Inicijalni test se piše prije uvođenja eksperimentalnih faktora. Uobičajeno je da inicijalno testiranje traje 45 minuta. Preporuča se raspon bodova na ljestvici 0-100 bodova radi lakše usporedbe s ostalim testovima i kasnije statističke analize.

4.1.1.2 DEFINIRANJE EKVIVALENTNIH GRUPA

Način na koji se bira uzorak ispitanika nad kojim ćemo provesti istraživanje o učinkovitosti, utjecat će na sakupljene informacije. Ako sami biramo uzorak, imamo mogućnost odabira ispitanika koji su najspremniji za suradnju ili imaju najprikladnije sposobnosti. Zato se trebamo služiti slučajnim odabirom na osnovu rezultata inicijalnog testa.

Najvažniji problem u vezi s predloženim eksperimentalnim modelom je kako stvoriti ekvivalentne grupe da bismo mogli dobiti što precizniju i bezpogovornu razliku u učincima eksperimentalnih faktora. Izjednačavanje grupa vrši se:

- određivanjem parova ispitanika uzimanjem po jednog ispitanika iz svake grupe kod kojih su podaci u vezi s pojedinim varijablama što sličniji; ispitanici koji previše odskoču iz neke distribucije ili koji se ne mogu spariti, izbacuju se iz eksperimenta,
- izjednačavanjem mjera centralne tendencije (obično aritmetička sredina) i mjera disperzije kao pokazatelja homogenosti podataka (obično standardna devijacija) uz ispitivanje statističke značajnosti uporabom t-testa.

Postoje dva načina izjednačavanja parova ispitanika [BACK2000]:

1. *Precizno izjednačavanje* (eng. exact matching)

Prvo treba sortirati ispitanike prema rezultatima inicijalnog testa. Zatim se pristupa traženju parova ispitanika s jednakim rezultatima inicijalnog testa. Možemo krenuti od kontrolne grupe i tražiti odgovarajuće ispitanike u eksperimentalnoj grupi, i obratno. Ukoliko postoji više ispitanika koji predstavljaju moguće uparivanje za odabranog ispitanika, pristupa se metodi slučajnog odabira.

Na primjer, promotrimo sljedeći uzorak ispitanika koji su slučajnim odabirom podijeljeni u dvije grupe koje treba izjednačiti (Tablica 4. 1.).

Tablica 4. 1. Rezultati inicijalnog testa

ID	Grupa	Inicijalni test	ID	Grupa	Inicijalni test
4	kontrolna	5	17	eksperimentalna	9
7	kontrolna	5	20	eksperimentalna	9
6	kontrolna	7	21	eksperimentalna	9
8	kontrolna	7	24	eksperimentalna	9
5	kontrolna	8	25	eksperimentalna	9
10	kontrolna	8	26	eksperimentalna	11
11	kontrolna	9	28	eksperimentalna	11
12	kontrolna	9	18	eksperimentalna	12
13	kontrolna	9	22	eksperimentalna	12
15	kontrolna	9	23	eksperimentalna	12
1	kontrolna	10	27	eksperimentalna	12
2	kontrolna	11	30	eksperimentalna	13
3	kontrolna	11	19	eksperimentalna	14
9	kontrolna	11	16	eksperimentalna	15
14	kontrolna	11	29	eksperimentalna	15

Krećemo od ispitanika iz eksperimentalne grupe i tražimo ispitanike s odgovarajućim brojem bodova u kontrolnoj grupi. Uzmimo ispitanika s ID 17 iz eksperimentalne grupe. Rezultat njegovog inicijalnog testa je 9. Potom tražimo u kontrolnoj grupi ispitanika s istim rezultatom na inicijalnom testu. U kontrolnoj grupi ispitanici s ID 11, 12, 13 i 15 imaju rezultat inicijalnog testa 9. Slučajnim odabirom između njih četvero biramo ispitanika s ID 12. Sljedeći na redu za uparivanje je ispitanik s ID 20 iz eksperimentalne grupe za kojeg ponavljamo prethodni postupak. Nakon provedenog preciznog izjednačavanja upareno je 40% ispitanika (6 od 15 mogućih parova - Tablica 4. 2.).

Tablica 4. 2. Parovi ispitanika nakon izjednačavanja

ID Eksperimentalna	Inicijalni test eksperimentalna	ID Kontrolna (slučajni odabir)
17	9	12
20	9	11
21	9	13
24	9	15
25	9	nema
26	11	2
28	11	3
18	12	nema
22	12	nema
23	12	nema
27	12	nema
30	13	nema
19	14	nema
16	15	nema
29	15	nema

2. Izjednačavanje u rasponu (eng. caliper matching)

Prvo je potrebno sortirati ispitanike prema rezultatima inicijalnog testa. Zatim, treba definirati raspon bodova koji dolaze u obzir kod uparivanja ispitanika prema rezultatima inicijalnog testa. Potom se pristupa traženju parova ispitanika s rezultatima inicijalnog testa koji su u definiranom rasponu bodova. Možemo krenuti od kontrolne grupe i tražiti odgovarajuće ispitanike u eksperimentalnoj grupi, i obratno. Prvo tražimo ispitanika s jednakom brojem bodova, zatim ako takav ne postoji, tražimo ispitanika kojem je rezultat inicijalnog testa u rasponu ± 1 , i tako dalje, povećavamo raspon bodova ili dok ne nađem odgovarajućeg ispitanika ili dok ne dostignemo definirani raspon. Ukoliko postoji više ispitanika koji predstavljaju moguće uparivanje za odabranog ispitanika, pristupa se metodi slučajnog odabira.

U prethodnom primjeru krećemo od ispitanika iz eksperimentalne grupe i tražimo ispitanike s odgovarajućim brojem bodova u kontrolnoj grupi. Raspon je ± 1 . Uzmimo ispitanika s ID 17 iz eksperimentalne grupe. Rezultat njegovog inicijalnog testa je 9. Potom tražimo u kontrolnoj grupi ispitanika s rezultatom na inicijalnom testu u rasponu od 8 do 10. U kontrolnoj grupi ispitanici s ID 5 i 10 imaju rezultat inicijalnog testa 8, ispitanici s ID 11, 12, 13 i 15 imaju rezultat inicijalnog testa 9, a ispitanik s ID 1 ima rezultat inicijalnog testa 10. Pošto prvo pristupamo preciznom izjednačavanju, slučajnim odabirom između četvero ispitanika s brojem bodova 9, biramo ispitanika s ID 12. Sljedeći na redu za uparivanje je ispitanik s ID 20 iz eksperimentalne grupe za kojeg ponavljamo prethodni postupak. Izjednačavanje u rasponu s ± 1 razlikom u bodovima daje, kod prethodnog primjera, dvostruko više parova učenika od preciznog izjednačavanja (67%, 10 od 15 mogućih parova - Tablica 4. 3.).

Tablica 4. 3. Izjednačavanje u rasponu

ID eksperimentalna	ID kontrolna (slučajni odabir) /Inicijalni test	Inicijalni test eksperimentalna	Caliper raspon (bod ± 1)
17	12/9	9	8-10
20	13/9	9	8-10
21	11/9	9	8-10
24	10/8	9	8-10
25	5/8	9	8-10
26	14/11	11	10-12
28	1/10	11	10-12
18	2/11	12	11-13
22	3/11	12	11-13
23	9/11	12	11-13
27	nema	12	11-13
30	nema	13	12-14
19	nema	14	13-15
16	nema	15	14-16
29	nema	15	14-16

Nakon provedenog izjednačavanja grupa prema rezultatima inicijalnog testa, pristupa se statističkoj provjeri ekvivalentnosti kontrolne i eksperimentalne grupe (veličine uzoraka su N_c i N_e). Sljedeći izračuni se provode samo ako su podaci normalno distribuirani prema rezultatima K-S testa za $\alpha=0.05$. Potrebno je prvo izračunati *aritmetičku sredinu* rezultata inicijalnog testa za svaku grupu po formuli:

$$\bar{X}_c = \frac{\sum X}{N_c} \quad \bar{X}_e = \frac{\sum X}{N_e} \quad (4.1)$$

Zatim se izračuna *standardna devijacija* (mjera za mjerenje varijabiliteta rezultata) za svaku grupu prema formuli:

$$\sigma_c = \sqrt{\frac{\sum (X - \bar{X}_c)^2}{N_c - 1}} \quad \sigma_e = \sqrt{\frac{\sum (X - \bar{X}_e)^2}{N_e - 1}} \quad (4.2)$$

Prema [PETZ1997] ako su i kontrolna i eksperimentalna grupa jednako velike, nije ni potrebno testirati F-testom da li im se varijance statistički značajno razlikuju ili ne, jer će se – ako se one razlikuju – pogreška u računu biti neznatna. Ako se te grupe razlikuju u veličini, onda treba ispitati statističku značajnost razlike između aritmetičkih sredina malih uzoraka. Zajedničku standardnu devijaciju smijemo izračunati samo onda ako se standardne devijacije samih grupa zaista značajno ne razlikuju, što se ispituje pomoću tzv. F-testa po formuli:

$$F = \frac{\text{veća } \sigma^2}{\text{manja } \sigma^2} \quad (4.3)$$

Ako su grupe jednako velike ili ako im se varijance ne razlikuju značajno, pristupamo izračunavanju *zajedničke standardne devijacije* koja se izračunava prema formuli:

$$\sigma_{\text{zajednicka}} = \sqrt{\frac{\sum (X - \bar{X}_c)^2 + \sum (X - \bar{X}_e)^2}{(N_c - 1) + (N_e - 1)}} \quad (4.4)$$

Standardna pogreška između dviju aritmetičkih sredina računa se prema formuli:

$$s_{\bar{x}_c - \bar{x}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}} \quad (4.5)$$

Ako je neka razlika između dvije aritmetičke sredine barem dva puta (točnije 1,96 puta) veća od svoje pogreške, onda se može smatrati statistički značajnom, jer je vrlo malo vjerojatno da će se tako velika razlika dogoditi slučajno. Uzima se razina značajnosti od 5%, što znači: ako zapravo ne postoji nikakva razlika između dviju aritmetičkih sredina, onda bi se takva konkretna razlika koju smo dobili, mogla slučajno dogoditi samo pet puta u 100 mjerenja, a to je malo vjerojatno, pa zato možemo uzeti da je razlika statistički značajna. Prema tome, razina slučajnosti od 5% znači zapravo šansu od 5% da smo pogriješili. Koliko je puta neka razlika veća od svoje pogreške možemo ustanoviti računanjem *t-vrijednosti*:

$$t = \frac{\text{razlika}}{\text{standardna pogreška te razlike}} = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{x}_c - \bar{x}_e}} \quad (4.6)$$

Broj stupnjeva slobode se izračunava prema formuli:

$$df=(N_c-1)+(N_e-1) \quad (4.7)$$

Ukoliko se pokaže da su prethodno definirane grupe statistički značajno različite (prema t-vrijednosti, razini značajnosti i stupnjevima slobode), potrebno ih je pokušati ponovno presložiti metodom slučajnog odabira i izjednačiti u rasponu, a zatim opet provjeriti da li su ekvivalentne.

4.1.1.3 IZRADA I PROVEDBA TESTOVA PROVJERE TE ZAVRŠNOG TESTA

Svi testovi koji se koriste u jednom eksperimentu trebaju biti potpuno komparativni. to znači da moraju biti bodovani na istoj ljestvici i da moraju ispitivati poznavanje istog područnog znanja. Ponavljanje istih pitanja nije dobro, ali je dobro zadržati dio gradiva iz inicijalnog testa i u testovima provjere i u završnom testu.

Obje grupe moraju pisati nekoliko 45-minutnih testova provjere, kao i 45-minutni završni test na kraju eksperimenta. Broj testova provjere se određuje prema trajanju eksperimenta (najmanje jedan mjesečno). Spomenuti testovi omogućavaju prikupljanje informacija o postojanju statistički značajnih razlika između dviju grupa tijekom trajanja eksperimenta. Raspon bodova ovih testova je također na ljestvici 0-100 bodova.

4.1.2. DEFINIRANJE NUL-HIPOTEZA

Nul-hipoteza u statistici znači „nema razlike“ među pojavama koje mjerimo. Stoga se za svaki provedeni test provjere i za završni test definiraju nul-hipoteze oblika:

H: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata xxx testa.“

gdje je xxx n-ti test provjere ili završni test.

U eksperimentima koji prate ovdje predloženu metodologiju definira se n hipoteza koje se odnose na n testova provjere, te jednu hipotezu koja se odnosi na završni test. Te hipoteze se tijekom ispitivanja statističke značajnosti rezultata odgovarajućeg testa, prihvaćaju ili odbacuju.

Da bi hipoteze mogli prihvatiti ili odbaciti, potrebno je izračunati razliku (eng. *gain*) između rezultata testova provjere i završnog testa s inicijalnim testom za kontrolnu i eksperimentalnu grupu. Zatim je potrebno izračunati t-vrijednost da bi se mogla potvrditi ili oboriti odgovarajuća nul-hipoteza, ali samo ako su podaci normalno distribuirani prema rezultatima K-S testa za $\alpha=0.05$. Izračun proveden nad razlikama u rezultatima najbolje je prikazati u Tablica 4. 4.:

Tablica 4. 4. Podaci potrebni za prihvaćanje ili odbacivanje nul-hipoteze

	KONTROLNA	EKSPERIMENTALNA
\bar{X}		
$\sum (X - \bar{X})^2$		
$\sigma^2 = \frac{\sum (X - \bar{X})^2}{N-1}$		
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}}$		
$F = \frac{\text{veća } \sigma^2}{\text{manja } \sigma^2}$		
$\sigma_{\text{zajednicka}} = \sqrt{\frac{\sum (X - \bar{X}_c)^2 + \sum (X - \bar{X}_e)^2}{(N_c - 1) + (N_e - 1)}}$		
$s_{\bar{x}_c - \bar{x}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$		
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{x}_c - \bar{x}_e}}, \text{ df, p, } \alpha=0.05$		

4.1.3. IZRAČUN VELIČINE UČINKA

Dok napravljeni testovi značajnosti govore nam o vjerojatnosti kojom se rezultati na testovima eksperimentalne grupe razlikuju od slučajnih rezultata, mjere veličine učinka govore o relativnoj magnitude eksperimentalnog faktora (u našem slučaju sustava xTEx-Sys). Veličina učinka je standardni način za usporedbu rezultata dva eksperimenta, odnosno za izradu meta-analiza.

U našem pristupu računamo *parcijalne veličine učinka* prema rezultatima svakog testa kao Glass-ove Δ (vidi formulu (3.53)), po sljedećim formulama u kojima koristimo aritmetičke sredine razlika između rezultata testova provjere i završnog testa te inicijalnog testa za obje grupe studenata, kao i standardnu devijaciju odgovarajuće razlike za kontrolnu grupu. Zbog usporedbe, izračunat ćemo i Hedges-ov g (prema formuli (3.62) koristeći zajedničku standardnu devijaciju).

$$\Delta_{tp1} = \frac{\bar{X}_{e_{tp1}} - \bar{X}_{c_{tp1}}}{\sigma_{c_{tp1}}}$$

$$\Delta_{tp2} = \frac{\bar{X}_{e_{tp2}} - \bar{X}_{c_{tp2}}}{\sigma_{c_{tp2}}}$$

-
-
-

$$\Delta_{tpn} = \frac{\bar{X}_{e_{tpn}} - \bar{X}_{c_{tpn}}}{\sigma_{c_{tpn}}}$$

$$\Delta_{završni} = \frac{\bar{X}_{e_{završni}} - \bar{X}_{c_{završni}}}{\sigma_{c_{završni}}}$$

(4.8)

Podatci potrebni za izračun parcijalnih veličina učinaka su sažeto prikazani u sljedećoj tablici:

Tablica 4. 5. Podaci potrebni za izračun parcijalnih veličina učinka

	RAZLIKE REZULTATA KONTROLNE GRUPE	RAZLIKE REZULTATA EKSPERIMENTALNE GRUPE
Test provjere n	$\bar{X}_{c_{tpn}} \quad \sigma_{c_{tpn}}$	$\bar{X}_{e_{tpn}}$
Završni test	$\bar{X}_{c_{završni}} \quad \sigma_{c_{završni}}$	$\bar{X}_{e_{završni}}$

Ukupna veličina učinka računa se kao prosječna vrijednost parcijalnih veličina učinka po formuli:

$$\Delta = \frac{\Delta_{tp1} + \Delta_{tp2} + \dots + \Delta_{tpn} + \Delta_{završni}}{n + 1} \quad (4.9)$$

4.1.4. ANALIZA I INTERPRETACIJA REZULTATA VREDNOVANJA UČINKA

Rezultati koji se dobiju primjenom neke metode za vrednovanje čine osnovu za pronalaženje odgovora na pitanje koje se postavlja prije samog istraživanja, tj. za interpretaciju tih rezultata. Ponekad je teško razlučiti analizu podataka od interpretacije. Analizom podataka se pojedinačni podaci sređuju i u njima se pronalaze smislene cjeline, tj. dolazi se do rezultata, dok interpretacija polazi od rezultata, uočava njihov značaj, traži se njihova međusobna povezanost i povezanost s

rezultatima nekih drugih istraživanja. Dakle, bez interpretacije svi koraci koji su prethodili dobivanju rezultata postaju nepotrebni [MUŽI1977].

Preduvjeti za ispravnu interpretaciju [MUŽI1977]:

1. Zaključci do kojih se dolazi interpretacijom trebaju biti logični i zasnovani na društvenoj i životnoj stvarnosti u kojoj istraživač živi.
2. Poznavanje područja koje se proučava, a osobito užeg problema koji je predmet istraživanja.
3. Poznavanje metodologije pedagoškog istraživanja zbog pravilnog odabira instrumenata pri statističkoj obradi podataka (npr. što nam neka određena mjera srednje vrijednosti kaže) i zbog određivanja da li je generalizacija rezultata moguća.

Neke pogreške pri interpretaciji [MUŽI1977]:

1. Prebrza generalizacija ili neprecizno izražena generalizacija. Ako uzorak na kojem se vrši interpretacija nije reprezentativan i samim time onemogućava valjanu generalizaciju, onda tu činjenicu treba istaknuti. Rješenje je u primjeni metode uzoraka.
2. Dokazivanje uzročno-posljedičnih veza. Pri interpretaciji rezultata koji su dobiveni računanjem korelacije može doći do pogrešaka pri određivanju uzročno-posljedičnih veza. Opasnost se krije u tome da se neka slučajna povezanost interpretira kao uzročno-posljedična povezanost.
3. Zaključivanje na temelju neistine. Javlja se kad istraživač prešućuje onaj rezultat koji nije onakav kakav bi on želio da bude i koji se ne slaže s hipotezom.

PRIHVAĆANJE ILI ODBACIVANJE NUL-HIPOTEZA

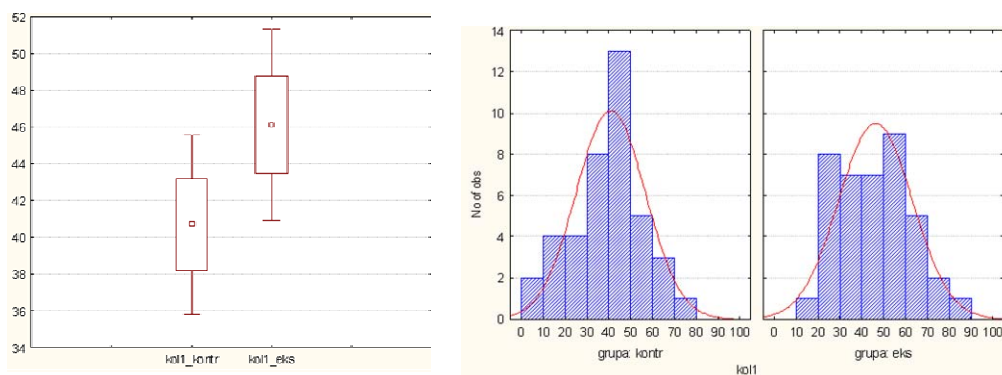
Nul-hipoteza se odbija ili kada t -vrijednost prijeđe kritičnu vrijednost na odabranoj razini značajnosti ili kada je p -vrijednost manja od razine značajnosti. Nul-hipoteza se prihvaća ili kada je t -vrijednost manja od kritične vrijednosti na odabranoj razini značajnosti ili kada je p -vrijednost veća od razine značajnosti $\alpha=0.05$.

Ako je apsolutna vrijednost dobivene t -vrijednosti manja od granične (za određene stupnjeve slobode) ili ako je p -vrijednost veća od α , to znači da razlika nije statistički značajna, odnosno potvrđujemo nul-hipotezu. Dakle, ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata određenog testa.

Ako je apsolutna vrijednost dobivene t -vrijednosti veća od granične (za određene stupnjeve slobode) ili ako je p -vrijednost manja od α , to znači da je razlika statistički značajna, odnosno odbacili smo nul-hipotezu. Dakle, postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata određenog testa.

Pozitivna t -vrijednost upućuje na to da je kontrolna grupa bolja od eksperimentalne prema rezultatima određenog testa uz $\alpha=0.05$. Negativna t -vrijednost upućuje na to da je eksperimentalna grupa bolja od kontrolne prema rezultatima određenog testa uz $\alpha=0.05$.

Usporedba aritmetičkih sredina i varijacija dviju grupa vizualno se prikazuje s *box-and-whisker plots* (Slika 4. 1.) i histogramima koji omogućavaju intuitivno vizualiziranje veličine povezanosti dviju grupa.



Slika 4. 1. Vizualna interpretacija razlike između dvije grupe

Rezultate provedene analize poželjno je sažeto prikazati u sljedećoj tablici:

Tablica 4. 6. Rezultati analize

	t-VRIJEDNOST p-VRIJEDNOST	POSTOJANJE STATISTIČKI ZNAČAJNE RAZLIKE	BOLJA GRUPA
Prvi test provjere n	t = , df=, p=, $\alpha=0.05$	da/ne	eksperimentalna/kontrolna
Završni test	t = , df=, p=, $\alpha=0.05$	da/ne	eksperimentalna/kontrolna

Ako postoji statistički značajna razlika kod svakog testa (test provjere i završni test) koja je ista ili raste, to znači da sustav e-učenja ima pozitivan učinak na učenikovo razumijevanje područnog znanja.

4.2. VREDNOVANJE UČINKA SUSTAVA *xTeX-Sys*

Vrednovanje učinka sustava *xTeX-Sys* (objašnjenog u 2. poglavlju) obavljeno je primjenom prethodno opisane metodologije. U eksperimentu je sudjelovalo 175 studenata koji su u akademskoj godini 2005/06 bili na prvoj godini Fakulteta prirodoslovno-matematičkih znanosti i kineziologije i Kemijsko-tehnološkog fakulteta Sveučilišta u Splitu. Eksperiment je obavljen u okviru kolegija Uvod u računarstvo koji se realizira istovremeno na oba fakulteta. Eksperiment je trajao od 15.10.2005. do 23.01.2006.

Jedina razlika u odnosu na definiranu metodologiju je u tome što su, zbog organizacijski problema, unaprijed definirane kontrolna i eksperimentalna grupa. U kontrolnoj grupi su bili studenti prve godine Kemijsko-tehnološkog fakulteta, a u eksperimentalnoj studenti prve godine Fakulteta prirodoslovno-matematičkih znanosti i kineziologije. Unaprijed određena pripadnost grupi nije predstavljala ograničenje provedenom eksperimentu, jer su grupe, nakon provedenih

svih testova, izjednačene i samim time su predstavljale prihvatljiv uzorak za statističku analizu i interpretaciju učinkovitosti.

Eksperiment je realiziran u okviru 7 etapa (Tablica 4. 7).

Tablica 4. 7. Etape eksperimenta

Etapa	Aktivnost	Datum
1.	45-minutni inicijalni test	15.10.2005.
2.	učenje i poučavanje: kontrolna grupa na tradicionalni način eksperimentalna grupa u sustavu xTEx-Sys	od 16.10.2005. do 19.11.2005.
3.	45-minutni prvi test provjere	19.11.2005.
4.	učenje i poučavanje: kontrolna grupa na tradicionalni način eksperimentalna grupa u sustavu xTEx-Sys	od 20.11.2005. do 17.12.2005.
5.	45-minutni drugi test provjere	17.12.2005.
6.	učenje i poučavanje: kontrolna grupa na tradicionalni način eksperimentalna grupa u sustavu xTEx-Sys	od 28.11.2005. do 23.01.2006.
7.	45-minutni završni test	23.01.2006.

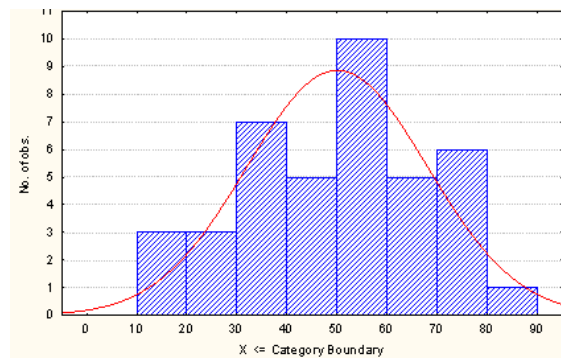
Na završetku eksperimenta je napravljena analiza prisustvovanja testovima u okviru eksperimenta. Ustanovljeno je sljedeće:

1. inicijalni test su pisala 73 studenta iz kontrolne grupe i 102 studenta iz eksperimentalne grupe (Prilog A)
2. prvi test provjere (Prilog B) su pisala 72 studenta iz kontrolne grupe (od toga 12 studenata koji nisu pisali inicijalni test) i 99 studenata iz eksperimentalne grupe (od toga 5 studenata koji nisu pisali inicijalni test)
3. drugi test provjere (Prilog C) je pisalo 57 studenata iz kontrolne grupe (od toga 9 studenata koji nisu pisali inicijalni test ili prvi test provjere) i 94 studenta iz eksperimentalne grupe (od toga 8 studenata koji nisu pisali inicijalni test ili prvi test provjere)
4. završni test (Prilog D) je pisalo 49 studenata iz kontrolne grupe (od toga 9 studenata koji nisu pisali inicijalni test ili prvi test provjere ili drugi test provjere) i 87 studenata iz eksperimentalne grupe (od toga 7 studenata koji nisu pisali inicijalni test ili prvi test provjere ili drugi test provjere)

Dakle, na kraju smo dobili da u kontrolnoj grupi ima 40 studenata koji imaju sva četiri testa, a u eksperimentalnoj grupi 80 studenata koji imaju sva četiri testa. Potrebno je stoga od 80 studenata u eksperimentalnoj grupi izabrati 40 tako da dobivena grupa bude ekvivalentna kontrolnoj grupi promatrajući rezultate inicijalnog testa.

4.2.1. DEFINIRANJE EKVIVALENTNIH GRUPA

Proučavajući rezultate inicijalnog testa *kontrolne grupe*, vidi se da je p-vrijednosti K-S testa za $\alpha=0.05$ veća od 0.05 (vidi Tablica 4. 8), pa možemo zaključiti da se radi o normalnoj distribuciji. Također se mogu uočiti sljedeće frekvencije bodova prikazane histogramima (na osi x su kategorije bodova, a na osi y broj pojedinih rezultata u određenoj kategoriji) (Slika 4. 2.):



Slika 4. 2. Histogram rezultata inicijalnog testa kontrolne grupe

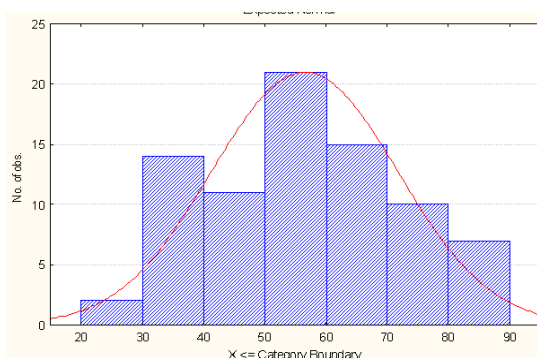
Tablica 4. 8. Tablica normalnosti za kontrolnu grupu

Test	K-S p
Inicijalni test	$p > 0,20$
Prvi test provjere	$p > 0,20$
Drugi test provjere	$p > 0,20$
Završni test	$p > 0,20$

Na temelju histograma rezultata inicijalnog testa kontrolne grupe trebalo je napraviti izbor od 40 studenata iz *proširene eksperimentalne grupe* od 80 studenata, koji bi tvorili statistički ekvivalent kontrolnoj grupi. Ta ekvivalentna eksperimentalna grupa trebala je, dakle, sadržavati:

- 0 studenata kojima je rezultat na inicijalnom testu između 0 i 10 bodova,
- 3 studenata kojima je rezultat na inicijalnom testu između 10 i 20 bodova,
- 3 studenata kojima je rezultat na inicijalnom testu između 20 i 30 bodova,
- 7 studenata kojima je rezultat na inicijalnom testu između 30 i 40 bodova,
- 5 studenata kojima je rezultat na inicijalnom testu između 40 i 50 bodova,
- 10 studenata kojima je rezultat na inicijalnom testu između 50 i 60 bodova,
- 5 studenata kojima je rezultat na inicijalnom testu između 60 i 70 bodova,
- 6 studenata kojima je rezultat na inicijalnom testu između 70 i 80 bodova,
- 1 studenata kojima je rezultat na inicijalnom testu između 80 i 90 bodova,
- 0 studenata kojima je rezultat na inicijalnom testu između 90 i 100 bodova.

Promotrimo frekvencije bodova inicijalnog testa za *proširenu eksperimentalnu grupu* prikazane histogramom (Slika 4. 3). P-vrijednosti K-S testa za $\alpha=0.05$ veća od 0.05 (vidi Tablica 4. 9), pa možemo zaključiti da se radi o normalnoj distribuciji.



Slika 4. 3. Histogram rezultata inicijalnog testa proširene eksperimentalne grupe

Tablica 4. 9. Tablica normalnosti za proširenu eksperimentalnu grupu

Test	K-S p
Inicijalni test	p>0,20
Prvi test provjere	p>0,20
Drugi test provjere	p>0,20
Završni test	p>0,20

Vidimo da *proširena eksperimentalna grupa* sadrži:

- 0 studenta kojima je rezultat na inicijalnom testu između 0 i 10 bodova,
- 0 studenta kojima je rezultat na inicijalnom testu između 10 i 20 bodova,
- 2 studenta kojima je rezultat na inicijalnom testu između 20 i 30 bodova,
- 14 studenta kojima je rezultat na inicijalnom testu između 30 i 40 bodova,
- 11 studenta kojima je rezultat na inicijalnom testu između 40 i 50 bodova,
- 21 studenta kojima je rezultat na inicijalnom testu između 50 i 60 bodova,
- 15 studenta kojima je rezultat na inicijalnom testu između 60 i 70 bodova,
- 10 studenta kojima je rezultat na inicijalnom testu između 70 i 80 bodova,
- 7 studenta kojima je rezultat na inicijalnom testu između 80 i 90 bodova,
- 0 studenta kojima je rezultat na inicijalnom testu između 90 i 100 bodova.

Prikažimo frekvencije bodova inicijalnog testa kontrolne grupe i proširene eksperimentalne grupe u sljedećoj tablici:

Tablica 4. 10. Frekvencije bodova inicijalnog testa kontrolne grupe i proširene eksperimentalne grupe

BODOVI	KONTROLNA	PROŠIRENA EKSPERIMENTALNA	EKVIVALENTNA EKSPERIMENTALNA
0-10	0	0	0
10-20	3	0	???
20-30	3	2	2+?
30-40	7	14	7
40-50	5	11	5
50-60	10	21	10
60-70	5	15	5
70-80	6	10	6
80-90	1	7	1
90-100	0	0	0

Nakon što smo usporedili dobivene frekvencije uvidjeli smo da *ekvivalentna eksperimentalna grupa* treba sadržavati jednaki broj studenata kao i kontrolna grupa, koji imaju rezultat na inicijalnom testu između 30 i 100 bodova. Problem smo uočili kod rezultata koji su između 10 i 30 bodova. Naime, u *proširenoj eksperimentalnoj grupi* nema studenata koji su na inicijalnom testu postigli rezultat koji je između 10 i 20 bodova, a postoje samo dva (od potrebnih tri) studenta koji su na inicijalnom testu postigli broj bodova između 20 i 30 bodova. Potrebno je odabrati još četiri studenta iz *proširene eksperimentalne grupe* tako da između eksperimentalna grupa i kontrolna grupa budu jednako brojne.

Izjednačavanje grupa po rasponu uz dodavanje četiri studenta slučajnim odabirom (označeni su podebljano u tablici) zbog postizanja jednakobrojnosti grupa, dalo je sljedeću strukturu grupa:

Tablica 4. 11. Izjednačavanje grupa po rasponu

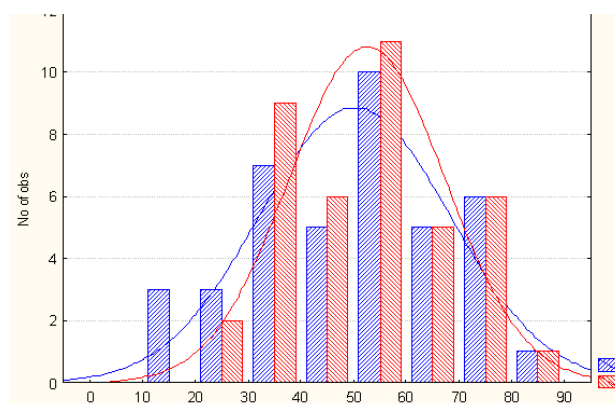
kontrolna grupa		ID eksperimentalna	Caliper raspon ± 3	
ID	inicijalni test			
student141	13	student37	10	16
student146	16	student52	13	19
student159	18	student45	15	21
student109	25	student9	22	28
student119	25	student59	22	28
student173	30	student89	27	33
student105	31	student12	28	34
student170	31	student39	28	34
student153	33	student58	30	36
student156	37	student65	34	40
student164	37	student67	34	40
student117	39	student66	36	42
student139	39	student72	36	42
student149	43	student43	40	46
student113	45	student68	42	48
student122	45	student73	42	48
student126	45	student14	42	48
student129	49	student44	46	52
student144	51	student36	48	54
student168	51	student93	48	54
student134	52	student4	49	55
student114	54	student13	51	57
student103	55	student16	52	58
student175	55	student42	52	58
student104	58	student33	55	61
student120	58	student15	55	61
student143	58	student25	55	61
student107	60	student91	57	63
student108	63	student64	60	66
student154	63	student101	60	66
student169	63	student40	60	66
student158	64	student28	61	67
student160	64	student46	61	67
student106	72	student26	69	75
student128	75	student7	72	78
student150	75	student34	72	78
student155	75	student27	72	78
student157	75	student71	72	78
student162	78	student6	75	81
student151	81	student79	78	84

Nakon izjednačavanja grupa po rasponu dobili smo sljedeće frekvencije eksperimentalne i kontrolne grupe (Tablica 4. 12.):

Tablica 4. 12. Frekvencije eksperimentalne i kontrolne grupe

BODOVI	KONTROLNA	EKSPERIMENTALNA
0-10	0	0
10-20	3	0
20-30	3	2
30-40	7	9
40-50	5	6
50-60	10	11
60-70	5	5
70-80	6	6
80-90	1	1
90-100	0	0
Ukupno:	$N_c=40$	$N_e=40$

Grafički prikaz frekvencija testnih grupa vidi se na Slika 4. 4.



Slika 4. 4. Zajednički histogram kontrolne i eksperimentalne grupe

Nad ovako definiranim grupama bilo je potrebno ispitati postojanje statistički značajnih razlika. Izračunata je aritmetička sredina, varijanca, standardna devijacija, zajednička standardna devijacija, standardna pogreška između dviju aritmetičkih sredina, t i p-vrijednosti (Tablica 4. 13).

Tablica 4. 13. Podaci potrebni za ispitivanje statistički značajne razlike

	KONTROLNA	EKSPERIMENTALNA
\bar{X}	50,00	52,31
$\sum (X - \bar{X})^2$	12653,15	8491,65
$\sigma^2 = \frac{\sum (X - \bar{X})^2}{N-1}$	324,44	217,73
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}}$	18,01	14,76
$F = \frac{\text{veća } \sigma^2}{\text{manja } \sigma^2}$	Nije potrebno računati jer su grupe jednakobrojne	
$\sigma_{\text{zajednicka}} = \sqrt{\frac{\sum (X - \bar{X}_c)^2 + \sum (X - \bar{X}_e)^2}{(N_c - 1) + (N_e - 1)}}$	16,46	
$s_{\bar{x}_c - \bar{x}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	3,68	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{x}_c - \bar{x}_e}}, \text{ df, p, } \alpha=0.05$	t=-0,63, df=78, p=0,5305	

Broj stupnjeva slobode se izračunava prema formuli $(N_c-1)+(N_e-1)$, a u našem slučaju je taj broj 78. Iz tablice graničnih t-vrijednosti, granična vrijednost t na razini značajnosti 5%, a uz 78 stupnjeva slobode iznosi 1,99. Dakle, pošto je apsolutna vrijednost dobivene t-vrijednosti manja od granične, a, p-vrijednost veća od 0.05, to znači da razlika nije statistički značajna, odnosno kontrolna i eksperimentalna grupa su ekvivalentne prema rezultatima inicijalnog testa pa nad njima može vršiti sva daljnja izračunavanja potrebna za dobivanje veličine učinka xTeX-Sys-a.

4.2.2. DEFINIRANJE NUL-HIPOTEZA

Za svaki provedeni test provjere i za završni test se definiraju nul-hipoteze:

H1: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata prvog testa provjere.“

H2: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata drugog testa provjere.“

H3: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata završnog testa.“

Da bi hipoteze mogli prihvatiti ili odbaciti, potrebno je izračunati razliku (eng. *gain*) između rezultata testova provjere i završnog testa s inicijalnim testom za kontrolnu i eksperimentalnu grupu (svi rezultati testova nalaze se u Prilogu E). Zatim je potrebno izračunati t-vrijednost da bi se mogla potvrditi ili oboriti odgovarajuća nul-hipoteza. Izračun proveden nad razlikama u rezultatima prikazan je za svaku hipotezu posebno.

HIPOTEZA H1

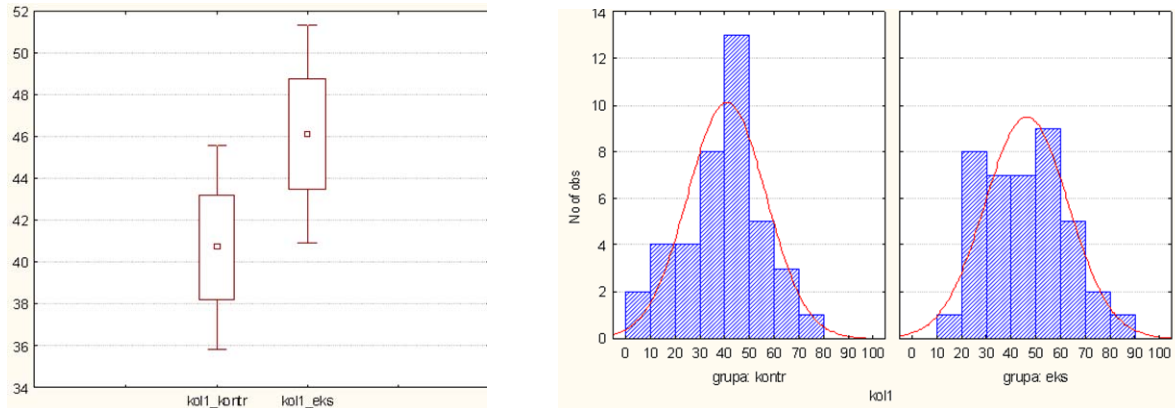
H1: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata prvog testa provjere.“

Tablica 4. 14. Podaci potrebni za prihvaćanje ili odbacivanje nul-hipoteze H1

	KONTROLNA	EKSPERIMENTALNA
\bar{X}	-9,28	-6,19
$\sum (X - \bar{X})^2$	12341,10	14018,40
$\sigma^2 = \frac{\sum (X - \bar{X})^2}{N-1}$	316,44	359,45
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}}$	17,79	18,96
$F = \frac{\text{veća } \sigma^2}{\text{manja } \sigma^2}$	Nije potrebno računati jer su grupe jednakobrojne	
$\sigma_{\text{zajednicka}} = \sqrt{\frac{\sum (X - \bar{X}_c)^2 + \sum (X - \bar{X}_e)^2}{(N_c - 1) + (N_e - 1)}}$	18,38	
$s_{\bar{x}_c - \bar{x}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	4,11	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{x}_c - \bar{x}_e}}, \text{ df, p, } \alpha=0.05$	t=-0,73, df=78, p=0,4676	

Dakle, pošto je apsolutna vrijednost dobivene t-vrijednosti manja od granične (1,99) i p-vrijednost veća od $\alpha=0.05$, to znači da razlika nije statistički značajna, odnosno potvrđujemo nul-

hipotezu H1. Dakle, ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata prvog testa provjere. Negativna t-vrijednost upućuje na to da je eksperimentalna grupa bolja od kontrolne prema rezultatima prvog testa provjere uz razinu značajnosti $\alpha=0.05$.



Slika 4. 5. Histogram i box-and-whisker plot za prvi test provjere

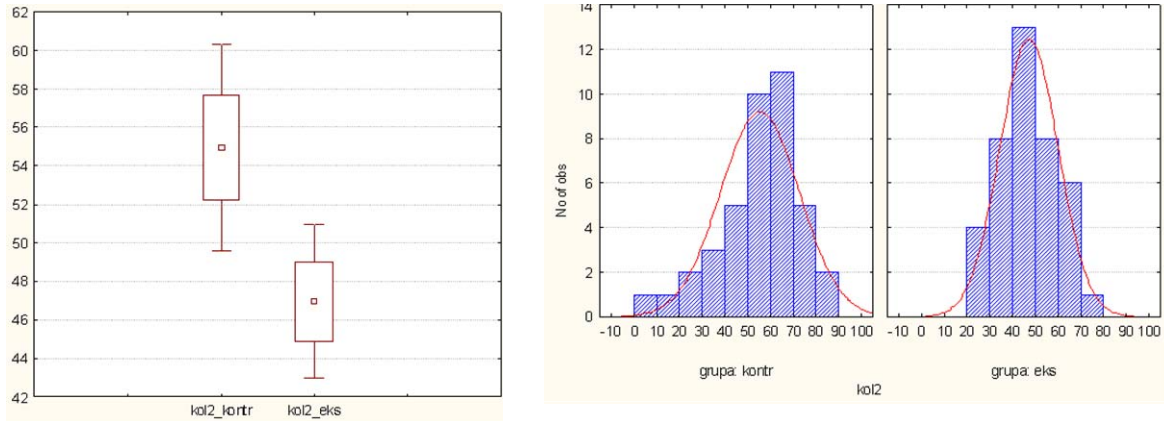
HIPOTEZA H2

H2: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata drugog testa provjere.“

Tablica 4. 15. Podaci potrebni za prihvaćanje ili odbacivanje nul-hipoteze H2

	KONTROLNA	EKSPERIMENTALNA
\bar{X}	4,95	-5,36
$\sum(X - \bar{X})^2$	18542,80	12573,94
$\sigma^2 = \frac{\sum(X - \bar{X})^2}{N-1}$	475,46	322,41
$\sigma = \sqrt{\frac{\sum(X - \bar{X})^2}{N-1}}$	21,80	17,96
$F = \frac{\text{veća } \sigma^2}{\text{manja } \sigma^2}$	Nije potrebno računati jer su grupe jednakobrojne	
$\sigma_{zajednicka} = \sqrt{\frac{\sum(X - \bar{X}_c)^2 + \sum(X - \bar{X}_e)^2}{(N_c - 1) + (N_e - 1)}}$	19,97	
$s_{\bar{x}_c - \bar{x}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	4,47	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{x}_c - \bar{x}_e}}, \text{ df, p, } \alpha=0.05$	t=2,31, df=78, p=0,0235	

Dakle, pošto je apsolutna vrijednost dobivene t-vrijednosti veća od granične (1,99) i p-vrijednost manja od $\alpha=0.05$, to znači da je razlika statistički značajna, odnosno oborili smo nul-hipotezu H2. Dakle, postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata drugog testa provjere. Pozitivna t-vrijednost upućuje na to da je kontrolna grupa bolja od eksperimentalne prema rezultatima drugog testa provjere uz razinu značajnosti $\alpha=0.05$.



Slika 4. 6. Histogram i box-and-whisker plot za drugi test provjere

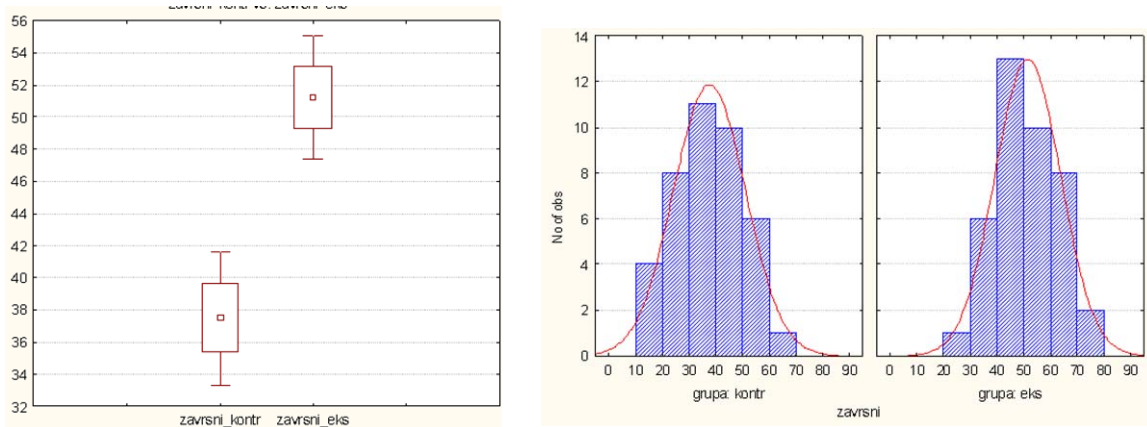
HIPOTEZA H3

H3: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata završnog testa.“

Tablica 4. 16. Podaci potrebni za prihvaćanje ili odbacivanje nul-hipoteze H3

	KONTROLNA	EKSPERIMENTALNA
\bar{X}	-12,52	-1,09
$\sum(X - \bar{X})^2$	8141,94	7429,24
$\sigma^2 = \frac{\sum(X - \bar{X})^2}{N-1}$	208,77	190,49
$\sigma = \sqrt{\frac{\sum(X - \bar{X})^2}{N-1}}$	14,45	13,80
$F = \frac{\text{veća } \sigma^2}{\text{manja } \sigma^2}$	Nije potrebno računati jer su grupe jednakobrojne	
$\sigma_{zajednicka} = \sqrt{\frac{\sum(X - \bar{X}_c)^2 + \sum(X - \bar{X}_e)^2}{(N_c - 1) + (N_e - 1)}}$	14,13	
$s_{\bar{x}_c - \bar{x}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	3,16	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{x}_c - \bar{x}_e}}, \text{ df, p, } \alpha=0.05$	t=-3,62, df=78, p=0,0005	

Dakle, pošto je apsolutna vrijednost dobivene t-vrijednosti veća od granične (1,99) i p-vrijednost manja od $\alpha=0.05$, to znači da je razlika statistički značajna, odnosno oborili smo nul-hipotezu H3. Dakle, postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata završnog testa. Negativna t-vrijednost upućuje na to da je eksperimentalna grupa bolja od kontrolne prema rezultatima završnog testa uz razinu značajnosti $\alpha=0.05$.



Slika 4. 7. Histogram i box-and-whisker plot za završni test

Prikažimo dobivene rezultate u sažeto tablici:

Tablica 4. 17. Rezultati ispitivanja nul-hipoteza

	t-VRIJEDNOST p-VRIJEDNOST $\alpha=0.05$, $df=78$	POSTOJANJE STATISTIČKI ZNAČAJNE RAZLIKE	BOLJA GRUPA
Prvi test provjere	$t = -0,73$, $p=0,4676$	ne	eksperimentalna
Drugi test provjere	$t = 2,31$, $p=0,0235$	da	kontrolna
Završni test	$t = -3,62$, $p=0,0005$	da	eksperimentalna

4.2.3. IZRAČUN VELIČINE UČINKA

Podatci potrebni za izračun parcijalnih veličina učinaka su sažeto prikazani u sljedećoj tablici:

Tablica 4. 18. Podaci potrebni za izračun parcijalnih veličina učinaka

	RAZLIKE REZULTATA KONTROLNE GRUPE	RAZLIKE REZULTATA EKSPERIMENTALNE GRUPE
Prvi test provjere	$\bar{X}_{c_{p1}} = -9,28$ $\sigma_{c_{p1}} = 17,79$	$\bar{X}_{e_{p1}} = -6,19$
Drugi test provjere	$\bar{X}_{c_{p2}} = 4,95$ $\sigma_{c_{p2}} = 21,80$	$\bar{X}_{e_{p2}} = -5,36$
Završni test	$\bar{X}_{c_{završni}} = -12,52$ $\sigma_{c_{završni}} = 14,45$	$\bar{X}_{e_{završni}} = -1,09$

Parcijalne veličine učinka su izračunate kao Glass-ove Δ . U našem slučaju dobili smo sljedeće rezultate:

$$\begin{aligned}\Delta_{ip1} &= \frac{\bar{X}_{e_{ip1}} - \bar{X}_{c_{ip1}}}{\sigma_{c_{ip1}}} = \frac{-6,19 - (-9,28)}{17,79} = 0,17 \\ \Delta_{ip2} &= \frac{\bar{X}_{e_{ip2}} - \bar{X}_{c_{ip2}}}{\sigma_{c_{ip2}}} = \frac{-5,36 - 4,95}{21,80} = -0,47 \\ \Delta_{završni} &= \frac{\bar{X}_{e_{završni}} - \bar{X}_{c_{završni}}}{\sigma_{c_{završni}}} = \frac{-1,09 - (-12,52)}{14,45} = 0,79\end{aligned}\quad (4.10)$$

Drugim riječima, eksperimentalna grupa je bila bolja na prvom testu provjere od kontrolne grupe za 0,17 standardne devijacije. Kontrolna grupa je bila bolja na drugom testu provjere od eksperimentalne grupe za 0,47 standardne devijacije. Eksperimentalna grupa je bila bolja na završnom testu od kontrolne grupe za 0,79 standardne devijacije.

Ukupna veličina učinka računa se kao prosječna vrijednost parcijalnih veličina učinka. U našem slučaju dobivamo:

$$\Delta = \frac{0,17 + (-0,47) + 0,79}{3} = 0,16 \quad (4.11)$$

Zbog usporedbe, izračunat ćemo i Hedges-ov g (prema formuli (3.62) koristeći zajedničku standardnu devijaciju).

$$\begin{aligned}\Delta_{ip1} &= \frac{\bar{X}_{e_{ip1}} - \bar{X}_{c_{ip1}}}{\sigma_{zajednicka_{ip1}}} = \frac{-6,19 - (-9,28)}{18,38} = 0,17 \\ \Delta_{ip2} &= \frac{\bar{X}_{e_{ip2}} - \bar{X}_{c_{ip2}}}{\sigma_{zajednicka_{ip2}}} = \frac{-5,36 - 4,95}{19,97} = -0,52 \\ \Delta_{završni} &= \frac{\bar{X}_{e_{završni}} - \bar{X}_{c_{završni}}}{\sigma_{zajednicka_{završni}}} = \frac{-1,09 - (-12,52)}{14,13} = 0,81\end{aligned}\quad (4.12)$$

Ukupna veličina učinka izračunata koristeći Hedges-ove veličine učinaka je:

$$\Delta = \frac{0,17 + (-0,52) + 0,81}{3} = 0,15 \quad (4.13)$$

Negativna t -vrijednost dobivena kod ispitivanja hipoteze $H1$ upućivala je na to da je eksperimentalna grupa bolja od kontrolne prema rezultatima prvog testa provjere uz razinu značajnosti $\alpha=0.05$, što odgovara veličini učinka 0,17 (koja je prema [COHE1969] mala veličina učinka) eksperimentalne grupe.

Positivna t -vrijednost kod ispitivanja hipoteze $H2$ upućivala je na to da je kontrolna grupa bolja od eksperimentalne prema rezultatima drugog testa provjere uz razinu značajnosti $\alpha=0.05$, što

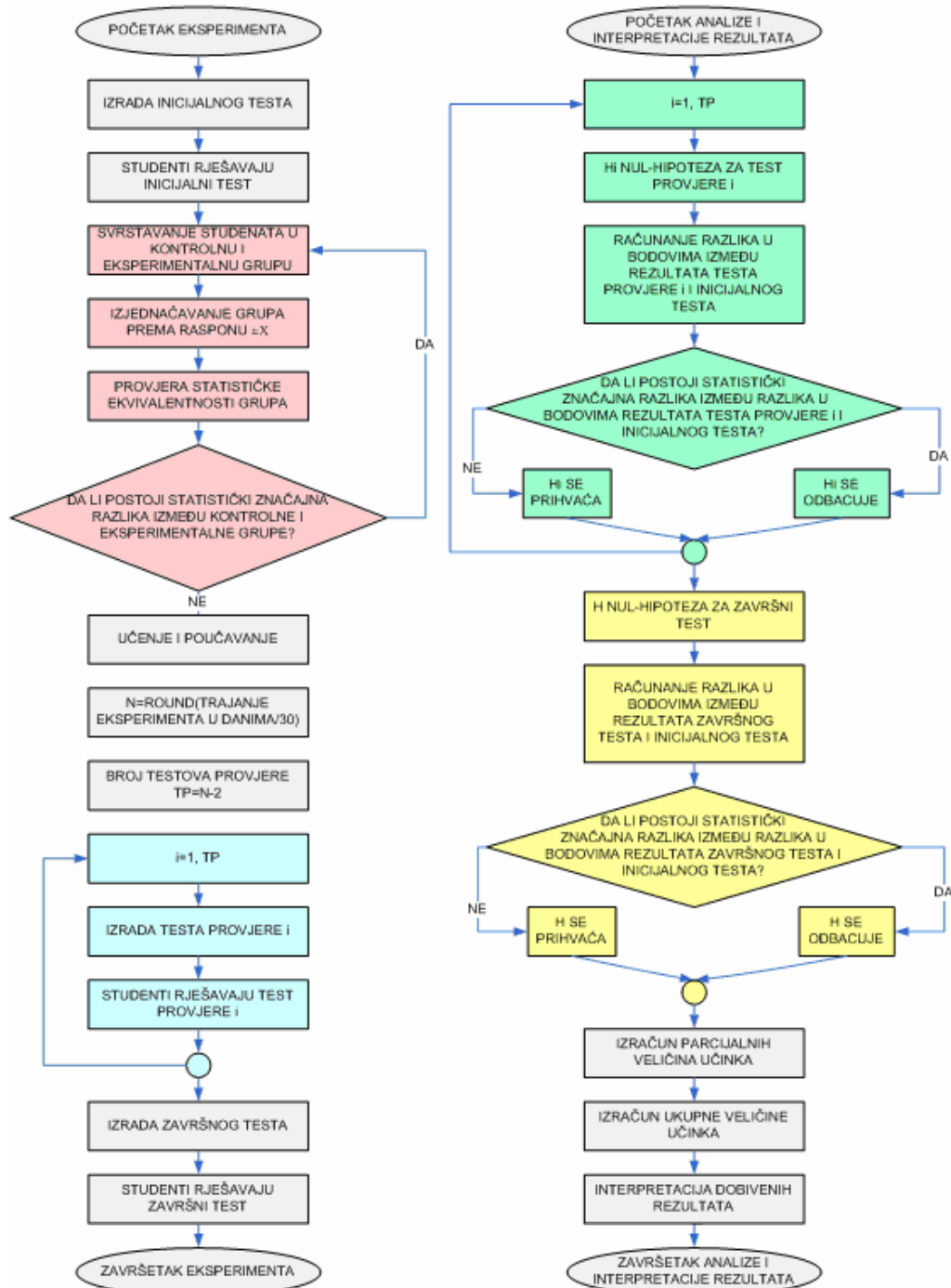
odgovara veličini učinka $-0,47$ (koja je prema [COHE1969] srednja veličina učinka) kontrolne grupe.

Negativna t-vrijednost kod ispitivanja hipoteze H3 upućivala je na to da je eksperimentalna grupa bolja od kontrolne prema rezultatima završnog testa uz razinu značajnosti $\alpha=0.05$, što odgovara veličini učinka $0,81$ (koja je prema [COHE1969] velika veličina učinka) eksperimentalne grupe.

Dobivena ukupna veličina učinka $0,16$ je prema [COHE1969] mala veličina učinka.

Vidimo da dobivene veličine učinaka, izračunate kao Glass-ove Δ ili kao Hedges-ov g , ne pokazuju linearan rast ili stagnaciju. Drugi test provjere je pokazao da je kontrolna grupa bolja od eksperimentalne, što je pak rezultat propusta u organizaciji eksperiment, kada je eksperimentalna grupa pisala drugi test provjere dva dana prije kontrolne grupe, uslijed čega je očito došlo do razmjene informacija između studenata.

Metodologija za vrednovanje učinka koja je u ovom poglavlju opisana, prikladna je za vrednovanje učinka bilo kojeg sustava e-učenja. Ona je u svrhu demonstracije, primijenjena za vrednovanje učinka inteligentnog tutorskog sustava, što je nipošto ne ograničava samo na tu klasu sustava. Slijed koraka koji se moraju izvršiti u realizacije opisane metodologije, može se prikladno prikazati dijagramom toka koji je prikazan na sljedećoj slici.



Slika 4. 8. Dijagram toka metodologije za vrednovanje

5. PROTOTIP PROGRAMSKOG ALATA ZA VREDNOVANJE UČINKA SUSTAVA E-UČENJA

Postoje brojne aplikacije za izradu i isporuku elektroničkih (eng. on-line) testova te izvještavanje o rezultatima tih testova. Tu se prvenstveno radi o upravljanju testovima (eng. test management), tj. izradi, isporuci i ocjenjivanju testova. Međutim, nema programskog alata koji bi nad tako dobivenim rezultatima testova proveo određenu statističku analizu. Moramo naglasiti da je upravljanje testovima nužno, ali ne i dovoljno, za vrednovanje učinka sustava e-učenja. Kao posljedica toga se javlja potreba za stvaranjem programskog alata za automatsko vrednovanje učinka sustava e-učenja. Naglasak je pri tom na riječ „automatsko“, jer se učinkovitost može izračunati „ručno“, koristeći različitu programsku podršku za statističku obradu podataka (npr. Statistica 7.0, SPSS,...), te se oslanjajući na neku od metodologija vrednovanja, što može biti zamoran i dugotrajan posao. Programski alat za automatsko vrednovanje učinka omogućava jednostavno izračunavanje učinkovitosti. Samim time, omogućava jednako tako jednostavnu provedbu meta-analiza nad rezultatima vrednovanja učinka dobivenih primjenom spomenutog programskog alata, jer su ti rezultati dobiveni primjenom iste metodologije vrednovanja.

Programski alat za automatsko vrednovanje učinka sustava e-učenja, kojeg ovdje opisujemo, nazvan je EVEDIN (EVALUATION of EDUCATIONAL INFLUENCE). EVEDIN izračunava veličinu učinka na temelju informacija o postignućima učenika na testovima, jer se znanje učenika utvrđuje primjenom nekoliko elektroničkih testova: inicijalnog testa, testova provjere te završnog testa. Rezultati spomenutih testova prolaze kroz EVEDIN-ov mehanizam za statističku analizu, te kao rezultat EVEDIN daje izračunatu veličinu učinka sustava e-učenja.

5.1. PRIMJERI ALATA ZA IZRADU I PROVEDBU ELEKTRONIČKIH TESTOVA ZNANJA

Provjera znanja je važan dio procesa učenja i poučavanja. Uporabom računala u procesu učenja i poučavanja došlo je i do promjene samog načina provjeravanja znanja, ponajviše primjenom elektroničkog ispitivanja (eng. online assessment) [JAME2002]. Ovakvim načinom testiranja moguće je postići uštede u vremenu i novcu, jer ono omogućava jednostavniju izradu i ispravljanje testova te nije potrebno fotokopirati testove, organizirati nastavnike koji trebaju paziti da li učenici prepisuju tijekom testiranja, omogućava da ispitu istovremeno pristupi velik broj učenika [DAVI2005]. Osnovno obilježje elektroničkog ispitivanja je uporaba računalne

mreže preko koje učenik pristupa ispitnom materijalu, rješava zadatke te dostavlja rješenja koja se potom automatski pohranjuju i ocjenjuju.

Postoji čitav niz dostupnih programskih rješenja koja služe za elektroničko ispitivanje. Odabir nekih od tih alata je napravljen nakon provedenog pretraživanja uporabom popularnih pretraživača WWW-a, te je u Tablica 5. 1. uz usporedbu njihovih značajki.

Tablica 5. 1. Usporedba alata za izradu i provedbu elektroničkih testova

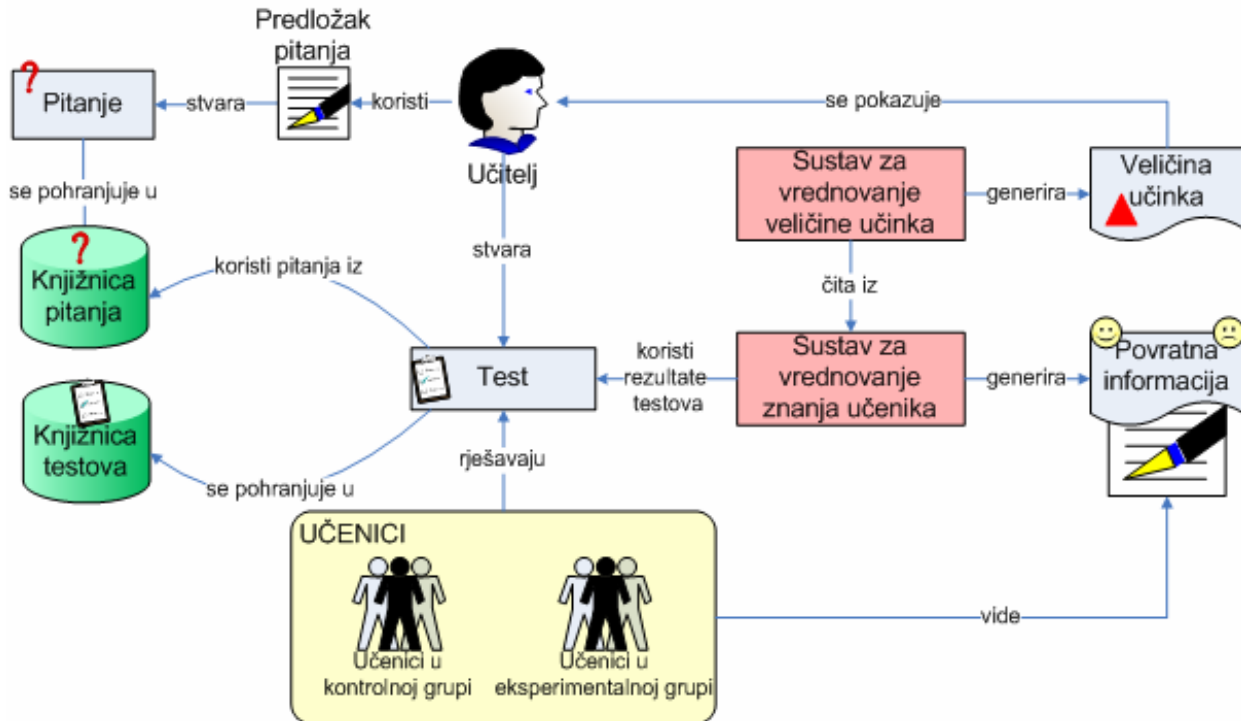
	Questionmark Perception	Question Writer	SmartLite XP	WebQuiz	Articulate Quizmaker
	[PERC2005]	[QUES2005]	[SMAR2005]		[ARTI2005]
Predloži za testove	+	-	-		-
Vrste pitanja					
jednostruki odabir	+	-	-		+
višestruki odabir	+	+	+		+
da/ne	+	+	+		+
povezivanje	+	+	-		+
redoslijed	+	-	-		+
odabir objekta	+	-	-		+
popunjavanje	+	+	+		+
esej	+	-	+		-
kratki odgovor	+	-	-		-
brojčani	+	-	-		+
Ponovna uporabljivost pitanja	+	-	-		-
Multimedija	+	+	+		+
Bodovanje	+	+	+		+
Rezultati	+	-	+		-
Čarobnjak	+/-	-	+		-
Pregled	+	+	+		+
HTML test	+	+	+		+
Planer	+	-	-		-
e-mail povratna informacija	-	+	+		+
Vrste izvještaja	+	-	-		-

Promatrani alati omogućavaju samo upravljanje testovima, tj. izradu, isporuku i ocjenjivanje testova, i ne mogu vrednovati učinak sustava e-učenja. Štoviše, njihove funkcionalnosti čine samo dio funkcionalnosti našeg prototipa programskog alata za vrednovanje učinka sustava e-učenja, jer je upravljanje testovima nužno, ali ne i dovoljno, za vrednovanje učinka sustava e-učenja.

5.2. FUNKCIONALNOST I ARHITEKTURA EVEDIN-A

Ka što je već spomenuto, naš je cilj bio napraviti prototip programskog alata za automatsko vrednovanje učinka sustava e-učenja. Metodologija vrednovanja koja je opisana u 4. poglavlju je implementirana i tvori središnji dio sustava EVEDIN.

EVEDIN omogućava upravljanje i izvođenje eksperimenata, tj. upravljanje kontrolnom i eksperimentalnom grupom, izradu i provođenje svih potrebnih testova, njihovo automatsko ispravljanje i ocjenjivanje, te, što je najvažnije, omogućava provjeru istinitosti nul-hipoteza, te izračun veličine učinka eksperimentalnog faktora (Slika 5. 1.).



Slika 5. 1. Funkcionalni model EVEDIN-a

EVEDIN se može promatrati sa gledišta njegovih sudionika: sa stajališta učitelja, te sa stajališta učenika. Okruženje za učitelja ima autorske značajke što znači da u tom autorskom okruženju učitelj stvara pitanja i testove, dok učenik isključivo pristupa testiranju znanja kroz pripremljene testove. Funkcionalnosti EVEDIN-a promatraju se zasebno kroz funkcionalnosti svakog od sudionika.

5.2.1. FUNKCIONALNOSTI UČITELJA

Učitelj u svom autorskom okruženju organizira eksperiment, tj. upravlja kontrolnom i eksperimentalnom grupom, testovima i pitanjima. Učitelj stvara pitanja i odgovore prema zadanim predlošcima. Učitelji, dakle, mogu:

- stvarati, mijenjati ili brisati svoje eksperimente
- dodavati, mijenjati ili brisati ispitanike
- pridruživati ispitanike u eksperiment

- stvarati, mijenjati ili brisati svoje testove
- pridruživati testove u eksperiment
- stvarati, mijenjati ili brisati svoja pitanja i odgovore koristeći predloške
- pridruživati pitanja u testove
- ručno unositi rezultate testiranja
- pregledavati rezultate svih testova koje su postigli svi ispitanici u pridruženom mu eksperimentu
- podijeliti ispitanike u kontrolnu i eksperimentalnu grupu
- provesti provjeravanje istinitosti nul-hipoteza
- izračunati veličinu učinka svog eksperimenta

UPRAVLJANJE EKSPERIMENTIMA

Učitelj može stvoriti novi eksperiment koji je definiran sljedećim atributima:

- naziv eksperimenta,
- naziv sustava koji se vrednuje,
- datum početka eksperimenta,
- datum završetka eksperimenta i
- autor eksperimenta (koji odgovara prijavljenom učitelju koji stvara eksperiment).

Eksperiment mora imati jedinstven naziv s obzirom na sustav koji se vrednuje. Učitelj može mijenjati ili brisati samo one eksperimente kojima je on autor. U slučaju brisanja eksperimenta, brišu se i svi pridruženi mu testovi istog autora, kao i svi pridruženi mu ispitanici, ukoliko nisu pridruženi nekim drugim eksperimentima.

UPRAVLJANJE ISPITANICIMA

Učitelj može dodavati nove ispitanike samo ako ima pridružen barem jedan eksperiment. Učitelj dodaje nove ispitanike koji su definirani sljedećim atributima:

- ime,
- prezime,
- korisničko ime (jedinstveno u cijelom sustavu EVEDIN) i
- lozinka.

Učitelj može u svoj eksperiment pridružiti i druge ispitanike koji pripadaju drugim eksperimentima. Učitelj može izbaciti iz eksperimenta bilo kojeg učenika.

UPRAVLJANJE TESTOVIMA

Učitelj može stvoriti novi test koji je definiran sljedećim atributima:

- naziv testa,
- datum početka testa,
- datum završetka testa,
- trajanje rješavanja testa u minutama i
- autor testa (koji odgovara prijavljenom učitelju koji stvara test).

Test mora imati jedinstven naziv s obzirom na autora. Učitelj može mijenjati ili brisati samo one testove kojima je on autor i koji nisu pridruženi drugim eksperimentima. U slučaju brisanja testa, brišu se i sva pridružena mu pitanja ukoliko nisu pridružena nekim drugim testovima, kao i svi rezultati odabranog testa.

UPRAVLJANJE PITANJIMA

Učitelj može dodavati nova pitanja samo ako ima pridružen barem jedan test. Učitelj može stvoriti novo pitanje koje je definirano sljedećim atributima:

- naziv pitanja,
- maksimalni broj bodova,
- tekst pitanja,
- vrsta i
- autor pitanja (koji odgovara prijavljenom učitelju koji stvara pitanje).

Automatski se pitanjima pridružuju i njihovi odgovori koji su definirani sljedećim atributima:

- identifikator pitanja,
- tekst odgovora i
- točnost.

Učitelj može u svoj test pridružiti i druga pitanja koji pripadaju drugim testovima. Učitelj može izbaciti iz testa bilo koje pitanje. Učitelj može izbrisati trajno samo ona pitanja kojima je autor i koja se ne koriste u drugim testovima.

Da bi se zadovoljile karakteristike valjanosti, pouzdanosti, osjetljivosti i diskriminativnosti pitanja učitelj mora voditi računa o uključivanju raznolikih pitanja. Budući se testiranje i ocjenjivanje obavljaju na računalu, isključena je mogućnost utjecaja bilo kakvih subjektivnih faktora na rezultat mjerenja, pa sigurno elektronički testovi imaju karakteristiku objektivnosti. Karakteristika baždarenosti se razvojem WWW tehnologije i Interneta sve lakše postiže jer se povećava mogućnost uspoređivanja rezultata pojedinih učenika sa širom skupinom, na primjer, sa svim učenicima iste razine obrazovanja u cijeloj zemlji.

Učitelj može stvoriti različita pitanja uporabom predložaka. Predlošci se uvode kao prijedlozi za pisanje pitanja i odgovora pri konstrukciji testova znanja. Svaki predložak ima svoje specifičnosti po kojima se razlikuje od ostalih i prema kojima će se sastavljati pitanja različitih tipova. Predlošci omogućavaju brzo i jednostavno pisanje pitanja i odgovora. Problem je što oduzimaju stvaralačku slobodu te često ne odgovaraju svim zahtjevima učitelja. Vrste pitanja koja učitelj može definirati u svom okruženju su sljedeće:

1. Jednostruki izbor (eng. Single choice)

Na ova pitanja je potrebno odgovoriti biranjem jedne od ponuđene dvije opcije od kojih je samo jedna točan odgovor. Ona su prikladna u slučajevima testiranja velike količine gradiva, jer se njima skraćuje vrijeme potrebno za testiranje, ali ona ne odražavaju u potpunosti znanje učenika jer je moguće pogadati odgovore uz 50% šanse za uspjeh.

2. Višestruki izbor (eng. Multiple choice)

Na ova pitanja odgovara se biranjem nekoliko od više ponuđenih odgovora od kojih su neka točna. Za razliku od pitanja jednostrukog izbora, ova pitanja su bolji pokazatelj znanja učenika jer se povećanjem broja odgovora smanjuje mogućnost pogađanja.

3. Da/Ne (eng. Yes/No)

Jednom od dvije ponuđene opcije Da ili Ne izabire se odgovor. Ne preporuča se da sadrže negacijske, potvrdne i druge riječi koje bi mogle sugerirati točnost odgovora. Pogodna su za ispitivanje stupnja usvojenosti činjeničnog znanja.

4. Povezivanje (eng. Matching)

Ova pitanja omogućuju ispitivanje znanja učenika kroz sparivanje danih pojmova. Posebno su korisni za testiranje onih područja znanja u kojima postoji veliki broj pojmova i činjenica. Ponuđene opcije, u obliku riječi, učenik treba upariti s nekom drugom riječi

5. Dopunjavanje (eng. Accessary)

Potrebno je dopuniti riječi koje nedostaju u pitanju odnosno, rečenici. Dopunjavanje se vrši odabirom nekog od ponuđenih odgovora, pa predstavlja varijaciju višestrukog izbora.

6. Redoslijed (eng. Sequence)

Svaki se element ponuđenog niza pojmova numerira jednim brojem, tako da na kraju dobijemo niz istih pojmova samo promijenjenog redoslijeda.

7. Brojčani odgovor (eng. Numerical)

Na pitanje se odgovara upisivanjem neke brojčane vrijednosti.

Pri stvaranju pitanja učitelj ponavlja proizvoljan broj puta sljedeće korake: izbor predloška za pitanje, unos teksta pitanja, unos teksta odgovora na pitanje (osim ako su pitanja tipa DA/NE), izbor točnog(ih) odgovora, te određivanje maksimalnog broja bodova koje učenik može postići odgovaranjem na pitanje.

UPRAVLJANJE REZULTATIMA TESTOVA

Učitelj može pregledati rezultate koje su ispitanici u odabranom eksperimentu postigli na testovima koji su mu pridruženi.

Zbog organizacijskih poteškoća koje su vrlo vjerojatne prilikom realizacije eksperimenta kada zbog nedostatka prostora nije moguće provesti testiranje učenika u okviru EVEDIN-a, učitelju je ostavljena mogućnost ručnog unošenja rezultata testova. U tu svrhu učitelj mora imati stvoren eksperiment i pridružene mu testove, kao i pridružene ispitanike. Tako uneseni rezultati se u sustavu označavaju kao „ručno uneseni“ i ne prikazuju se učeniku u popisu riješenih testova.

UPRAVLJANJE KONTROLNOM I EKSPERIMENTALNOM GRUPOM

Nakon provedenog inicijalnog testiranja, učitelj može pristupiti definiranju kontrolne i eksperimentalne grupe. U tu svrhu na raspolaganju su mu ručno i automatsko definiranje eksperimentalnih grupa. Preporuča se prvo pokrenuti automatsko izjednačavanje grupa koje radi na principu izjednačavanja po rasponu (opisanog u prethodnom poglavlju) uz proizvoljno odabrani raspon bodova. Zatim učitelj može prijeći na ručno izjednačavanje grupa. U bilo kojem trenutku učitelj može provjeriti ekvivalentnost definiranih grupa. Kada učitelj dobije ekvivalentne grupe koje odgovaraju njegovim zahtjevima i koje su ekvivalentne, on evidentira postojanje te ekvivalentnosti.

ISPITIVANJE NUL-HIPOTEZA I IZRAČUN VELIČINE UČINKA

Ispitivanje istinitosti nul-hipoteza može se provesti samo za one eksperimente koji su završili i koji imaju evidentiranu ekvivalentnost kontrolne i eksperimentalne grupe. Tada se definiraju podgrupe kontrolne i eksperimentalne grupe kojima pripadaju samo oni ispitanici koji su riješili sve testove pridružene određenom eksperimentu. Zatim učitelj može provjeriti ekvivalentnost tako dobivenih grupa. Ako su podgrupe ekvivalentne, pristupa se ispitivanju istinitosti nul-hipoteza i računanju veličine učinka.

5.2.2. FUNKCIONALNOSTI UČENIKA

Učenik pristupa procesu testiranja znanja prema definiranim testovima koji su mu na raspolaganju (ovisno o pripadnosti eksperimentu i o vremenskoj raspoloživosti testa), a nakon završetka istog procesa dobiva uvid u povratnu informaciju o rezultatima koje je postigao s ispisom bodova. Učenici pristupaju rješavanju pitanja onim redoslijedom kojim žele, pošto se sva pitanja prikazuju odjednom. Ovo je napravljeno zbog toga da se što više simulira rješavanje testa na papiru, čime se povećava uspjeh i smanjuje nervoza, a osim toga osigurana je i mogućnost prilagođavanja rješavanja testa samom učeniku. Testiranje završava ili kada učenik sam odabere završetak ili kad prođe za to predviđeno vrijeme. Ako se definira trajanje testiranja, tada se približavanje isteka vremena prikazuje učeniku. Sustav automatski ispravlja test, ocjenjuje ga te pokaže rezultate učeniku. Učenikov rezultat se pohranjuje u odgovarajuću bazu podataka da bi se mogao koristiti u daljnjoj analizi i pri računanju veličine učinka.

Učenik odgovaranjem na pitanja prikuplja bodove za svako pitanje na sljedeći način:

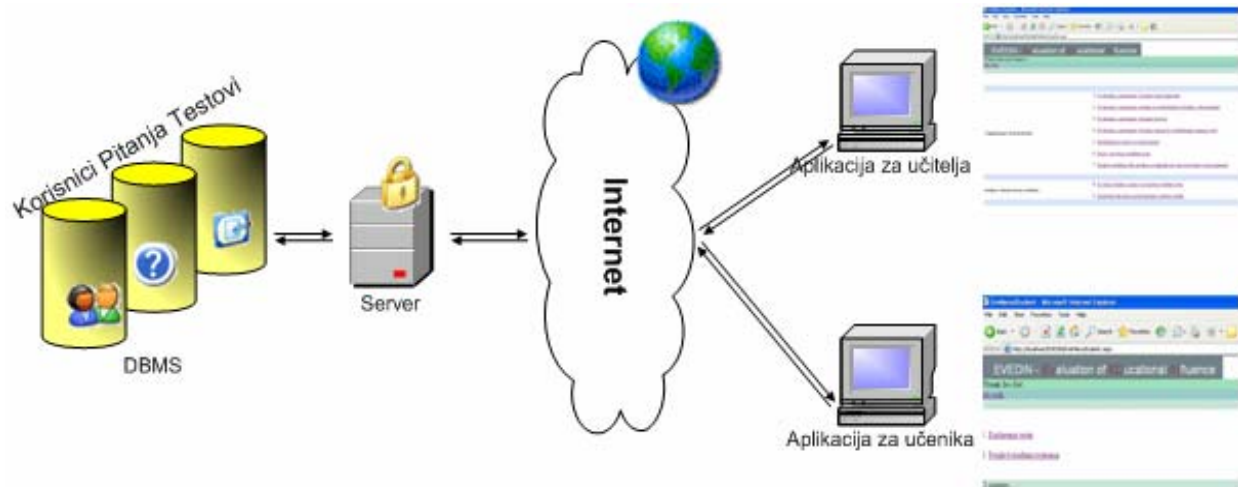
- ako odgovori na pitanje u potpunosti točno postiže maksimalni broj bodova pitanja
- ako ne odgovori na pitanje pridružuje mu se 0 bodova
- ako odgovori djelomično točno, odnosno x točnih od y mogućih točnih odgovora, postiže $(x/y) * (\text{maksimalni broj bodova pitanja})$ bodova

EVEDIN učitelju omogućava uštedu vremena kod izrade testova, a posebno kod ispravljanja testova, jer automatski ispravlja testove te prezentira učeniku i učitelju rezultate. Učitelji mogu koristiti ovaj alat za štampanje izvještaja o rezultatima testiranja. Najvažnija značajka ovog alata je njegova mogućnost izračuna veličine učinka eksperimenta primjenom opisane metodologije.

5.2.3. STRUKTURA PROGRAMSKOG RJEŠENJA

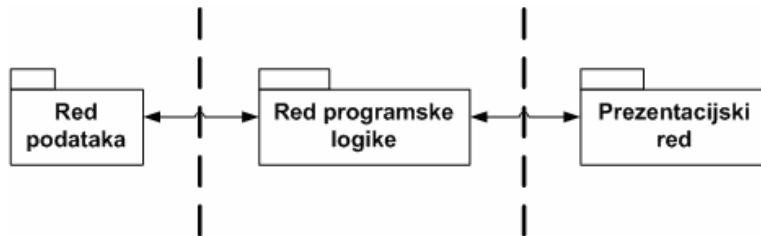
Sustav EVEDIN realiziran je pomoću troredne arhitekture s izdvojenim redovima podataka, aplikacijske logike i korisničkog sučelja. EVEDIN čine tri neovisne komponente (Slika 5. 2.): sustav za upravljanje bazom podataka (eng. database management system - DBMS) – red podataka, server aplikacija – red programske logike, klijent aplikacija – prezentacijski red.

DBMS pohranjuje eksperimente, korisnike, pitanja, testove i rezultate u SQL bazi, ASP.NET aplikacija predstavlja sučelje između klijenta i DBMS-a, Internet web server omogućava pristup HTML stranicama, a klijent aplikacija se sastoji od dvije aplikacije: aplikacija za učitelja i aplikacija za učenika. Aplikacija za učitelja služi za organiziranje eksperimenta: ona omogućava upravljanje kontrolnom i eksperimentalnom grupom, pitanjima i testovima.



Slika 5. 2. Arhitektura EVEDIN-a

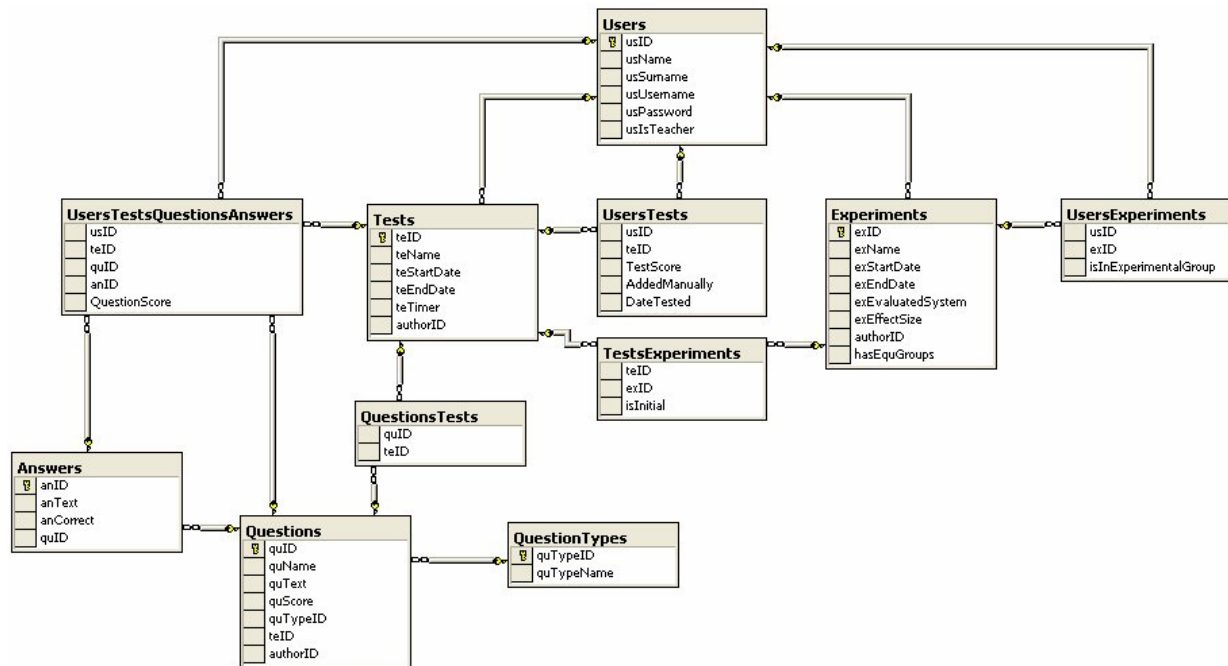
Dijagramom na Slika 5. 3. je opisana troredna arhitektura EVEDIN-a gdje je svaki red predstavljen paketom. Red aplikacijske logike kao i red korisničkog sučelja sustava implementiran je u Microsoftovom .NET okruženju. Ova dva reda sustava se funkcionalno dijele prema funkcionalnostima namijenjenim pojedinim sudionicima. Komunikacija reda aplikacijske logike i reda podataka odvija se razmjenom podataka korištenjem protokola SOAP (Simple Object Access Protocol) koji se zasniva na prijenosu poruka strukturiranih XML-om (eXtensible Markup Language), dok je komunikacija s redom podataka određena internim podešavanjima.



Slika 5. 3. EVEDIN – dijagram komponenti

5.2.3.1 RED PODATAKA

Red podataka implementiran je unutar Microsoftovog sustava za upravljanje relacijskim bazama podataka SQL Server. Red podataka čini baza podataka koja sadrži podatke o svim sudionicima sustava, svim eksperimentima, pripadajućim im testovima, pitanjima i odgovorima. U ovoj bazi podataka se uz osnovne informaciju o sudioniku (ime, prezime, korisničko ime, zaporka) nalazi i podatak o pripadnosti sudionika eksperimentalnoj ili kontrolnoj grupi u nekom eksperimentu, o tome da li je sudionik učitelj ili učenik, pa samim time i o ograničenjima u korištenju pojedinih funkcija sustava (Slika 5. 4.).



Slika 5. 4. Struktura baze podataka

XML shema baze:

```

<xs:element name="Users" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="usID" msdata:ReadOnly="true" msdata:AutoIncrement="true" type="xs:int" />
  <xs:element name="usName" type="xs:string" />
  <xs:element name="usSurname" type="xs:string" />
  <xs:element name="usUsername" type="xs:string" />
  <xs:element name="usPassword" type="xs:string" />
  <xs:element name="usIsTeacher" type="xs:boolean" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="Tests" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="teID" msdata:ReadOnly="true" msdata:AutoIncrement="true" type="xs:int" />
  <xs:element name="teName" type="xs:string" />
  <xs:element name="teStartDate" type="xs:dateTime" />
  <xs:element name="teEndDate" type="xs:dateTime" />
  <xs:element name="teTimer" type="xs:int" minOccurs="0" />
  <xs:element name="authorID" type="xs:int" />
</xs:sequence></xs:complexType></xs:element>
  
```

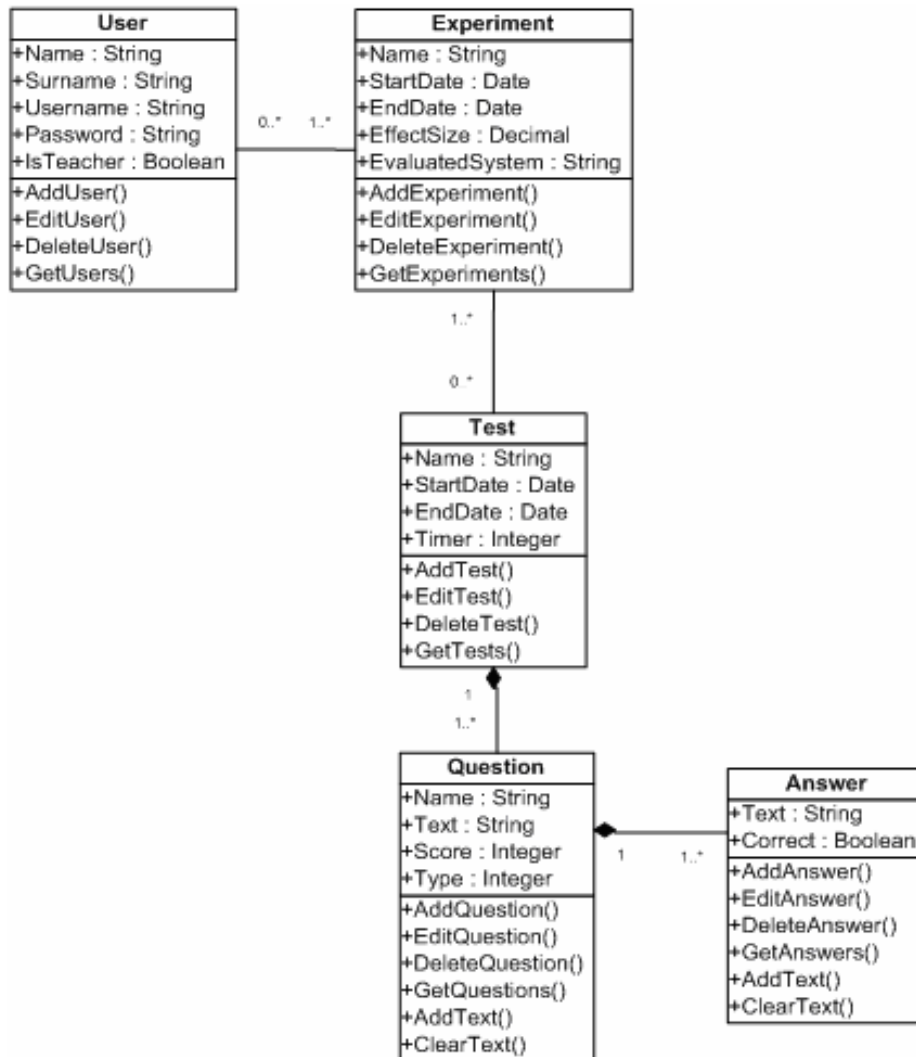


```

<xs:element name="Questions" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="quID" msdata:ReadOnly="true" msdata:AutoIncrement="true" type="xs:int" />
  <xs:element name="quName" type="xs:string" />
  <xs:element name="quText" type="xs:string" />
  <xs:element name="quScore" type="xs:int" />
  <xs:element name="quTypeID" type="xs:int" />
  <xs:element name="authorID" type="xs:int" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="UsersTestsQuestionsAnswers" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="usID" type="xs:int" />
  <xs:element name="teID" type="xs:int" />
  <xs:element name="quID" type="xs:int" />
  <xs:element name="anID" type="xs:int" />
  <xs:element name="QuestionScore" type="xs:float" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="Experiments" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="exID" msdata:ReadOnly="true" msdata:AutoIncrement="true" type="xs:int" />
  <xs:element name="exName" type="xs:string" />
  <xs:element name="exStartDate" type="xs:dateTime" minOccurs="0" />
  <xs:element name="exEndDate" type="xs:dateTime" minOccurs="0" />
  <xs:element name="exEvaluatedSystem" type="xs:string" />
  <xs:element name="exEffectSize" type="xs:float" minOccurs="0" />
  <xs:element name="authorID" type="xs:int" />
  <xs:element name="hasEquGroups" type="xs:boolean" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="UsersTests" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="usID" type="xs:int" />
  <xs:element name="teID" type="xs:int" />
  <xs:element name="TestScore" type="xs:int" minOccurs="0" />
  <xs:element name="AddedManually" type="xs:boolean" minOccurs="0" />
  <xs:element name="DateTested" type="xs:dateTime" minOccurs="0" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="TestsExperiments" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="teID" type="xs:int" />
  <xs:element name="exID" type="xs:int" />
  <xs:element name="isInitial" type="xs:boolean" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="QuestionsTests" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="quID" type="xs:int" />
  <xs:element name="teID" type="xs:int" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="UsersExperiments" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="usID" type="xs:int" />
  <xs:element name="exID" type="xs:int" />
  <xs:element name="isInExperimentalGroup" type="xs:int" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="QuestionTypes" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="quTypeID" type="xs:int" />
  <xs:element name="quTypeName" type="xs:string" />
</xs:sequence></xs:complexType></xs:element>
<xs:element name="Answers" xmlns:xs="http://www.w3.org/2001/XMLSchema"><xs:complexType><xs:sequence>
  <xs:element name="anID" msdata:ReadOnly="true" msdata:AutoIncrement="true" type="xs:int" />
  <xs:element name="anText" type="xs:string" />
  <xs:element name="anCorrect" type="xs:boolean" />
  <xs:element name="quID" type="xs:int" />
</xs:sequence></xs:complexType></xs:element>
    
```

Osnovne klase paketa koji predstavlja red podataka su **Experiment**, **User**, **Test**, **Question** i **Answer**, koje svojim metodama implementiraju dodavanje, mijenjanje i brisanje istoimenih elemenata: eksperimenata (**AddExperiment()**, **EditExperiment()**, **DeleteExperiment()**), korisnika (**AddUser()**, **EditUser()**, **DeleteUser()**), testova (**AddTest()**, **EditTest()**, **DeleteTest()**), pitanja (**AddQuestion()**, **EditQuestion()**, **DeleteQuestion()**), te odgovora (**AddAnswer()**, **EditAnswer()**, **DeleteAnswer()**).

Dodavanje i brisanje pitanja i odgovora testa se vrši preko klasa `Question` i `Answer`. Pitanje i odgovor se sastoji od teksta, stoga ove dvije klase sadrže metodu `AddText()` za dodavanje teksta pitanjima i odgovorima.

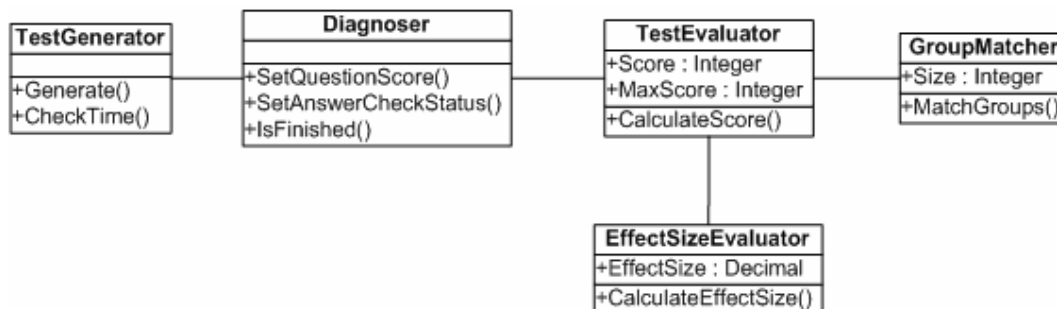


Slika 5. 5. EVEDIN – red podataka

Ostale klase paketa koji predstavlja red podataka su `UsersExperiments`, `TestsExperiments`, `UsersTests`, `QuestionsTests`, `UsersTestsQuestionsAnswers` koje svojim metodama implementiraju dodavanje i brisanje redom ispitanika u eksperiment, testova u eksperiment, rezultata koje je neki ispitanik postignuo na testu, zatim dodavanje i brisanje pitanja u test, te na kraju dodavanje odgovora na pitanja u testu nekog ispitanika.

5.2.3.2 RED PROGRAMSKE LOGIKE

Klasama paketa reda podataka pristupaju klase iz reda programske logike koje čitaju elemente baze znanja te čitaju i zapisuju osnovne elemente eksperimenta. Klase paketa koji predstavlja red programske logike su prikazane na Slika 5. 6. i mogu se podijeliti globalno na klase odgovorne za postavljanje testa i na klase kojima se vrši generiranje pitanja i odgovora, testiranje i ocjenjivanje.



Slika 5. 6. EVEDIN – red programske logike

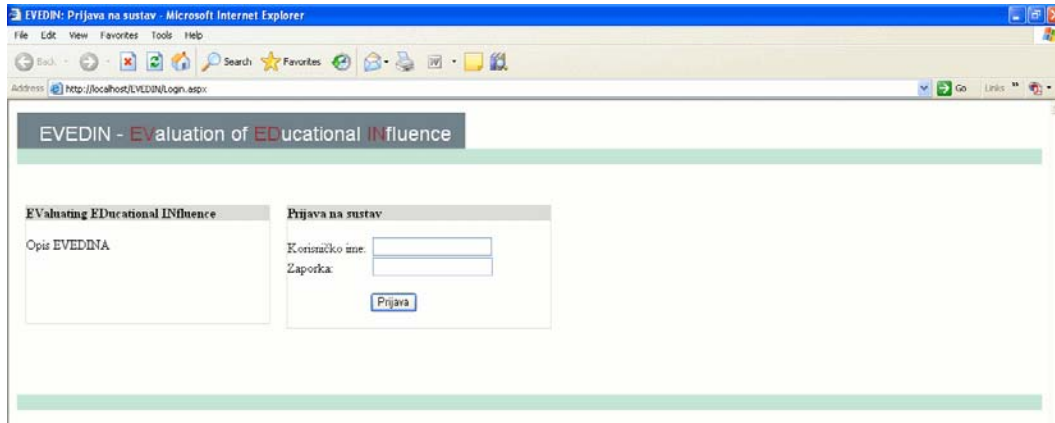
Klasa **Generator** generira test. Klasa **Diagnoser** služi za dijagnosticiranje stanja učenikova znanja, odnosno sadrži metode čije izvođenje, određuje da li je došlo do kraja kviza (**IsFinished()**). Nakon završenog testa slijedi izračunavanje ukupnog broja bodova pomoću metode **CalculateScore()** klase **TestEvaluator**. Izjednačavanje eksperimentalne i kontrolne grupe na temelju rezultata inicijalnog testa vrši metoda **MatchGroups()** klase **GroupMatcher**. Izračunavanje veličine učinka nekog sustava izvršava metoda **CalculateEffectSize()** klase **EffectSizeEvaluator**.

5.2.3.3 PREZENTACIJSKI RED

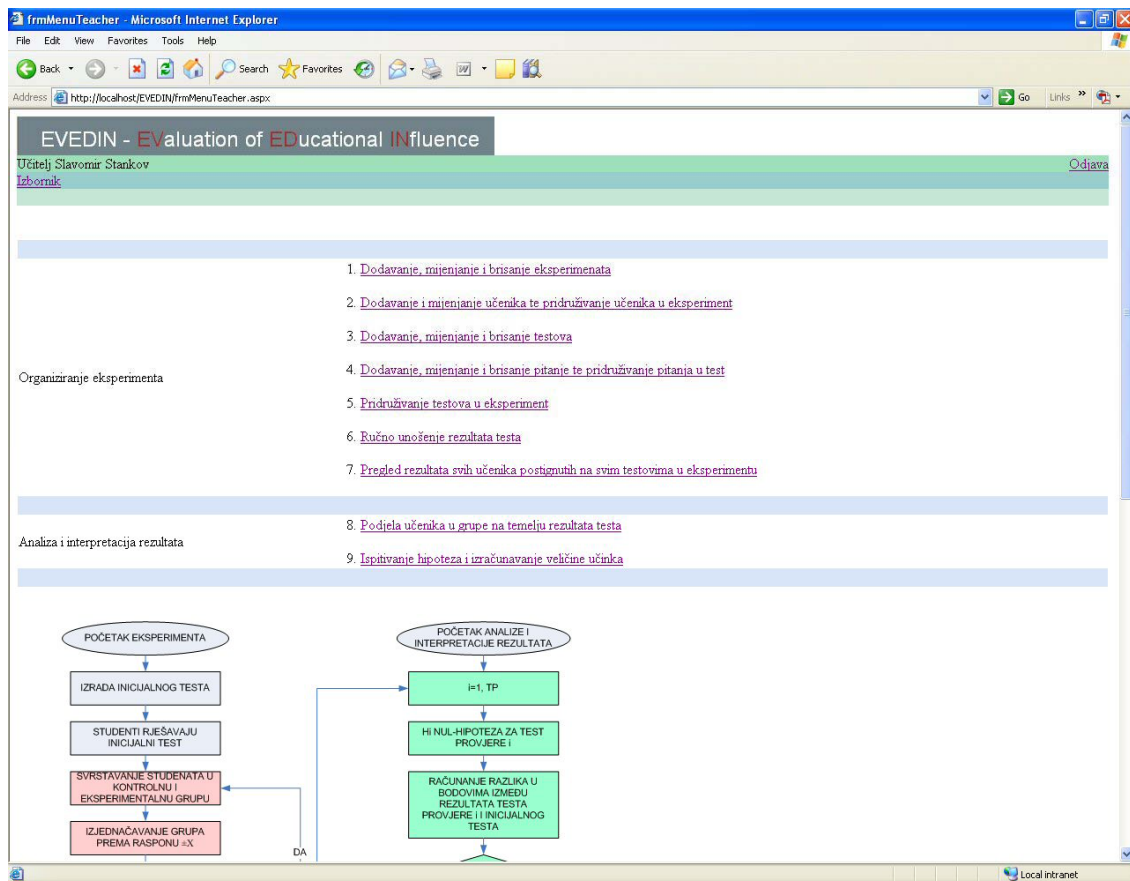
Prezentacijski red čini nekoliko Web forme i Web kontrola koje su implementirane ASP.NET tehnologijom:

1. forma za prijavu na sustav (Slika 5. 7)
2. početna forma za učitelja (Slika 5. 8)
3. forma za dodavanje i brisanje eksperimenta, te mijenjanje postavki eksperimenta (Slika 5. 9)
4. forma za dodavanje, mijenjanje i brisanje učenika, te pridruživanje učenika u eksperiment (Slika 5. 10)
5. forma za dodavanje, mijenjanje i brisanje testa (Slika 5. 11)
6. forma za pridruživanje testova u eksperiment (Slika 5. 12)
7. forma za dodavanje, mijenjanje i brisanje pitanja, te pregled i brisanje pitanja u testu (Slika 5. 13)
8. forma za pregled testa (Slika 5. 14)
9. forma za ručno unošenje rezultata testa (Slika 5. 15)
10. forma za ukupni pregled rezultata testa za odabrani eksperiment (Slika 5. 16)
11. forma za izjednačavanje grupa (Slika 5. 17)

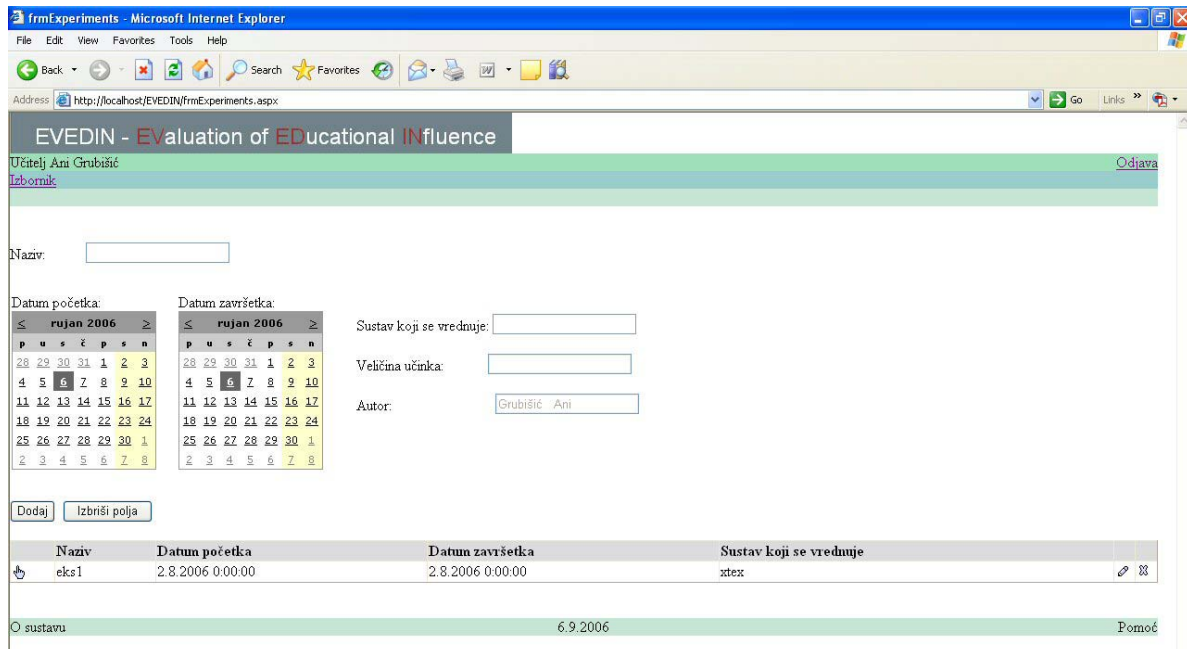
12. forma za ispitivanje nul-hipoteza i vrednovanje učinka (Slika 5. 18)
13. početna forma za učenika (Slika 5. 19)
14. forma za rješavanje testa (Slika 5. 20)
15. forma za pregled rezultata pojedinog testa za jednog učenika (Slika 5. 21)
16. forma za odjavu iz sustava (Slika 5. 22)



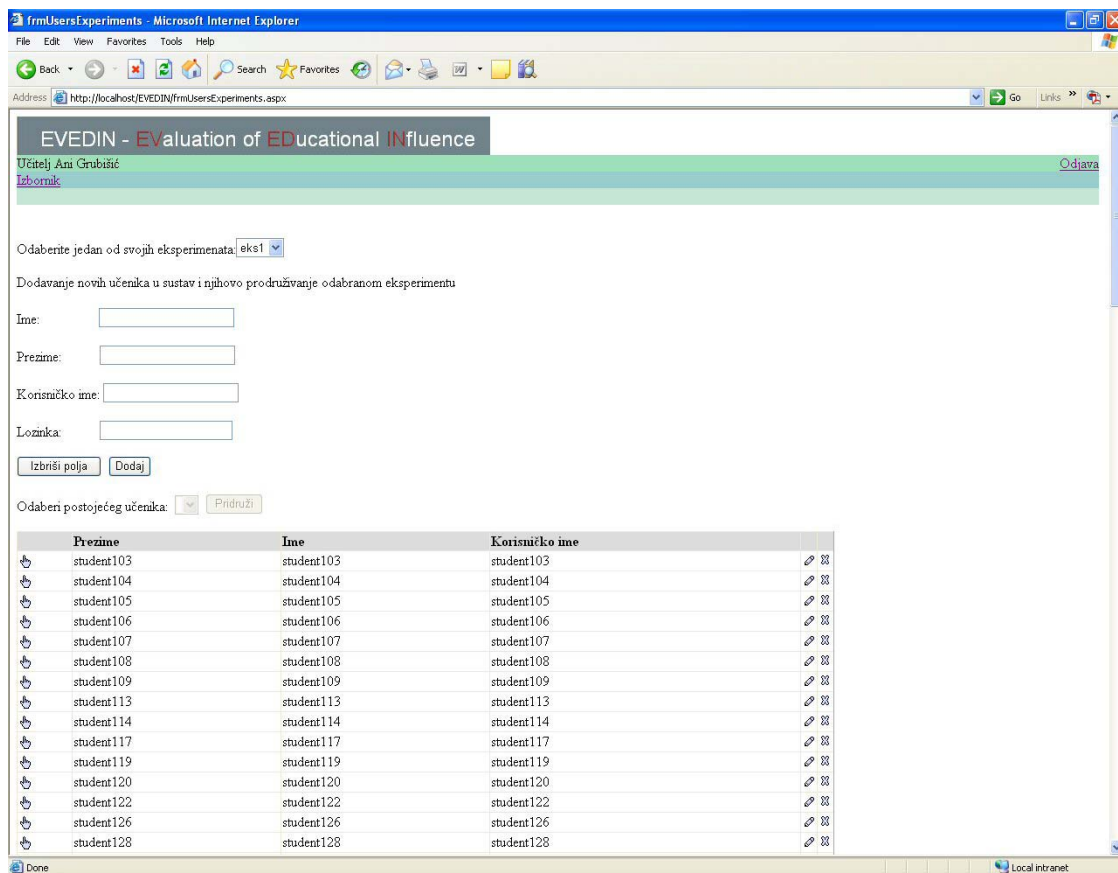
Slika 5. 7. Forma za prijavu na sustav



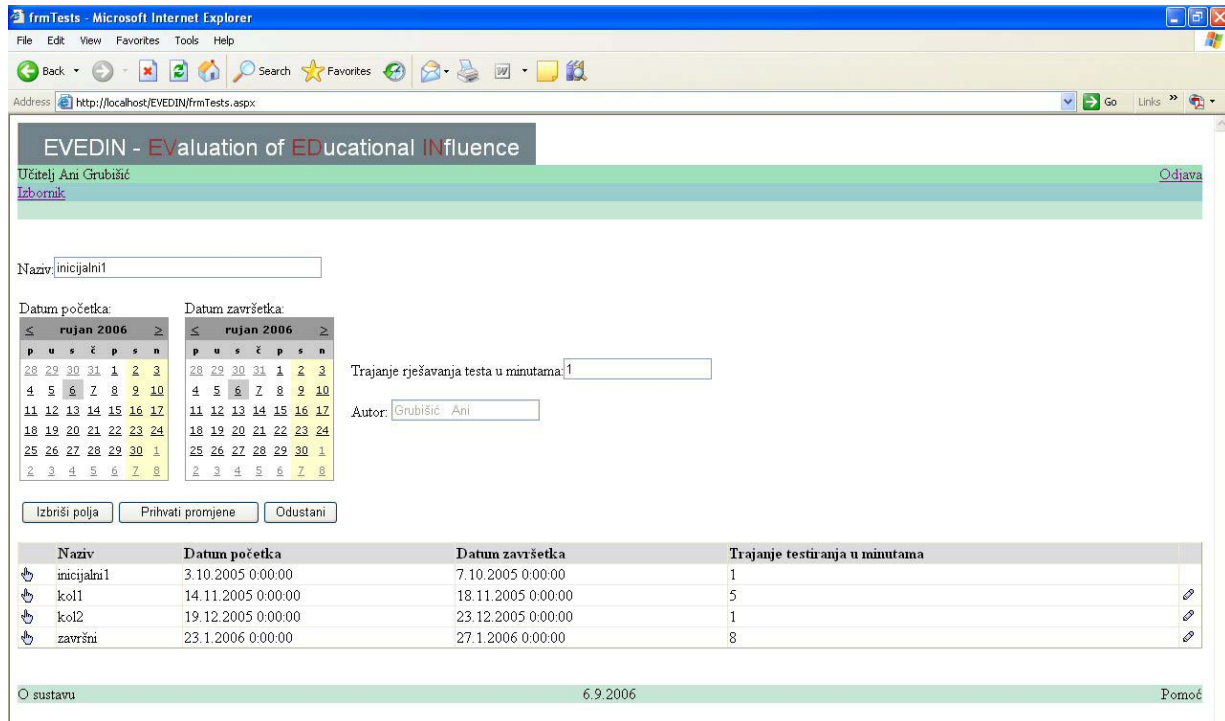
Slika 5. 8. Početna forma za učitelja



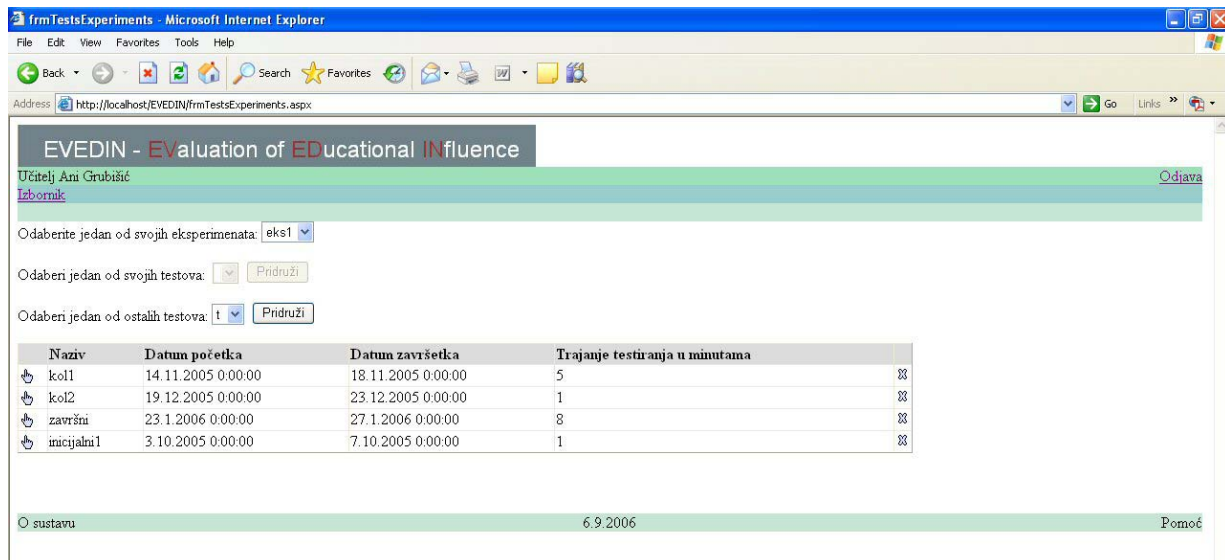
Slika 5. 9. Forma za dodavanje i brisanje eksperimenta, te mijenjanje postavki eksperimenta



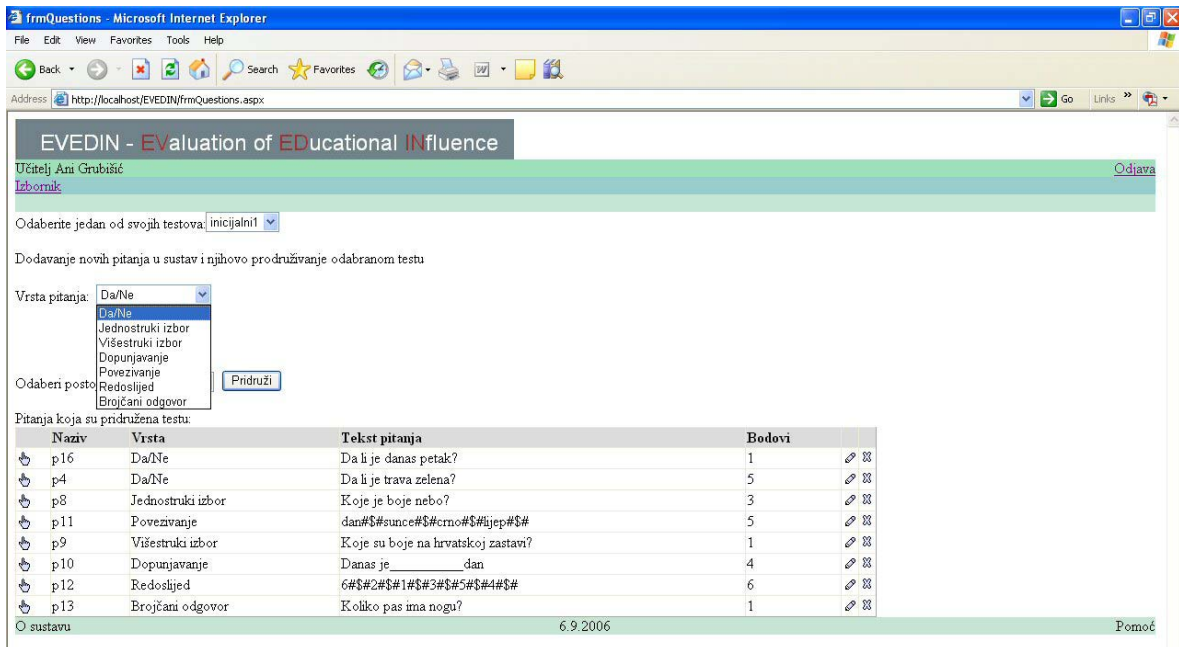
Slika 5. 10. Forma za dodavanje, mijenjanje i brisanje učenika, te pridruživanje učenika u eksperiment



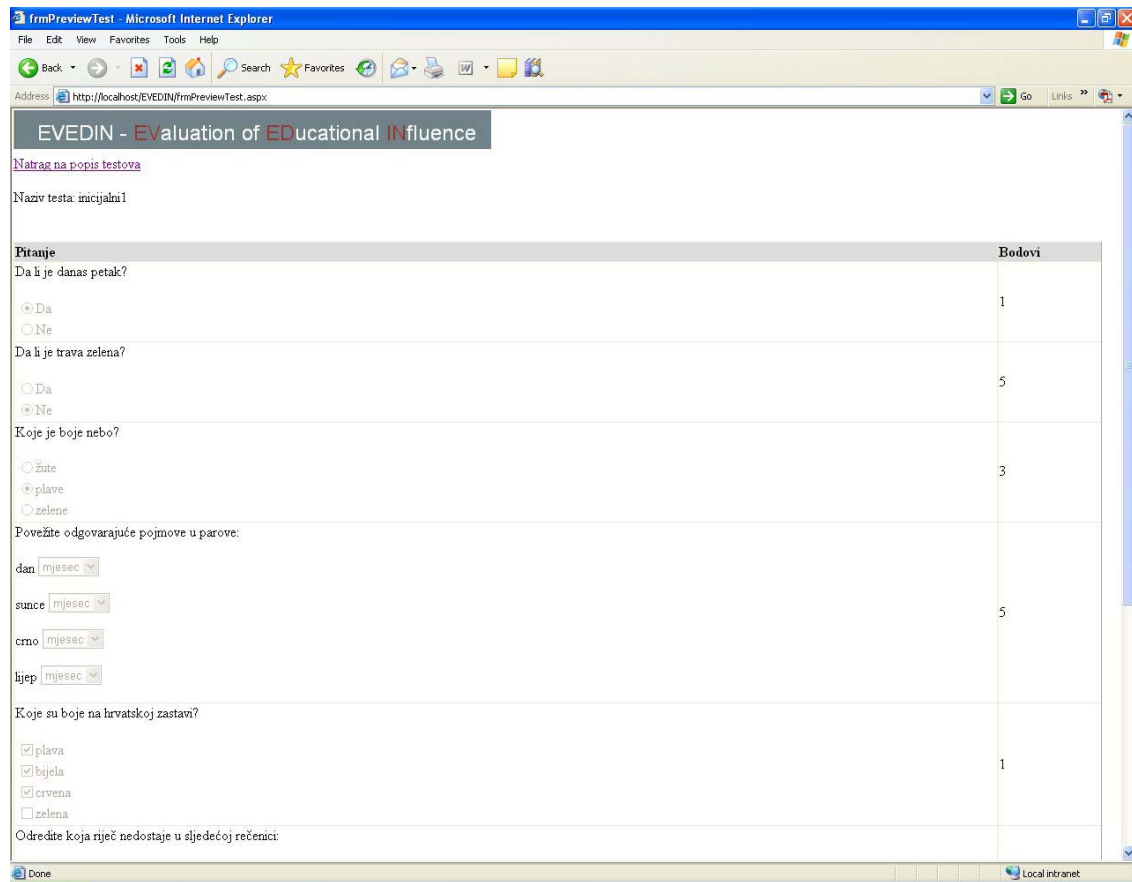
Slika 5. 11. Forma za dodavanje, mijenjanje i brisanje testa



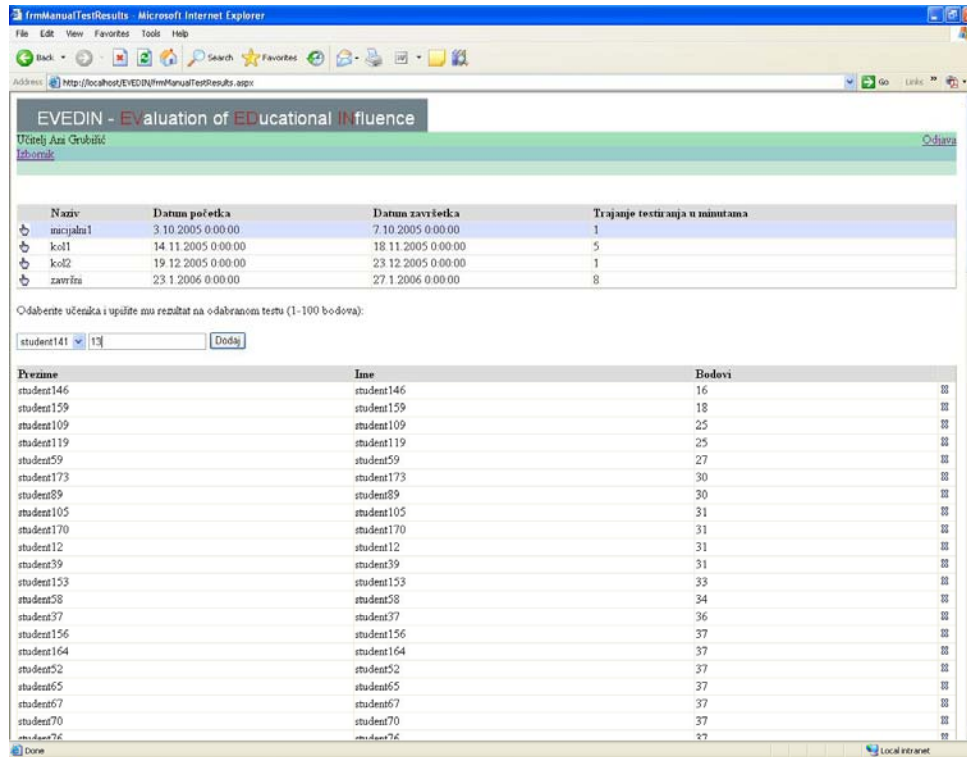
Slika 5. 12. Forma za pridruživanje testova u eksperiment



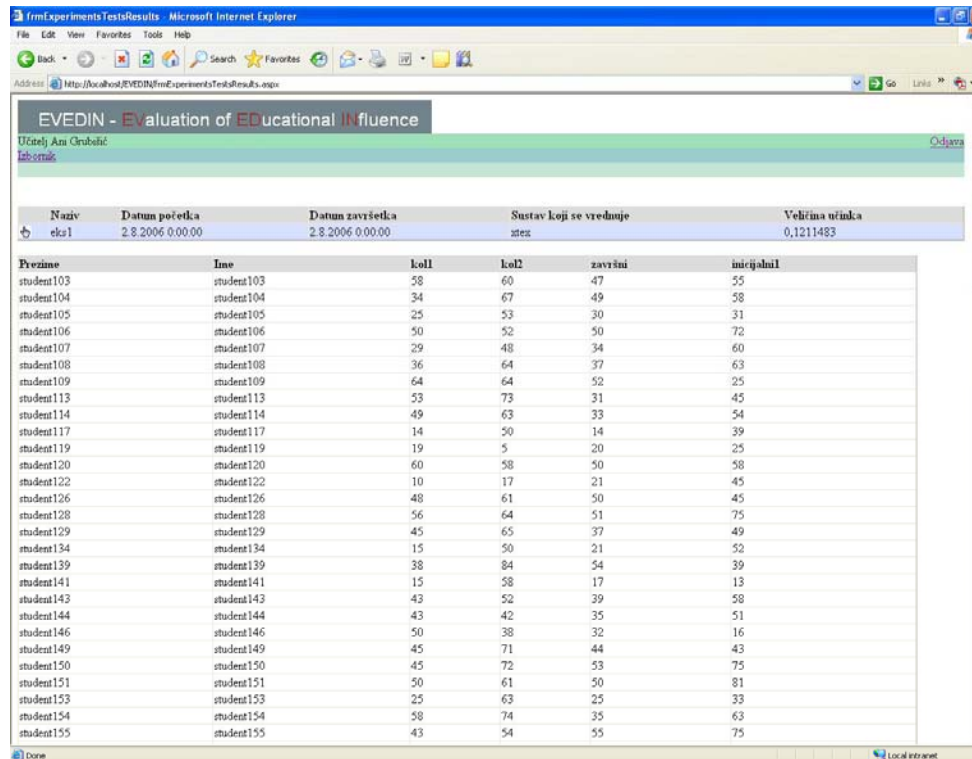
Slika 5. 13. Forma za dodavanje, mijenjanje i brisanje pitanja, te pregled i brisanje pitanja u testu



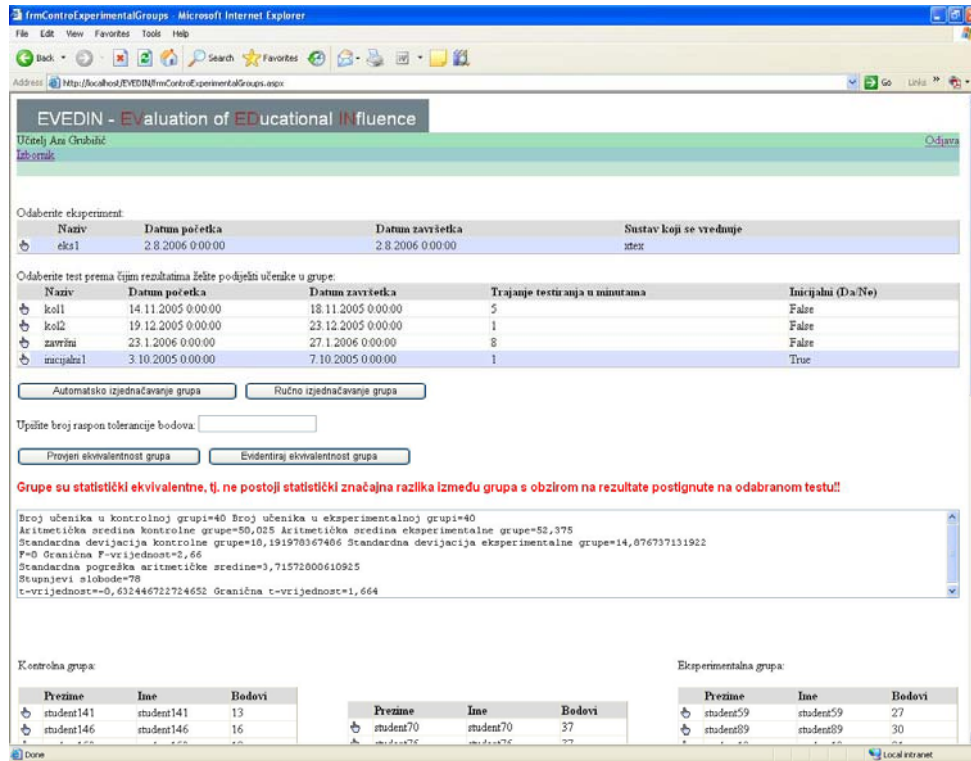
Slika 5. 14. Forma za pregled testa



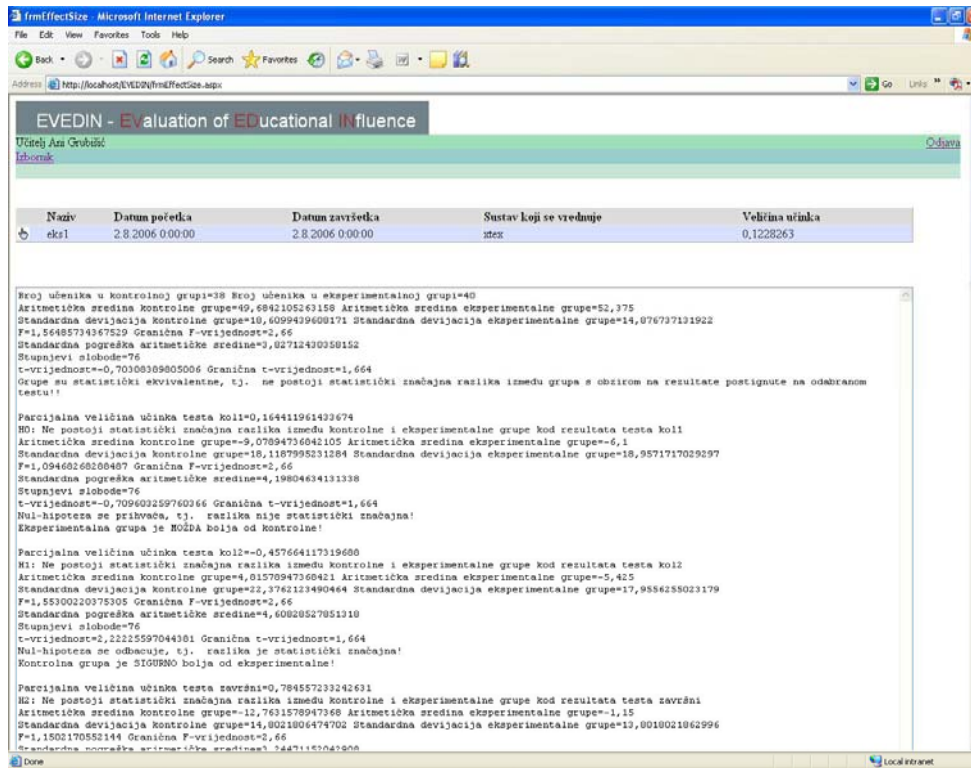
Slika 5. 15. Forma za ručno unošenje rezultata testa



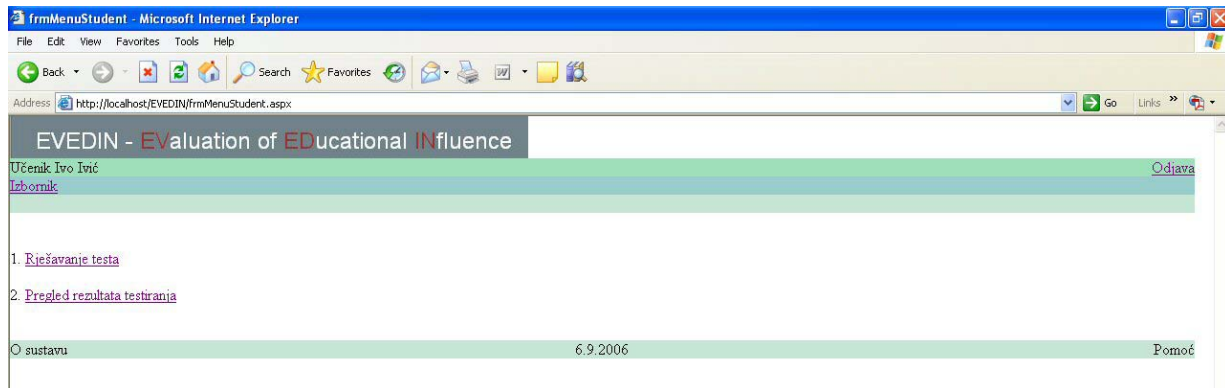
Slika 5. 16. Forma za ukupni pregled rezultata testa za odabrani eksperiment



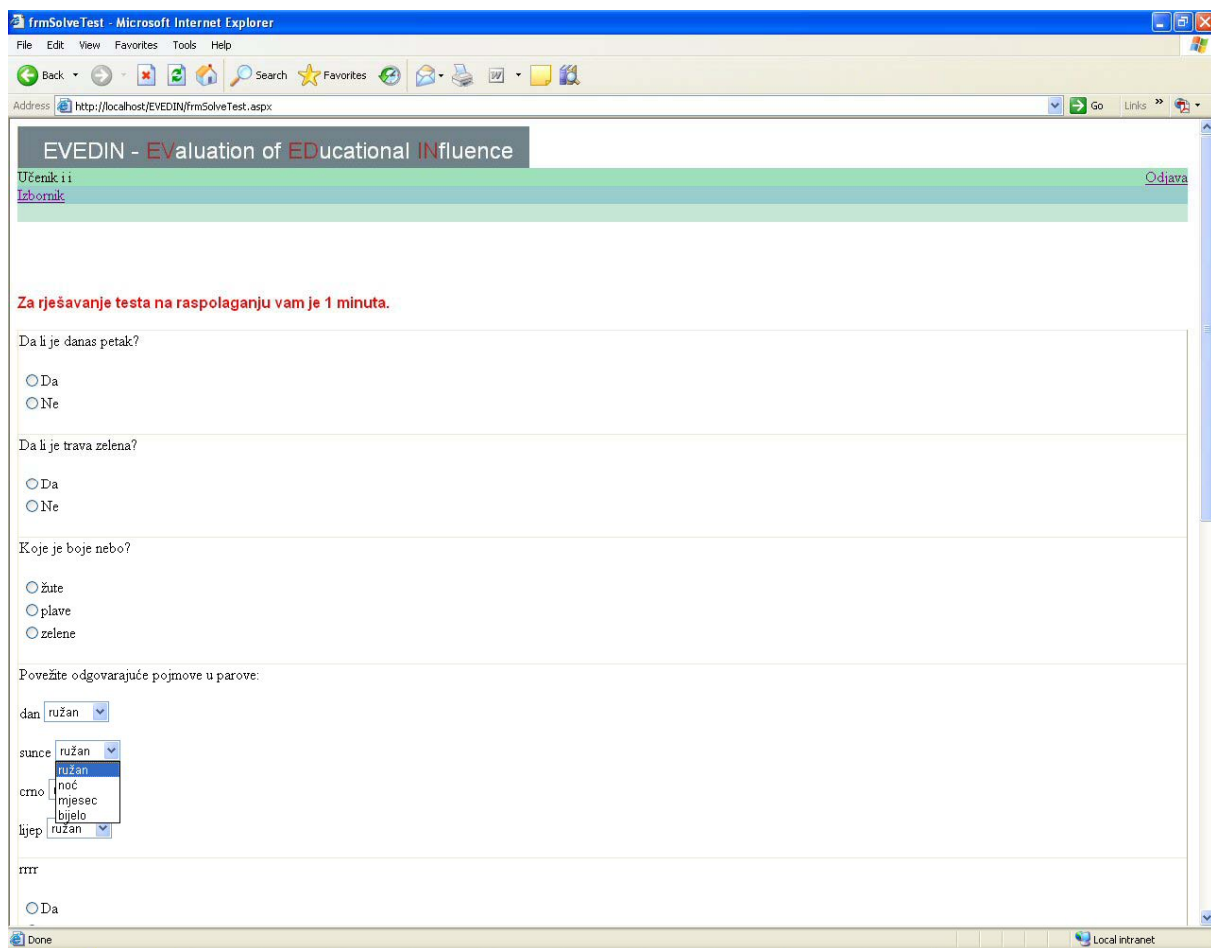
Slika 5. 17. Forma za izjednačavanje grupa



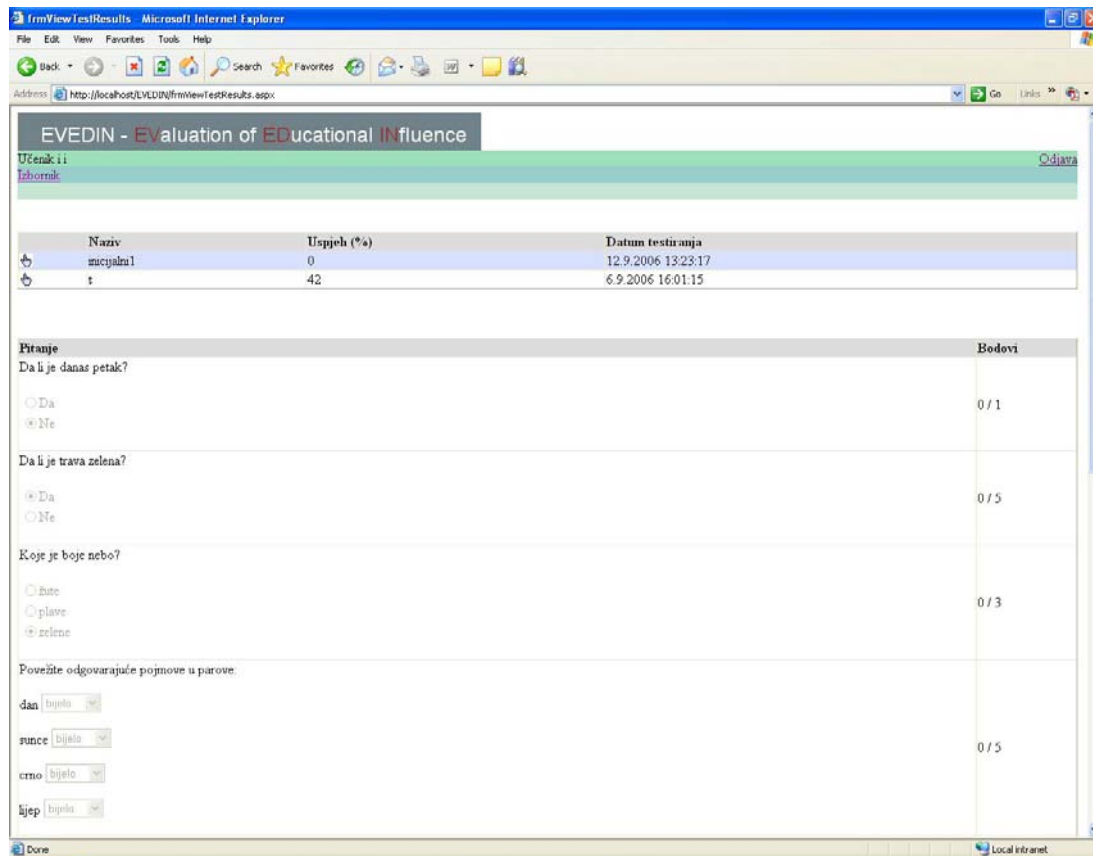
Slika 5. 18. Forma za ispitivanje nul-hipoteza i vrednovanje učinka



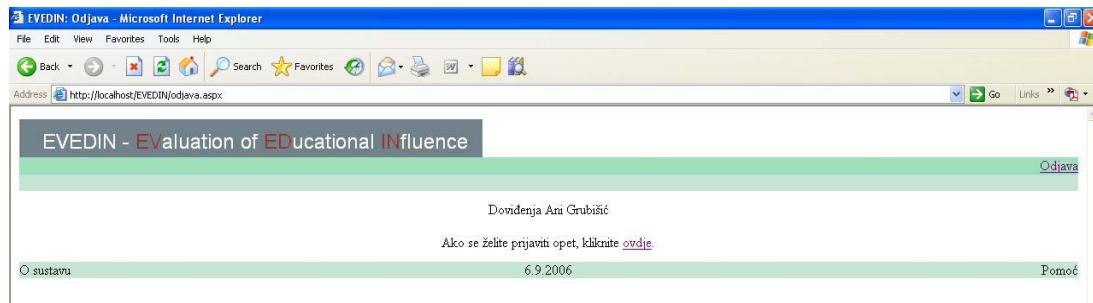
Slika 5. 19. Početna forma za učenika



Slika 5. 20. Forma za rješavanje testa



Slika 5. 21. Forma za pregled rezultata pojedinog testa za jednog učenika



Slika 5. 22. Forma za odjavu iz sustava

Sustavu EVEDIN se može pristupiti bilo kada s bilo kojeg mjesta koje ima pristup Internetu. On omogućava oblikovanje različitih vrsta pitanja, trenutno ispravljanje i ocjenjivanje testova, pregled izgleda cijelog testa, planiranje vremena za test. Na taj način učenici mogu odmah po završetku testa dobiti povratnu informaciju o svom postignuću. Učenik mora biti svjestan svoje uspješnosti i napredovanja i preuzeti odgovornost za svoje učenje. Sustav može poslati povratnu informaciju o rezultatima testa putem elektronske pošte učenicima i učiteljima. Povrh svega, on omogućava automatsko vrednovanje učinka sustava e-učenja.

Da bi primjena i provedba metodologije za vrednovanje učinka, definirane u prethodnom poglavlju, bila što jednostavnija, pristupili smo oblikovanju i implementaciji sustava za automatsko vrednovanje učinka sustava e-učenja. Ovaj sustav omogućava izradu i provedbu elektroničkih testova, čime se eliminira uporaba papirnatih testova, što podupire ekološku svijest, olakšava ispravljanje i ocjenjivanje testova. Rezultati tih elektroničkih testova koriste se kao ulazni podaci mehanizma za vrednovanje učinka, koji ih obrađuje i izračunava veličinu učinka sustava e-učenja.

Očekujemo da će primjena EVEDIN-a, sustava za automatsko vrednovanje učinka, povećati broj vrednovanih sustava e-učenja, što će doprinijeti uniformnosti kvalitativnih karakteristika istih. Naime, razvojem informacijske i komunikacijske tehnologije došlo je do povećanja primjene različitih kategorija sustava e-učenja u procesu učenja i poučavanja. Naravno da svi ti sustavi nisu jednako kvalitetni i da su im karakteristike različite, pa ih je stoga potrebno karakterizirati nekom uniformnom kvalitativnom mjerom koja će omogućiti jednostavniji odabir sustava e-učenja.

6. ZAKLJUČAK

S obzirom na rastuće interese za razvoj sustava koji podupiru proces e-učenja i svekoliku prisutnost e-učenja u svijetu obrazovanja, potrebno je osigurati valjano i učinkovito okruženje u kojem bi se e-učenje odvijalo. Danas se skoro na svakom fakultetu na kojem postoji odjel za računarstvo, pristupa oblikovanju i implementaciji sustava e-učenja. Iako se ti sustavi razvijaju u svrhu potpore nastavi, mali je broj vrednovanih sustava. Naime, vrednovanje sustava, a pogotovo vrednovanje učinka sustava, zahtijeva dugotrajan angažman kako nastavnika koji te sustave koriste u nastavi, tako i studenata koji prisustvuju toj nastavi.

Upravo zbog činjenice da sustavi e-učenja zbog svoje prisutnosti u procesu učenja i poučavanja, utječu na taj isti proces, kao i na postignuća učenika, smatramo da se u nastavi mogu koristiti samo oni sustavi čija je učinkovitost provjerena. Pošto je proces vrednovanja učinka dugotrajan, a izbor metoda za vrednovanje velik, odlučili smo pristupiti definiranju vlastite metodologije za vrednovanje učinka sustava e-učenja.

U ovom radu smo prikazali odnos e-učenja i inteligentnih tutorskih sustava gdje smo inteligentne tutorske sustave prikazali kao posebnu vrstu inteligentnih sustava e-učenja, a samim time i sustava e-učenja općenito. Osvrnuli smo se na model inteligentne hipermedijske autorske ljuške za izgradnju inteligentnih tutorskih sustava TEx-Sys, te smo analizirali sva provedena testiranja nad spomenutim modelom.

Smatrali smo nužnim objasniti neke od osnovnih statističkih pojmova kao što su varijabla, populacija, uzorak, distribucija, nul-hipoteza, statistička značajnost, te opisati karakteristike distribucije frekvencija, položaj pojedinog rezultata u grupi, t-test, korelaciju, te na kraju, grafički način prikazivanja rezultata. Ovaj repetitorij statističkih pojmova je dan u svrhu što boljeg razumijevanja metoda za vrednovanje sustava e-učenja.

Posebna je pažnja posvećena objašnjavanju različitih pristupa vrednovanju sustava e-učenja. Na temelju spoznaja iz literature dali smo pregled i klasifikaciju metoda vrednovanja, objasnili smo postupke i instrumente prikupljanja podataka, kao i karakteristike tih instrumenata, te smo posebno objasnili eksperiment kao metodu vrednovanja zbog njegove sveprisutnosti u psihološkim i pedagoškim istraživanjima. Objasnili smo što je to veličina učinka i kako se može izračunati, te kako se može primijeniti kod meta-analiza.

U svrhu što boljeg razumijevanja složenosti i općeprisutnosti problematike vrednovanja učinkovitosti, analizirali smo nekoliko studija o vrednovanju različitih sustava e-učenja. To nam je omogućilo uočavanje sličnosti različitih metodologija, kao i problema koje svaki pristup nosi.

Zatim smo, na temelju spoznaja iz literature, pristupili definiranju vlastite metodologije za vrednovanje učinka procesa učenja i poučavanja u sustavima e-učenja. U tu svrhu je potrebno odrediti vrstu veličine učinka koja će se računati, zatim definirati strukturu eksperimenta, postaviti nul-hipoteze, pokazati kako se provodi ispitivanje statističke značajnosti, te kako analizirati i interpretirati rezultate. Naša metodologija predstavlja sintetički pristup vrednovanju

učinka kojim se povezuju eksperiment kao najčešće korištena metoda vrednovanja i ideja o provjeravanju učinka sustava u što većem broju međustanja. Ostale metodologije vrednovanja učinka sustava e-učenja promatraju učinak sustava samo na kraju eksperimenta, čime gube uvid u utjecaj sustava na postignuće učenika tijekom samog procesa učenja i poučavanja.

Definiranu metodologiju smo koristili za vrednovanje učinka sustava xTEx-Sys, kao jedne od izvedenica modela TEx-Sys, koji pripada klasi inteligentnih tutorskih sustava, a samim time i sustavima e-učenja. Metodologija za vrednovanje učinka koja je u ovom radu opisana, prikladna je za vrednovanje učinka bilo kojeg sustava e-učenja. Ona je u svrhu demonstracije, primijenjena za vrednovanje učinka inteligentnog tutorskog sustava, što je nipošto ne ograničava samo na tu klasu sustava.

Nakon što smo definirali vlastitu metodologiju za vrednovanje učinka sustava e-učenja i proveli eksperiment kojim smo vrednovali učinkovitost sustava xTEx-Sys, uočili smo niz poteškoća koje su se pojavile tijekom samog procesa vrednovanja. Naime, uočili smo veliki materijalni trošak prilikom tiskanja papirnatih testova, što predstavlja i loš utjecaj na ekološku svijest, zatim veliki gubitak vremena na ispravljanje i ocjenjivanje papirnatih testova, te poteškoće pri uporabi alata za statističku analizu. Sve je to utjecalo na razvoj ideje o automatiziranju samog postupka vrednovanja, te realizaciju iste. Stoga smo implementirali programski alat za automatsko vrednovanje učinka sustava e-učenja koji je nazvan EVEDIN (EVALuation of EDucational INfluence). EVEDIN izračunava veličinu učinka na temelju informacija o postignućima učenika jer se znanje učenika utvrđuje primjenom nekoliko „online“ testova: inicijalnog testa, testova provjere te završnog testa. Rezultati spomenutih testova prolaze kroz EVEDIN-ov mehanizam za statističku analizu, te kao rezultat EVEDIN daje izračunatu veličinu učinka sustava e-učenja.

Dakle, možemo zaključiti da je svrha ovog rada sistematizacija metoda vrednovanja koje se danas u svijetu koriste za vrednovanje učinka sustava e-učenja, definiranje vlastite metodologije za vrednovanje temeljem spoznaja iz literature, primjena te metodologije na vrednovanje učinka inteligentne autorske ljuske xTEx-Sys, te implementacija prototipa sustava za automatsko vrednovanje učinka sustava e-učenja.

Nastavak istraživanja je orijentiran prema primjeni definirane metodologije, kao i sustava EVEDIN, za vrednovanje učinka što većeg broja sustava e-učenja različitih kategorija, što bi omogućilo provedbu meta-analize i procjenu sveukupne veličine učinka sustava e-učenja općenito, kao i različitih podkategorija sustava e-učenja.

7. LITERATURA

- [AAAI1997] xxx, AAAI (1997) *Intelligent Tutoring System Authoring Tools*, Cambridge
- [AINS2003] Ainsworth, S.E., Williams, B.C & Wood, D.J. (2003) *Comparing the learning effectiveness of REDEEM and CBT*. 11th International Conference on Artificial Intelligence in Education. In U. Hoppe, F Verdejo. & J. Kay Proceedings of the 11th International Conference on Artificial Intelligence in Education. (pp 123-130) IOS Press: Amsterdam. ISBN 1586033565
- [ALBA2000] Albacete P.L. and VanLehn K.A. (2000) *Evaluating the Effectiveness of a Cognitive Tutor for Fundamental Physics Concepts*, in Proceedings of the 22nd Annual Meeting of the Cognitive Science Society
- [ALBE2001] Albert, D. (2001) *E-learning Future – The Contribution of Psychology*. (Keynote). In: Roth, R., Lowenstein, L., Trent, D. (eds.): *Catching the Future: Women and Men in Global Psychology*, Proceedings of the 59th Annual Convention, International Council of Psychologists, Winchester, England Lengerich, Germany, Pabst Science Publishers, 30-53
- [ALEX1994] Alexander, S., & Hedberg, J.G. (1994) *Evaluating technology-based learning: which model*. In K. Beattie, C. McNaught, & S. Wills (Eds), *Multimedia in Higher Education: Designing for change in teaching and Learning*. Amsterdam: Elsevier.
- [ANDE1985] Anderson, J.R. and Reiser, R.B. (1985) *The LISP tutor: it approaches the effectiveness of a human tutor*, BYTE, v.10 n.4, p.159-175
- [ANDE1988] Anderson, J. R. (1988) *The Expert Module*. In Polson, M.C., Richardson, J.J. (eds.) *Foundations of Intelligent Tutoring Systems*. Lawrence Erlbaum;
- [ARRU1997] Arruarte, A., Fernandez-Castro, I., Ferrero, B. & Greer, J. (1997) *The IRIS shell: How to build ITSs from pedagogical and design requisites*. International J. of Artificial Intelligence in Education. Vol. 8 , No. 3-4, pp. 341-381.
- [ARTI2005] Articulate Quizmaker 2.0, <http://www.articulate.com/quizmaker.html>
- [ASTD2001] ASTD (2001) *A Vision of E-Learning for America's Workforce*. Report of the Commission on Technology and Adult Learning, <http://www.astd.org>
- [BACK2000] Backer L.A., (2000) *Online syllabus - Basic and Applied Research Methods*, <http://web.uccs.edu/lbecker/Psy590/default.html>
- [BECK1996] Beck, J., Stern, M., and Haugsjaa, E. (1996) *Applications of AI in Education*, ACM Crossroads.

-
- [BLES1997] Blessing, S.B. (1997) *A programming by demonstration authoring tool for model tracing tutors*. Int. J. of Artificial Intelligence in Education. Vol. 8 , No. 3-4, pp 233-261.
- [BLOO1956] Bloom B. S. (1956) *Taxonomy of Educational Objectives*, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.
- [BLOO1984] Bloom, B.S. (1984) *The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring*. Educational Researcher, 13, 4-16
- [BOŽI2001] Božičević, J. (2001) *Zašto zagovaramo sustavski pristup?* Mislimo sustavski, Hrvatsko društvo za sustave
- [BROW1978] Brown, J. S., & Burton, R. R. (1978) *Diagnostic models for procedural bugs in basic mathematical skills*. Cognitive Science, 2, 155-192.
- [BRUS1996] Brusilovsky, P, Schwartz, E. & Weber, G. (1996) *ELM -ART: An Intelligent Tutoring System on the Work Wide Web*. In Proceedings of ITS-96, Frasson, Gauthier, Lesgold (Eds.), Springer: Berlin, 1996. pp. 261-269.
- [BRUS1997] Brusilovsky, P., Ritter, S., & Schwarz, E. (1997) *Distributed intelligent tutoring on the Web*. In: B. du Boulay and R. Mizoguchi (eds.) Proceedings of AI-ED'97, 8th World Conference on Artificial Intelligence in Education, 18-22 August 1997. Amsterdam: IOS, pp. 482-489.
- [BURN1988] Burns H. L., Capps C. G. (1988) *Foundations of intelligent tutoring systems: an introduction*, M.C.Poison, J. J.Richardson (Ed.) Foundations of intelligent tutoring systems, Lawrence Erlbaum, London, 1-19
- [BURT1982] Burton, R. R. (1982) *Diagnosing bugs in a simple procedural skill*. In D. Sleeman & J. S. Brown (Eds.), Intelligent Tutoring Systems (pp. 157-183). New York: Academic Press.
- [CARB1970] Carbonell, J.R. (1970) *AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction*. IEEE Transactions on Man-Machine Systems, 11 (4), pp. 190-202
- [CARB1973] Carbonell, J.R., Collins, A. (1973) *Natural Semantics in Artificial Intelligence*. Proceedings of Third International Joint Conference on Artificial Intelligence, pp. 344-351.
- [CARD1983] S. K. Card, T. P. Moran, A. Newell (1983) *The Psychology of Human-Computer Interaction*, Erlbaum, Hillsdale
- [CHER1985] Cherniak, E., McDermot, D. (1985) *Introduction to Artificial Intelligence*, Addison-Wesley Pub Co.
- [CLAN1988] Clancey, W. & Joerger, K. (1988) *A Practical Authoring Shell for Apprenticeship Learning*. Proceedings of ITS-88, 67-74. June 1988, Montreal.
-

-
- [COE2000] Coe, R. (2000) *What is an 'effect size'? A guide for users*. University of Durham, Curriculum, Evaluation and Management Centre, Evidence-Based Education-UK
<http://www.cemcentre.org/ebeuk/research/effectsize/ESguide.htm>
- [COHE1969] Cohen, J. (1969) *Statistical Power Analysis for the Behavioral Sciences*. NY: Academic Press.
- [COLL1975] Collins, A., Warnock, E. H. and Passafiume, J. J. (1975) *Analysis and synthesis of tutorial dialogues*. In: Bower, G. H. (ed.) *The Psychology of Learning and Motivation: Advances in Research and Theory*, pp.49-87, New York: Academic Press.
- [DAVI2005] Davies, W. M. and Davis, H. C. (2005) *Designing Assessment Tools in a Service Oriented Architecture*. In *Proceedings of 1st International ELeGI Conference on Advanced Technology for Enhanced Learning BCS Electronic Workshops in Computing (eWiC)*, Vico Equense, (Napoli), Italy.
- [DEMP2004] Dempster, J. (2004) *CAP e-Learning guides: Evaluating E-Learning*, Centre of Academic Excellence, University of Warwick, www.warwick.ac.uk/go/cap/resources/eguides
- [DEVE2003] Devedžić V. (2003) *Think ahead: evaluation and standardization issues for e-learning applications*, *Int. J. Continuing Engineering Education and Lifelong Learning*, Vol. 13, Nos. 5/6, pp. 556-566.
- [DRAP1996] Draper S.W., Brown M.I., Henderson F.P. & McAteer E. (1996) *Integrative evaluation: an emerging role for classroom studies of CAL*, *Computers & Education* 26(1-3) pp 17-32.
- [FLET2003] Fletcher J.D. (2003) *Evidence for Learning From Technology-Assisted Instruction*, in H.F. O'Neal, R.S. Perez (Eds.), *Technology applications in education: a learning view*, Mahwah, NJ: Lawrence Erlbaum Associates, pp.79-99
- [GARS2006] Garson, D. (2006) *Statnotes: Topics in Multivariate Analysis*. <http://www2.chass.ncsu.edu/garson/pa765/statnote.htm>
- [GLAS1981] Glass, G.V., McGaw, B. and Smith, M.L. (1981) *Meta-Analysis in Social Research*. London: Sage.
- [HEDG1985] Hedges, L. and Olkin, I. (1985) *Statistical Methods for Meta-Analysis*. New York: Academic Press.
- [HEMP2005] Hempel, S. (2005) *Reliability*. In: J.N.V. Miles & P. Gilbert (Ed.), *A handbook of research methods for clinical & health psychology*. Oxford: Oxford University Press.
- [IQBA1999] 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, pp. 169-181.
-

-
- [ITSW1996] xxx, ITS'96 (1996) Workshop on Architectures for ITS, Position Papers and Statements of Interest, Montreal, June 10th, <http://advlearn.lrdc.pitt.edu/its-arch/papers/index.html>
- [JACK1998] Jackson, B. (1998) *Evaluation of Learning Technology Implementation*. In N. Mogyey (Ed.), LTDI Evaluation Studies: Learning Technology Dissemination Initiative. <http://www.icbl.hw.ac.uk/ltidi/evalstudies/esevalimp.htm>
- [JAME2002] James, R., McInnis, C., Marcia D. M. (2002) *Assessing Learning in Australian Universities*, Centre for the Study of Higher Education, University of Melbourne, <http://www.carrickinstitute.edu.au/carrick/webdav/site/carricksite/users/siteadmin/public/learning.pdf>
- [JONE1991] Jones, M. & Wipond, K. (1991) *Intelligent Environments for Curriculum and Course Development*. In Goodyear (Ed.), Teaching Knowledge and Intelligent Tutoring. Norwood, NJ: Ablex.
- [KIYA1997] Kiyama, M., Ishiuchi, S., Ikeda, K., Tsujimoto, M. & Fukuhara, Y. (1997) *Authoring Methods for the Web-Based Intelligent CAI System CALAT and its Application to Telecommunications Service*. In the Proceedings of AAAI-97, Providence, RI.
- [KHAN2001] Khan, B. H. (2001) *A framework for Web-based learning*. In: Khan, B. H. (ed.): Web-based training. Englewood Cliffs, NJ: Educational Technology Publications
- [LELO1999] Lelouche, R. (1999) *Intelligent Tutoring Systems from Birth to Now*, KI-Künstliche Intelligenz, 4, pp. 5-11
- [LIPS2001] Lipsey M.W., and Wilson, D. B. (2001) *Practical Meta-Analysis*. Applied Social Research Methods Series (Vol. 49). Thousand Oaks, CA: SAGE Publications.
- [MARK1993] Mark M.A. and Greer J.E. (1993) *Evaluation methodologies for intelligent tutoring systems*, Journal of Artificial Intelligence and Education, 4 (2/3), pp. 129-153.
- [MAXW2004] Maxwell, S. E., & Delaney, H. D. (2004) *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates
- [MCTA2001] Mc Taggart J (2001) *Intelligent Tutoring Systems and Education for the Future*. In: 512X Literature Review April 30, pp 2.
- [MERI2001] Merrill, M. D. (2001) *Components of instruction: toward a theoretical tool for instructional design*. Instructional Science. 29(4/5), 291-310.
- [MITR1999] Mitrovic A. and Ohlsson S. (1999) *Evaluation of a Constraint-Based Tutor for a Database Language*, International Journal of Artificial Intelligence in Education 10, pp. 238-256.
-

- [MITR2005] Mitrovic, A. (2005) *The Effect of Explaining on Learning: a Case Study with a Data Normalization Tutor*. In: C-K Looi, G. McCalla, B. Bredeweg, J. Breuker (eds) Proc. Artificial Intelligence in Education AIED 2005, IOS Press, pp. 499-506,
- [MOHA1998] Mohammad N.Y. (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
- [MUNR1997] Munro, A., Johnson, M.C., Pizzini, Q.A., Surmon, D.S., Towne, D.M, & Wogulis, J.L. (1997) *Authoring simulation-centered tutors with RIDES*. International J. of Artificial Intelligence in Education. Vol. 8 , No. 3-4, pp. 284-316.
- [MURR1996] T. Murray (1996) *Having It All, Maybe: Design Tradeoffs in ITS Authoring Tools*, In Proceedings of the Third International Conference on Intelligent Tutoring Systems, Montreal
- [MURR1998] Murray, T. (1998) *Authoring knowledge-based tutors: Tools for content, instructional strategy, student model, and interface design*. J. of the Learning Sciences, Vol. 7. No. 1, pp. 5-64.
- [MURR2003] Murray, Ainsworth, & Blessing (eds.) (2003) *Authoring Tools for Advanced Technology Learning Environment* © 2003 Kluwer Academic Publishers. Printed in the Netherlands. pp. 493–546.
- [MUŽI1977] V. Mužić (1977) *Metodologija pedagoških istraživanja*, Svjetlost, Sarajevo
- [NKAM1996] Nkambou, R., Gauthier, R., & Frasson, M.C. (1996) *CREAMTools: an authoring environment for curriculum and course building in an ITS*. In Proceedings of the Third International Conference on Computer Aided Learning and Instructional Science and Engineering. New York: Springer-Verlag.
- [OHLS1987] Ohlsson, S. (1987) *Some Principles of Intelligent Tutoring*. In Lawler & Yazdani (Eds.), Artificial Intelligence and Education, Volume 1. Ablex: Norwood, NJ, pp. 203-238.
- [PASK1965] Pask, G. (1965) *A cybernetic model of concept learning*, Proceedings of 3rd, Congress International. Assoc. Cybernetics, Namur 1961, Gauthier-Villars
- [PERC2005] Perception Questionmark, <http://www.questionmark.com/us/perception/index.htm>
- [PETZ2004] Petz, B. (2004) *Osnovne statističke metode za nematematičare*, Naklada Slap, Jastrebarsko, 5. izdanje
- [PFAH2004] D. Pfahl, O. Laitenberger, G. Ruhe, J. Dorsch, T. Krivobokova (2004) *Evaluating the Learning Effectiveness of Using Simulations in Software Project Management Education: Results from a Two Times Replicated Experiment*. Information and Software Technology (Elsevier), Vol 46 pp 81-147.
- [PHIL2002] Phillips, R. A., & Gilding, A. (2002) *Approaches to evaluating the effect of ICT on student learning*. Learning and Teaching Support Network.

-
- [PONB2004] Pon-Barry, H., Clark, B., Bratt, E., Schultz, K., & Peters, S. (2004) *Evaluating the effectiveness of SCoT: a Spoken Conversational Tutor*. In Mostow, J. & Tedesco, P. (Eds.), ITS 2004 Workshop on Dialog-based Intelligent Tutoring Systems, 23-32, Retrieved October 25, 2005, from, http://www-csli.stanford.edu/semlabhold/muri/papers/ITS_2004_Workshop.pdf.
- [QIUR2003] Qiu, L., Riesbeck, C.K., and Parsek, M.R. (2003) *The Design and Implementation of an Engine and Authoring Tool for Webbased Learn-by-doing Environments*. Proc. of World Conf. on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA 2003). June 23-28, 2003, Honolulu, HA. AACE.
- [QUES2005] Question Writer 2.0, <http://www.questionwriter.com/product.html>
- [RAZZ2004] Razzaq, L. & Heffernan, N. T (2004) *Tutorial dialog in an equation solving intelligent tutoring system*. Proceedings of 7th Annual Intelligent Tutoring Systems Conference, Maceio, Brazil, <http://nth.wpi.edu/ITS2004-WS/W8Proceedings1.pdf>
- [REDF1996] Redfield, C.L., (1996) *Demonstration of the experimental advanced instructional design advisor*. In the Third International Conference on Intelligent Tutoring Systems, Montreal, Quebec, Canada, June 12-14, 1996
- [RICK1989] Rickel, J. (1989) *Intelligent Computer Aided Instruction: A Survey Organized Around System Components*, IEEE Transactions on Systems, Man, and Cybernetics, 19 (1), pp. 40-57, Jan/Feb
- [ROSI2000] Rosić, M. (2000) *Establishing of Distance Education Systems within the Information Infrastructure*. Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia, MS Thesis
- [RUSS1988] Russell, D. (1988) *IDE: The Interpreter*. In Psootka, Massey, & Mutter (Eds.), Intelligent Tutoring Systems, Lessons Learned. Hillsdale, NJ:Lawrence Erlbaum.
- [SCHE1999] Schewe, S., Reinhardt, B., Bestz, C. (1999) *Experiences with a Knowledge Based Tutoring System for Student Education in Rheumatology*. In XPS-99: Knowledge Based Systems: Survey and Future Direction, 5th Biannual German Conference on Knowledge Based Systems, Lecture Notes in Artificial Intelligence 1570, Springer.
- [SCOR2004] xxx, *SCORM 2004 Overview*, Advanced Distributed Learning, 2004. <http://www.adlnet.org/>
- [SELF1990] Self, J. (1990) *Theoretical Foundations for Intelligent Tutoring Systems*. Journal of Artificial Intelligence in Education, 1(4), 3-14
- [SHUT1996] Shute, V. J., and Psootka, J., (1996) *Intelligent tutoring systems: Past, present, and future*. In D. Jonassen (Ed.), Handbook of Research for Educational Communications and Technology (pp. 570-600). New York, NY: Macmillan.
-

-
- [SHUT2000] Shute, V. J. & Torreano, L., & Willis, R. (2000) *DNA: Towards an automated knowledge elicitation and organization tool*. In S. P. Lajoie (Ed.) *Computers as Cognitive Tools*, Volume 2. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 309-335.
- [SLEE1982] Sleeman, D., & Brown, J. S. (1982) *Introduction: Intelligent Tutoring Systems*. In D. Sleeman & J. S. Brown (Eds.), *Intelligent Tutoring Systems* (pp. 1-11). New York: Academic Press.
- [SMAR2005] Smartlite WebQuiz XP, <http://www.smartlite.it/en2/products/webquiz/index.asp>
- [SPAR1999] Sparks, R. Dooley, S., Meiskey, L. & Blumenthal, R. (1999) *The LEAP authoring tool: supporting complex courseware authoring through reuse, rapid prototyping, and interactive visualizations*. *Int. J. of Artificial Intelligence in Education*.
- [STAN1997] Stankov, S. (1997) *Isomorphic Model of the System as the Basis of Teaching Control Principles in an Intelligent Tutoring System*. Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split, Croatia, PhD Thesis
- [STAN2001] S. Stankov, V. Glavinić, A. Granić, M. Rosić (2001) *Inteligentni tutorski sustavi – istraživanje, razvoj i primjena*, CARNET - Časopis Edupoint (elektronička verzija – <http://www.carnet.hr>); I – dio: godište II, broj 1, Zagreb; 20.12.2001.
- [STAN2003] S. Stankov (glavni istraživač): Tehnologijski projekt MZT: Web orijentirana inteligentna hipermedijska autorska ljuska (TP-02/0177-01), za razdoblje 2003-2005.
- [STAN2004a] Stankov, S., Grubišić, A., Žitko B. (2004) *E-learning paradigm & Intelligent tutoring systems*. In: Kniewald, Z. (ed.): *Annual 2004 of the Croatian Academy of Engineering*. Croatian Academy of Engineering, Zagreb, 21-31
- [STAN2004b] Stankov, S., Glavinić V., Grubišić A. (2004) *What is our effect size: Evaluating the Educational Influence of a Web-Based Intelligent Authoring Shell?* In: Nedeveschi, S., Rudas, I. J. (eds.): *Proceedings INES 2004 / 8th International Conference on Intelligent Engineering Systems*. Cluj-Napoca, Faculty of Automation and Computer Science, Technical University of Cluj-Napoca, 545-550
- [STAN2005] Stankov, S., Grubišić, A., Žitko, B., Krpan, D. (2005) *Vrednovanje učinkovitosti procesa učenja i poučavanja u sustavima za e-učenje*, Školski Vjesnik - časopis za pedagoška i školska pitanja, 54 , 1-2; 21-31
- [THAL2002] Thalheimer, W., & Cook, S. (2002) *How to calculate effect sizes from published research articles: A simplified methodology*. http://work-learning.com/effect_sizes.htm.
- [TOUR1992] Touritzky, D. S. (1992) *Inheritance Hierarchy*. In I. Shapiro, C. Stuart, editors, *Artificial Intelligence – Encyclopedias*, John Wiley & Sons, Inc, pp. 690-701.
-

-
- [TWID1993] Twidale, M. B. (1993) *Redressing the balance: The advantages of informal evaluation techniques for Intelligent Learning Environments*. Journal of Artificial Intelligence In Education, 4(2/3):155--178.
- [UHRL1969] Uhr, L. (1969) *Teaching machine programs that generate problems as a function of interaction with students*. Proceedings of the 24th National Confernece, 125-134.
- [URBA1996] Urban-Lurain, Mark. (1996) *Intelligent Tutoring Systems: An Historic Review in the Context of the Development of Artificial Intelligence and Educational Psychology*. <http://aral.cse.msu.edu/Publications/ITS/its.htm>.
- [VANJ1996] Van Joolingen, W.R. & Jong, T. de (1996) *Design and implementation of simulation-based discovery environments: The SMISLE solution*. Int. J. of Artificial Intelligence in Education 7(3/4). pp. 253-276.
- [VANM1998] Van Marcke, K. (1998) *GTE: An epistemological approach to instructional modeling*. Instructional Science, Vol. 26, pp 147-191.
- [WENG1987] Wenger, E. (1987) *Artificial Intelligence and Tutoring Systems*. Los Altos, California: Morgan Kaufmann Publishers.
- [WIEN1948] N. Wiener (1948) *Cybernetics or Control and Communication in the Animal and the Machine*, J. Wilay, New York
- [WINN1988] Winne, P. & Kramer, L. (1988) *Representing and Inferencing with Knowledge about Teaching: DOCENT*. Proceedings of ITS-88. June 1988, Montreal, Canada.
- [WISH2003] Wisher R.A. and Olson T.M. (2003) *The Effectiveness of Web-based Training*, U.S. Army Research Institute for the Behavioral and Social Sciences, Research Report 1802
- [WONG1997] Wong, W.K. & Chan, T.W. (1997) *A Multimedia authoring system for crafting topic hierarchy, learning strategies, and intelligent models*. International J. of Artificial Intelligence in Education, Vol. 8, No 1, pp. 71-96.
- [ZHAN2006] Zhang, D., Zhou, L., Briggs, R.O., and Nunamaker, J.F. Jr. (2006) *Instructional Video In E-learning: Assessing the Impact of Interactive Video on Learning Effectiveness*. Information and Management, 43(1), pp. 15–27.

8. PRILOZI

- A Inicijalni test
- B Prvi test provjere
- C Drugi test provjere
- D Završni test
- E Rezultati testova

Ime i prezime: _____

Studijska grupa: _____

Završena srednja škola: _____

1. Navedite nazive komponenti računala koje poznajete:

2. Izaberite uređaje koji spadaju u izlazne jedinice!

- a. Tipkovnica
- b. Štampač
- c. Monitor

3. Što je MS DOS?

- a. Tablični kalkulator
- b. Operacijski sustav
- c. Ulazna jedinica

4. Napišite neku e-mail adresu.

5. Odgovorite:



Koju funkciju smo upotrijebili?

6. Koliko bajtova ima 4 kilobajta?

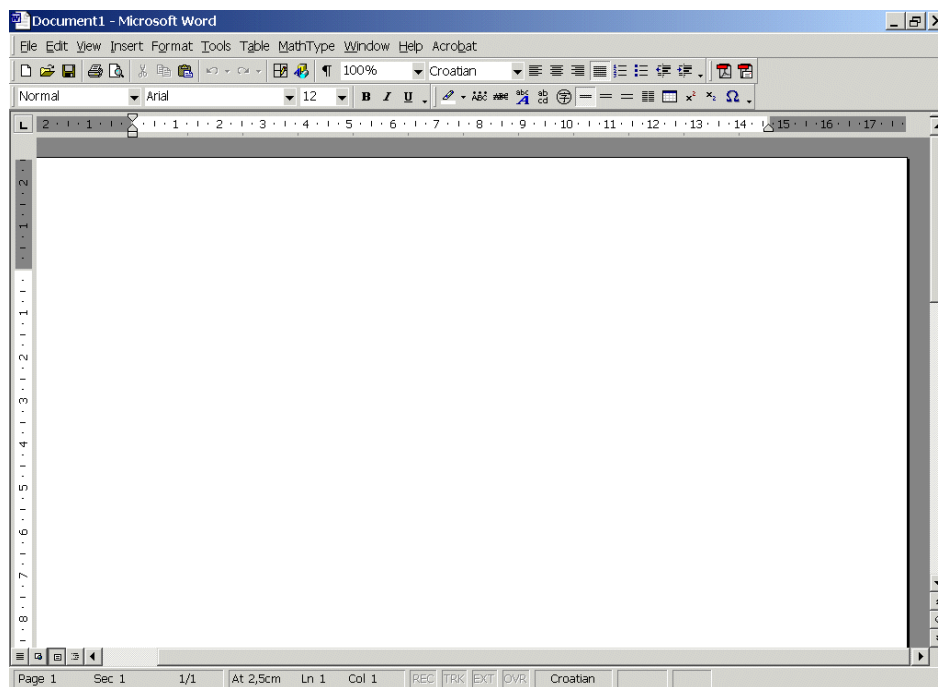
- a. 4064
- b. 4096
- c. 4004

7. Datoteka napravljena u MS Wordu ima ekstenziju:

- a. .doc
- b. .ppt
- c. .gif

8. Što je to floppy?

9. Označite elemente MS Windows prozora odgovarajućim slovom:



- a. Traka naslova
- b. Traka s izbornicima
- c. Traka s alatima
- d. Ravnalo
- e. Radna površina
- f. «Scroll bar»
- g. Oznake za minimiziranje, maksimiziranje i zatvaranje prozora

10. Objasnite sljedeće engleske računarske izraze na hrvatskom jeziku:

Web – _____

Site - _____

- Homepage - _____
- Chat - _____
- Browser - _____
- Server - _____
- Resource - _____
- Download - _____
- Transmit - _____
- Default - _____
- Desktop - _____
- Printer - _____
- Link - _____
- Fonts - _____
- Undo - _____

11. Zaokružite: Koja instrukcija zbraja cijele brojeve stupca D?

	A	B	C	D	E
1					
2	1	2	3	4	
3	5	6	7	8	
4	9	10	11	12	
5	13	14	15	16	
6	17	18	19	20	
7					
8					
9					
10					

- a) SUM(D2+D6)
 b) SUM(D2*D6)
 c) SUM(D2:D6)
 d) SUM(D2-D6)

12. Binarnom broju 1001110 dekadski ekvivalent je:

- a. 78
 b. 69
 c. 85

13. Pretvorite oktalni broj 563 u binarni (prikažite svaki korak pretvorbe):

14. Pridijelite nazivu programskih sustava objašnjenje njihove funkcionalnosti. Upišite odgovarajući broj u kvadratić.

MS Windows	<input type="text"/>	1 Aplikacija za komprimiranje/ dekomprimiranje datoteka
MS Word	<input type="text"/>	2 Aplikacija za organiziranje datoteka i direktorija na disku
MS Excel	<input type="text"/>	3 Operacijski sustav s grafičkim korisničkim sučeljem
MS PowerPoint	<input type="text"/>	4 Preglednik web stranica
MS Access	<input type="text"/>	5 Aplikacija za rad s e-mail-om
Internet Explorer	<input type="text"/>	6 Aplikacija za izradu tabličnih proračuna
Windows Explorer	<input type="text"/>	7 Aplikacija za izradu prezentacija
Outlook Express	<input type="text"/>	8 Aplikacija za izradu relacijskih baza podataka
WinZip	<input type="text"/>	9 Aplikacija za upis i obradu tekstualnih podataka

Ime i prezime: _____

Studijska grupa: _____

E-mail adresa: _____

1. Nabrojite temeljne funkcije računalnog sustava i ukratko ih opišite.

2. Nabrojite generacije računalnih sustava i ukratko ih opišite.

3. Koja je razlika između programske podrške i tehničke podrške?

- a) Programska podrška je naziv za sve programe koje računalo može izvršiti, a tehnička podrška je naziv za sve dijelove računala;
- b) Programska podrška je naziv za sve programe koje računalo može izvršiti i podatke s kojima programi rade, a tehnička podrška je naziv za sve dijelove računala;
- c) Programska podrška je naziv za sve programe koje računalo izvršava, a tehnička podrška su podaci koje obrađuje programska podrška.

4. Koja je razlika između sistemske i aplikacijske programske podrške?

- a) Sistemsku programsku podršku razvija proizvođač te je prilagođena konfiguraciji tehničke podrške računala s kojim je isporučena. Aplikacijsku programsku podršku razvija sam korisnik u skladu sa svojim potrebama;
- b) Aplikacijsku programsku podršku razvija proizvođač te je prilagođena konfiguraciji tehničke podrške računala s kojim je isporučena. Sistemsku programsku podršku razvija sam korisnik u skladu sa svojim potrebama;
- c) Sistemsku programsku podršku razvija proizvođač u skladu s potrebama korisnika. Aplikacijsku programsku podršku razvija sam korisnik obzirom na tehničku podršku od koje mu se sastoji računalo.

5. Nacrtajte i opišite John Von Neumann-ov koncept sustava računala:

6. Objasnite što je to memorija, kako ju dijelimo, te ukratko opišite svaku vrstu memorije koju poznajete. Što je kapacitet memorije?

7. Što je mikroprocesor? Koji je prvi mikroprocesor?

8. RAM je :

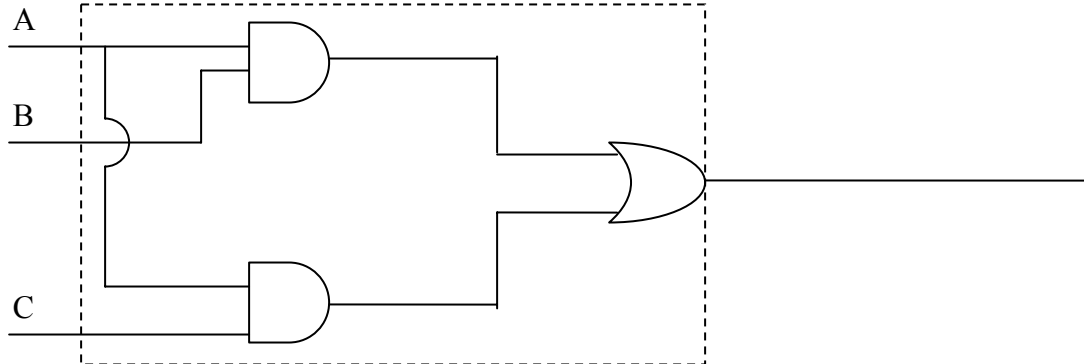
- a) kratica od random access memory (memorija sa slučajnim pristupom)
- b) radna memorija - može se samo pročitati
- c) masovna memorija
- d) stalna memorija

9. Što je to CD-RW?

10. Što se podrazumijeva pod računalnom snagom CPU?

11. U usporedbi paralelnog i serijskog porta, kojim se istodobno prenosi više bitova i zašto?

12. *Koja je logička jednažba zadanog logičkog sklopa? Primjeni postupak minimizacije i nacrtaj minimizirani sklop (sklop s istim djelovanjem, ali s manjim brojem osnovnih logičkih sklopova). Napiši tablicu istinitosti za minimizirani sklop.*



13. Na koji se port najčešće priključuje modem?

14. Koje vrste modema postoje s obzirom na način povezivanja s računalom?

15. Monitor spajamo s računalom ?

- a) USB priključkom
- b) VGA konektorom
- c) PS/2 konektorom
- d) Serijski
- e) Paralelno

16. Kako se izražava rezolucija monitora i što ta oznaka znači?

17. Što je to virus?

18. Pretvori broj iz jednog brojevnog sustava u drugi:

11011011₍₂₎

(8)

(10)

(16)

19. Izvršite sljedeće konverzije brojevnik sustava:

$$(1E0B)_{16}=(\quad)_{10} \qquad (10101110110111)_2=(\quad)_8$$

20. Brojeve $125_{(8)}$ i $125_{(10)}$, ako zbrojimo, dobijemo rezultat koji bi u binarnom brojevnom bio:

- a) 11010011
- b) 11010010
- c) 10010010
- d) 11000010

Ime i prezime: _____

Studijska grupa: _____

E-mail adresa: _____

1. Što je i čemu služi operacijski sustav?

2. Nabrojite i ukratko objasnite osnovne komponente operacijskog sustava MS-DOS.

3. Datoteka CONFIG.SYS. Glavne značajke.

4. *Datoteka AUTOEXEC.BAT. Glavne značajke.*

5. *Što je FAT?*

6. *Neka je vaš tekući (aktivni) direktorij u okruženju operacijskog sustava MSDOS glavni (root) direktorij diska C (tj. C:\). Napišite redom sve instrukcije MSDOS-a koje morate koristiti da bi napravili sljedeću strukturu direktorija*

```
C:\>
|__MAT
|__FIZ
    |__OPCA
    |__ATOM
|__KEM
```

7. *U tekućem direktoriju se nalaze sljedeće datoteke: abc.txt, acb.txt, abc.doc. Imena kojih datoteka će ispisati naredbe:*

a) `dir *.txt`

b) `dir a*.*`

8. Nabrojite metode obrade podataka i ukratko ih opišite.

9. Što je sažimanje podataka?

10. Što je kodna stranica?

11. Što je ASCII kod?

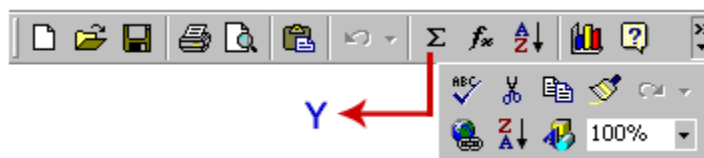
12. Koliko je bite-ova: 1 KB, 1 MB, 1 TB?

13. Koje od sljedećih formula možemo upisati u ćeliju G2 da bi se u njoj prikazao prosjek prodaje proizvoda od siječnja do travnja?

	A	B	C	D	E	F	G
1		siječanj	veljača	ožujak	travanj	ukupno	prosjeak
2	proizvod1	30	40	44	53	167	
3	proizvod2	21	11	22	23	77	
4	proizvod3	445	323	444	456	1668	
5	proizvod4	22	33	33	45	133	
6	proizvod5	12	11	1	2	26	
7						2071	

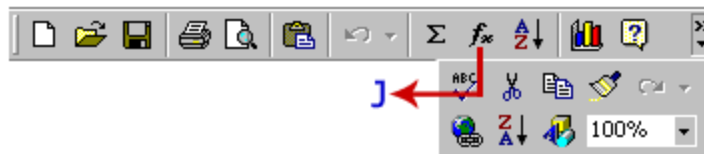
- a) =sum(average(B2:E2))
- b) =sum(B2-E2)/4
- c) =sum(B2:E2)/4
- d) =sum(F2/4)

14. Za što nam služi alat označen slovom Y?



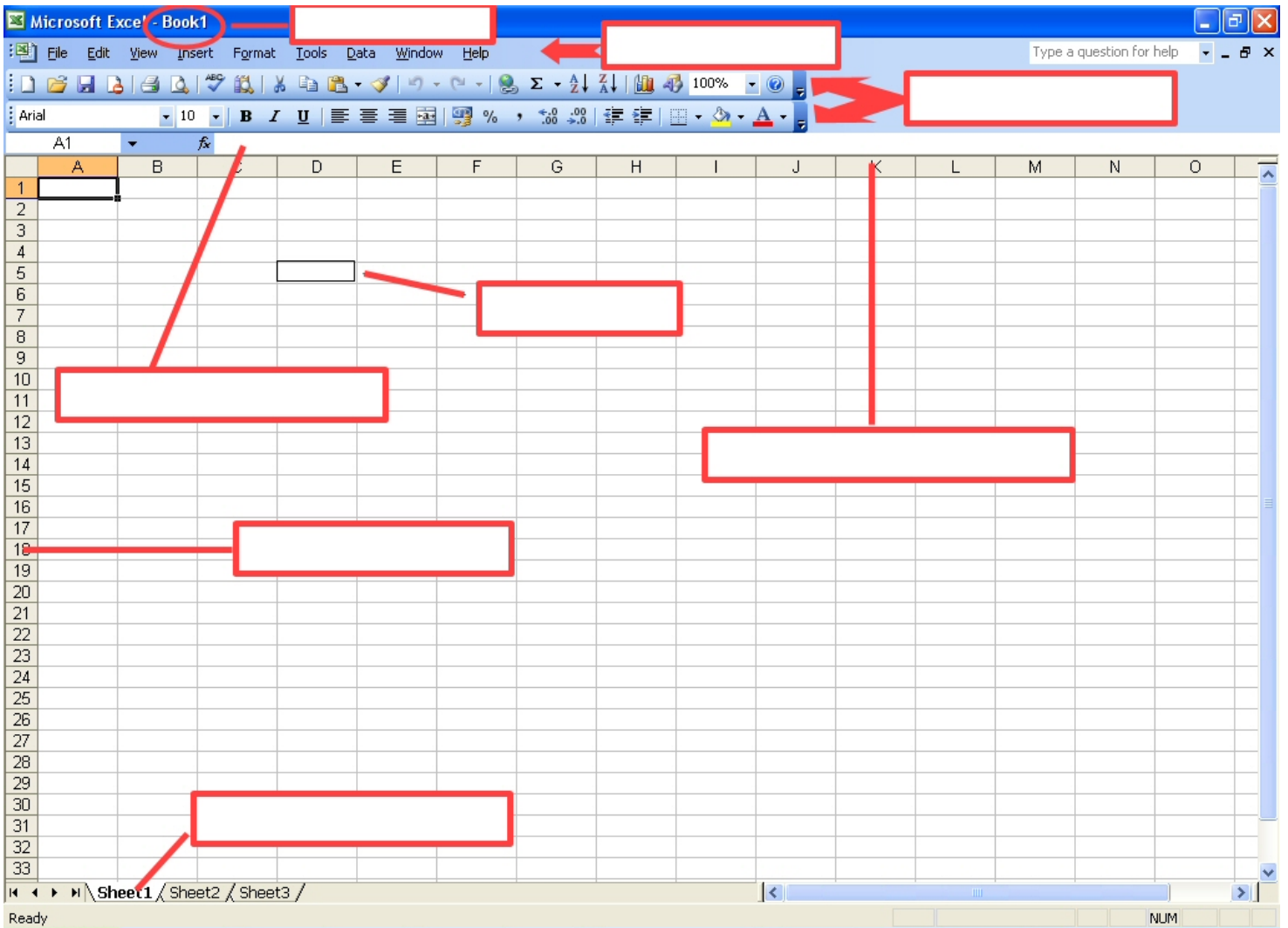
- a) Za računanje prosječne vrijednosti označenih brojeva
- b) Za zbrajanje vrijednosti u označenim ćelijama
- c) Za spajanje označenih ćelija u jednu
- d) Za prikazivanje popisa funkcija

15. Za što nam služi alat označen slovom J?



- a) Za računanje prosječne vrijednosti označenih brojeva
- b) Za zbrajanje vrijednosti u označenim ćelijama
- c) Za spajanje označenih ćelija u jednu
- d) Za prikazivanje popisa funkcija

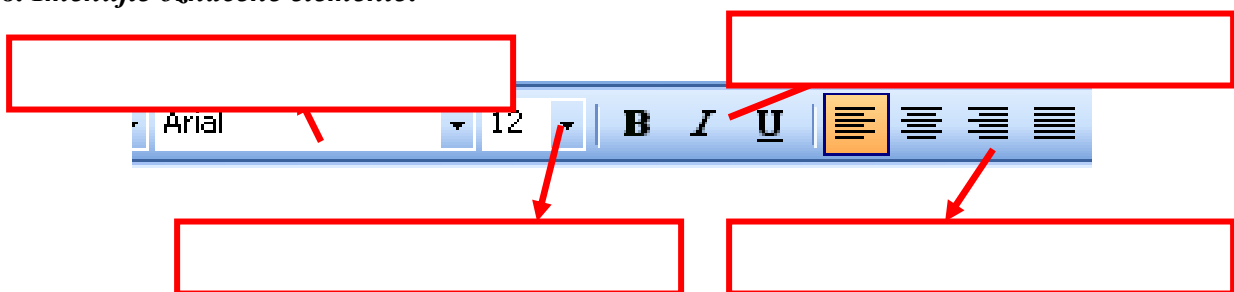
16. Imenujte označene elemente.



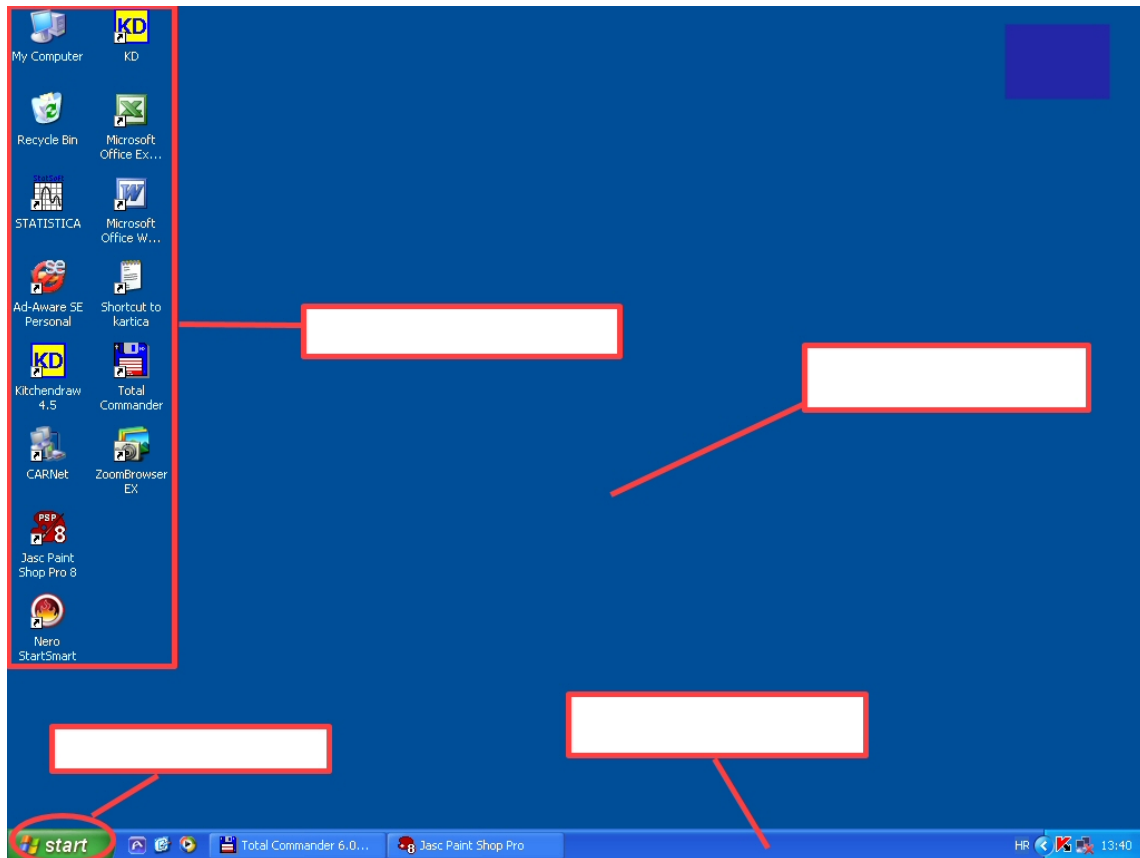
17. Skup predefinirane grafike unutar MS Word-a naziva se:

- clip art
- clip graphics
- clip photos
- clips

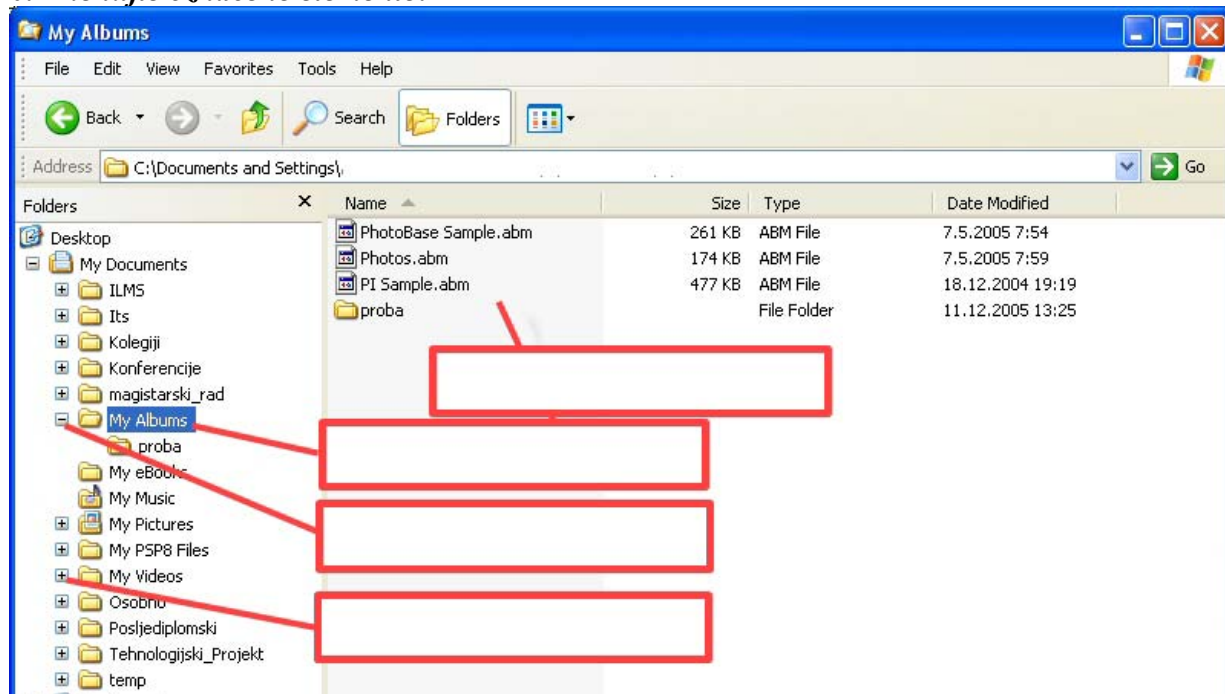
18. Imenujte označene elemente:



19. Imenujte označene elemente:



20. Imenujte označene elemente:



Ime i prezime: _____

Studijska grupa: _____

E-mail adresa: _____

1. Nabroji i objasni temeljne funkcije računalnog sustava:

2. Koja je glavna razlika između RAM i ROM memorije?

3. Kako je na disketi 3,5" moguće privremeno onemogućiti brisanje i upis podataka?

4. CPU je skraćeni naziv za...

5. Koji od navedenih uređaja mogu biti ULAZI ?

- a) tipkovnica
- b) miš
- c) monitor
- d) štampač
- e) disketa

6. Opiši funkcije tipki Shift, AltGr, Backspace, Delete, Enter, Esc.

7. Čemu služe prekidači u naredbama operacijskog sustava MS-DOS?

8. *Kamo se kopira datoteka ako se ne navede drugi parametar u naredbi copy operacijskog sustava MS-DOS?*

9. *Zbog čega se provodi sažimanje podataka?*

10. *15000B= KB 1,2 GB= MB*

11. *Čemu služe alati na slici?*



12. Objasni funkcije sum i count. Navedi primjer.

13. Kako ćeš razlikovati da li je podatak upisan u ćeliju broj ili tekst?

14. Kojim simbolom započinje formula?

15. Što je i čemu služi računalna mreža?

16. Što je Internet i koje sve usluge pruža Internet?

17. Što je to hipertekst?

18. Navedi jedan URL i objasni njegove dijelove.

19. Objasni sliku



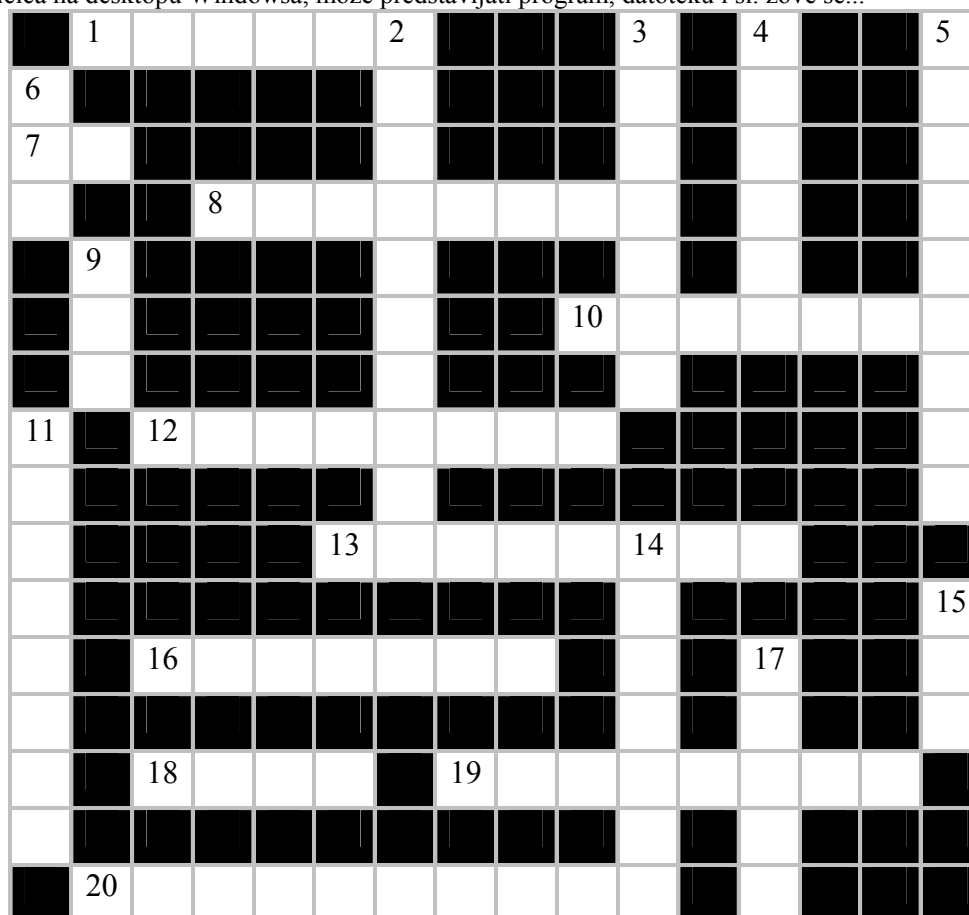
20. Križaljka

Vodoravno:

1. Ulazna jedinica, služi za unos slike, fotografije i sl.
7. Vrsta računala koja se pojavila početkom 80-tih, zasnovana je na mikroprocesoru i koristi se do danas.
8. Traka pri vrhu prozora na kojoj se nalaze naredbe zove se...
10. Osnivač Microsoft-a je Bill...
12. Udaljeno "glupo" računalo koje ima samo tipkovnicu i monitor zove se...
13. Sklopovski, opipljivi dio računala zove se (engl.)...
16. Popularni naziv za pregledavanje web stranica
18. Neželjene e-mail poruke (reklame, poruke o brzjoj zaradi i sl.) zovu se...
19. Postupak "skidanja" datoteka sa Interneta na vlastito računalo zove se (engl.)...
20. Vrsta teksta koji sadrži veze na druge dijelove teksta ili druge dokumente zove se...

Okomito:

2. Podatak o broju pixela po horizontali i vertikali ekrana zove se...
3. Traka uz gumb Start zove se...
4. Kad smo spojeni na Internet kažemo da smo...
5. Vrsta software-a u koju spada operacijski sustav je...
6. Najpopularniji format komprimiranih glazbenih datoteka je...
9. Usluga na Internetu koja omogućuje prijenos datoteka s jednog računala na drugo zove se...
11. Port na koji spajamo miša i modem.
14. Mreža iz koje se razvio Internet zvala se je...
15. Vrsta monitora tj. ekrana s tekućim kristalima.
17. Mala sličica na desktopu Windowsa, može predstavljati program, datoteku i sl. zove se...



kontrolna grupa				
	inicijalni test	prvi test provjere	drugi test provjere	završni test
student103	55	58	60	47
student104	58	34	67	49
student105	31	25	53	30
student106	72	50	52	50
student107	60	29	48	34
student108	63	36	64	37
student109	25	64	64	52
student113	45	53	73	31
student114	54	49	63	33
student117	39	14	50	14
student119	25	19	5	20
student120	58	60	58	50
student122	45	10	17	21
student126	45	48	61	50
student128	75	56	64	51
student129	49	45	65	37
student134	52	15	50	21
student139	39	38	84	54
student141	13	15	58	17
student143	58	43	52	39
student144	51	43	42	35
student146	16	50	38	32
student149	43	45	71	44
student150	75	45	72	53
student151	81	50	61	50
student153	33	25	63	25
student154	63	58	74	35
student155	75	43	54	55
student156	37	24	57	13
student157	75	64	47	42
student158	64	49	31	36
student159	18	10	22	21
student160	64	44	66	38
student162	78	35	27	50
student164	37	40	52	24
student168	51	39	66	43
student169	63	61	85	63
student170	31	39	55	22
student173	30	39	34	23
student175	55	71	73	58

proširena eksperimentalna grupa				
	inicijalni test	prvi test provjere	drugi test provjere	završni test
student1	60	50	50	55
student2	70	76	44	77
student4	52	59	31	43
student5	69	55	39	53
student6	78	73	67	68
student7	72	50	41	54
student8	90	59	43	60
student9	61	24	21	43
student10	84	63	53	71
student11	69	54	33	55
student12	31	11	31	23
student13	54	51	50	65
student14	42	26	50	40
student15	57	28	27	46
student16	55	48	36	53
student18	81	19	39	58
student19	55	30	38	33
student21	69	50	42	55
student22	49	38	35	55
student23	54	29	35	55
student25	57	56	34	47
student26	72	38	45	58
student27	75	53	53	60
student28	64	53	39	46
student29	66	18	34	23
student30	66	26	29	35
student31	69	65	51	60
student33	58	49	44	53
student34	73	63	60	68
student36	51	35	26	33
student37	36	34	53	41
student38	39	21	35	51
student39	31	24	53	47
student40	60	66	57	62
student41	55	56	27	49
student42	55	53	61	80
student43	43	24	41	53
student44	49	61	66	65
student45	42	66	50	59
student46	64	85	69	71
student48	69	61	80	63
student49	66	38	35	60
student50	54	30	28	41
student51	54	59	53	66
student52	37	55	39	61
student53	67	56	35	61
student54	70	58	28	62
student55	78	50	38	52
student57	49	46	39	41

student58	34	43	43	44
student59	27	66	42	38
student61	52	41	45	54
student62	51	34	44	48
student63	54	41	23	43
student64	63	29	42	59
student65	37	59	57	50
student66	39	23	53	43
student67	37	75	71	62
student68	45	34	40	32
student70	37	35	37	19
student71	75	38	45	49
student72	39	28	54	44
student73	45	41	28	40
student76	37	33	36	12
student77	82	59	55	55
student78	42	25	32	24
student79	81	50	62	59
student81	84	40	73	61
student87	46	20	39	31
student88	46	26	37	39
student89	30	36	37	31
student90	39	25	14	41
student91	60	50	68	62
student93	51	38	46	56
student94	39	30	44	32
student95	52	40	59	50
student96	70	66	75	55
student99	67	60	36	42
student100	84	69	46	53
student101	63	56	46	41

ekvivalentna eksperimentalna grupa				
	inicijalni test	prvi test provjere	drugi test provjere	završni test
student4	52	59	31	43
student6	78	73	67	68
student7	72	50	41	54
student9	61	24	21	43
student12	31	11	31	23
student13	54	51	50	65
student14	42	26	50	40
student15	57	28	27	46
student16	55	48	36	53
student25	57	56	34	47
student26	72	38	45	58
student27	75	53	53	60
student28	64	53	39	46
student33	58	49	44	53
student34	73	63	60	68
student36	51	35	26	33
student37	36	34	53	41
student39	31	24	53	47
student40	60	66	57	62
student42	55	53	61	80
student43	43	24	41	53
student44	49	61	66	65
student45	42	66	50	59
student46	64	85	69	71
student52	37	55	39	61
student58	34	43	43	44
student59	27	66	42	38
student64	63	29	42	59
student65	37	59	57	50
student66	39	23	53	43
student67	37	75	71	62
student68	45	34	40	32
student71	75	38	45	49
student72	39	28	54	44
student73	45	41	28	40
student79	81	50	62	59
student89	30	36	37	31
student91	60	50	68	62
student93	51	38	46	56
student101	63	56	46	41