

PRIRODOSLOVNO-MATEMATIČKI FAKULTET
SVEUČILIŠTE U SPLITU

Josipa Raos

EKSPERIMENTALNI DIZAJN U VREDNOVANJU UČINKOVITOSTI SUSTAVA E-UČENJA

DIPLOMSKI RAD

SPLIT, rujana, 2008.

Studijska grupa: Matematika i informatika

Predmet: Primjena računala u nastavi

EKSPERIMENTALNI DIZAJN U VREDNOVANJU UČINKOVITOSTI SUSTAVA E-UČENJA

DIPLOMSKI RAD

Student: Josipa Raos

Mentor: prof.dr.sc. Slavomir Stankov

Neposredni voditelj: mr.sc. Ani Grubišić

SPLIT, rujan, 2008.

Sadržaj:

1. UVOD.....	1
2. ISTRAŽIVAČKI DIZAJNI.....	3
2.1. TEMELJNI KONCEPTI ISTRAŽIVAČKIH DIZAJNA	4
2.2. EKSPERIMENTALNI DIZAJN	4
2.3. KVAZI-EKSPERIMENTALNI DIZAJN	7
2.4. VALJANOST EKSPERIMENTA	11
3. EKSPERIMENTALNI DIZAJN	15
3.1. POSTAVLJANJE I ETAPE EKSPERIMENTA.....	20
3.2. PRINCIPI EKSPERIMENTALNIH DIZAJNA	24
3.3. EKSPERIMENTALNA POGREŠKA	27
3.2.1. <i>Analiza pogreške</i>	29
4. VREDNOVANJE UČINKA UČENJA I POUČAVANJA UZ POMOĆ SUSTAVA XTEX-SYS.....	31
4.1. SUSTAV XTEX-SYS	31
4.2. PRIPREMA EKSPERIMENTA.....	32
4.3. TIJEK ISTRAŽIVANJA.....	32
4.4. ANALIZA REZULTATA ZA ŠESTI RAZRED	34
4.5. ANALIZA REZULTATA ZA OSMI RAZRED	37
4.6. ANALIZA ANKETA DANIH EKSPERIMENTALNIM GRUPAMA.....	40
4.6.1. <i>Analiza ankete očekivanja</i>	40
4.6.2. <i>Analiza završne ankete</i>	44
5. ZAKLJUČAK.....	48
6. LITERATURA	49
7. PRILOZI.....	50
A. INICIJALNI TEST 6.R.	51
B. INICIJALNI TEST 8.R.	52
C. MEDUTEST 6.R.	53
D. MEDUTEST 8.R.	55
E. ZAVRŠNI TEST 6.R.	57
F. ZAVRŠNI TEST 8.R.	60
G. ANKETA OČEKIVANJA	62
H. ZAVRŠNA ANKETA.....	64

1. UVOD

Čovjek svijet oduvijek doživljava kao uređeni sustav u kojem vlada red uzroka i posljedica. Prirodno nam se čini da se u svakodnevnom ponašanju, mišljenju i govoru izražavamo u kategorijama zavisnosti i uzročnosti kao da je to samo po sebi razumljivo. U znanosti je dokazivanje uzročno-posljedičnih veza jedna od temeljnih zadaća, ali, kao što ćemo vidjeti, ne ostvaruje se nimalo jednostavno. Upravo zbog te teškoće u dokazivanju uzročno-posljedičnih veza nastao je veliki broj istraživačkih metoda kojima se nastoji objasniti ili, što je još važnije, dokazati odnos uzroka i posljedica određenih fenomena. Tim istraživačkim metodama, odnosno istraživačkim dizajnima se namjerno, u strogo kontroliranom okruženju izaziva neka pojava u svrhu opažanja i mjerenja. Glavna svrha je da se provjeri djeluje li određeni faktor ili varijabla na neki jasno definirani fenomen.

Iako su *istraživački dizajni* jako raznoliki i mnogobrojni, ipak ih možemo klasificirati u tri osnovne skupine: eksperimentalni dizajn, kvazi-eksperimentalni dizajn i ne-eksperiment. *Eksperimentalni dizajn* je najsnažniji od svih istraživačkih dizajna i osigurava najbolje rezultate zbog toga što su subjekti slučajnim odabirom dodijeljeni eksperimentalnoj i kontrolnoj grupi. Iako je on najjači dizajn njega je veoma teško provesti zbog toga što nije uvijek praktičan, etičan ili čak moguć postupak slučajnog odabira ljudi za izradu eksperimentalnih i kontrolnih grupa. Drugi najbolji izbor bi bio kvazi-eksperimentalni dizajn. Razlika između njega i eksperimentalnog dizajna je što u *kvazi-eksperimentalnom* dizajnu subjekti nisu naizmjenično podijeljeni grupama. Jako je bitno da istraživač bude svjestan posebnih problema koji mogu proizići iz toga i trebali bi poduzeti određene korake za rješavanje tog problema. Najslabiji oblik dizajna je ne-eksperiment, no to ne znači da je on i beznačajan. To je jedan od najčešćih oblika istraživanja, i za neka pitanja, posebno opisnog tipa, je veoma jak dizajn.

U drugom poglavlju ćemo osim ove osnovne klasifikacije istraživačkih dizajna opisati i način na koji se u literaturi eksperimentalni dizajn dijeli na deset različitih podtipova, a kvazi-eksperimentalni dizajni u dvije velike skupine od kojih svaka opet sadrži različite podtipove dizajna. Istraživač koji izvodi eksperiment je dužan organizirati svoj eksperiment na način da pomoću njega može, ne samo utvrditi učinak na zavisnu varijablu, nego i dokazati da su te posljedice rezultat djelovanja baš određene varijable, a ne nekog drugog faktora. Kontrola eksperimenta uključuje pažljivu analizu i određivanje koji faktori mogu utjecati na rezultate eksperimenta. Velik broj *parazitarnih faktora* se može neutralizirati eliminiranjem, a u suprotnom se treba poslužiti određenim metodama kojima se oni mogu držati pod kontrolom. Mnogo je faktora koji mogu ograničavati *valjanost* eksperimenta. Valjanost je povezana s dizajnom i istraživač može utjecati na učinke faktora u eksperimentu mijenjajući eksperimentalni dizajn u onaj koji sprječava ugrožavanje valjanosti. Razlikujemo unutarnju valjanost i vanjsku valjanost eksperimenta. Kažemo da je proučavanje unutarnje valjanosti kada je rezultat ili učinak na zavisnoj varijabli pripisan nezavisnoj varijabli i nijednom drugom faktoru, dok pod vanjskom

valjanošću smatramo da se rezultati istraživanja sastoje od uzorka populacije koja uistinu predstavlja cijelu populaciju.

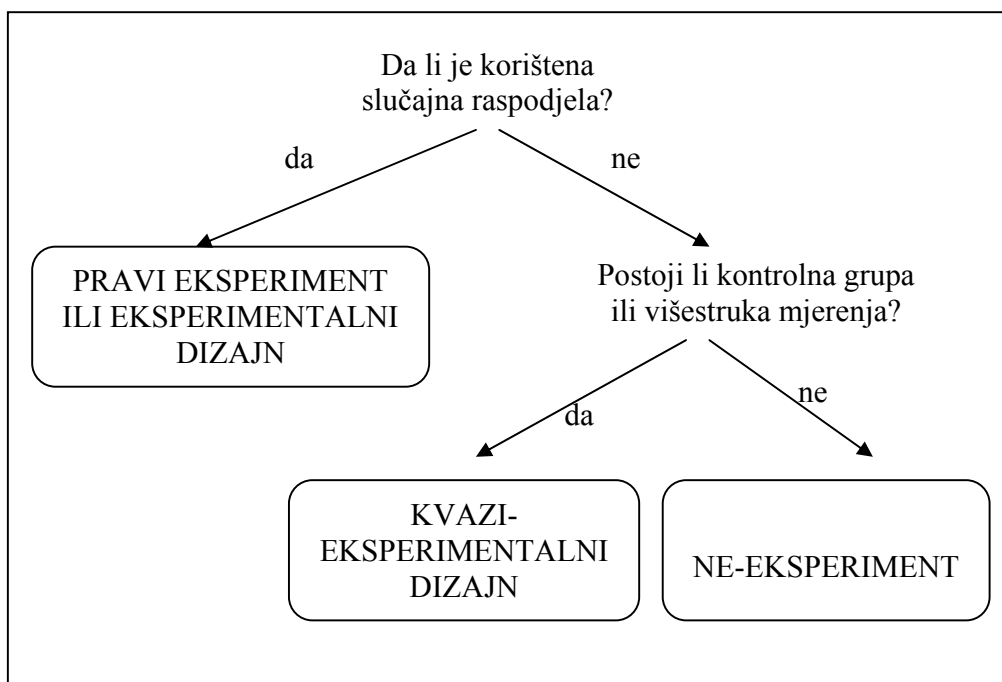
U trećem poglavlju ćemo detaljno opisati provođenje eksperimentalnog dizajna. U njemu namjerno, u strogo kontroliranim uvjetima izazivamo neku pojavu u svrhu njenog opažanja i mjerenja. Glavna svrha eksperimenta je da se provjeri djeluje li određeni faktor ili varijabla na jasno definirani fenomen. Upravo taj fenomen koji je predmet proučavanja nazivamo zavisnom varijablom, dok faktor koji se u eksperiment namjerno unosi i mijenja da bi se provjerilo utječe li on na zavisnu varijablu, nazivamo nezavisna varijabla. Objasniti ćemo, na raznim primjerima, što je to *eksperimentalna pogreška* i na koji se način analizira, odstupanje, ponavljanje i onemogućavanje. Pokazati ćemo čemu služe eksperimentalna i kontrolna grupa, te što se misli pod slijepi i dvostruko slijepi pokus. Naravno, ukoliko želimo eksperimentom proizvesti smislene i korisne rezultate, trebalo bi uložiti mnogo brige i planiranja. Stoga ćemo ovom poglavlju razmotriti još i način na koji bi se trebalo provesti planiranje koje je potrebno da bi definirali i pokrenuli službeni eksperiment.

U četvrtom poglavlju opisujemo dva eksperimentalna istraživanja provedena u svrhu ispitivanja vrednovanja učinka učenja i podučavanja uz pomoć sustava xTeX-Sys. Definirali smo varijantu *eksperimentalnog dizajna s paralelnim grupama uz parcijalno ispitivanje stanja* kod kojeg postoje dvije grupe ispitanika kontrolna i eksperimentalna. Kontrolna skupina se poučavala na tradicionalni način, a eksperimentalna skupina uz pomoć inteligentnog tutorskog sustava. Osim inicijalnog i završnog stanja ispitanika, koje se mjeri inicijalnim i završnim testom, izračunali smo i parcijalnu veličinu učinka pomoću rezultata međutesta. Za vrednovanje učinka sustava koristili smo programski alat EVEDIN (EValuation of EDucational INfluence) [GRUB2007]. EVEDIN omogućuje automatsko upravljanje kontrolnom i eksperimentalnom grupom, izradu i provedbu svih testova, njihovo ispravljanje i ocjenjivanje, te računanje veličine učinka eksperimentalnog faktora.

2. Istraživački dizajni

Postoji više različitih metoda, odnosno istraživačkih dizajna kojima namjerno, u strogo kontroliranom okruženju izazivamo neku pojavu u svrhu opažanja i mjerenja. Glavna svrha svih istraživačkih dizajna je da se provjeri djeluje li određeni faktor ili varijabla na neki jasno definirani fenomen.

Istraživačke dizajne dijelimo na tri široka područja, eksperimentalni dizajn, kvazi-eksperimentalni dizajn i ne-eksperiment. Dizajn je eksperimentalan ako su subjekti dodijeljeni eksperimentalnoj i kontrolnoj grupi slučajnim odabirom. Osnovna prednost eksperimentalnog dizajna nad drugim oblicima znanstvenih istraživanja je mogućnost *slučajne raspodjele* (eng. randomization) subjekata u grupe. Slučajni odabir nam omogućuje dugoročnu kontrolu varijabli koje nisu izravno uključene u istraživanje. Za razliku od eksperimentalnog, u kvazi-eksperimentalnom, odnosno korelacijskom dizajnu, se ne manipulira varijablama kako bi se pratio njihov učinak, već se utvrđuju veze koje postoje u prirodnim uvjetima. Stoga se na temelju korelacije (povezanosti) ne može zaključivati o uzročno-posljedičnoj povezanosti. Osim povezanosti dviju ili više varijabli na temelju korelacijskih nacrti istraživanja moguće je zaključivanje, odnosno predviđanje o vjerojatnom stanju neke varijable na temelju poznavanja druge. Iako kod kvazi-eksperimentalnog dizajna subjekti nisu dodijeljeni grupama slučajnim odabirom, još uvijek postoji kontrolna grupa. Karakterizira ga nemogućnost raspodjele ispitanika



Slika 2.1. Klasifikacija eksperimentalnih dizajna

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). Dakle, dizajn je kvazi-eksperimentalan ako subjekti nisu dodijeljeni grupama slučajnim odabirom, ali još uvijek postoji kontrolna grupa. U slučaju da kontrolna grupa ne postoji tada govorimo o ne-eksperimentu. Ne-eksperiment je općenito najslabiji kada govorimo o *unutarnjoj valjanosti* (eng. internal validity), no to ne znači da je on i beznačajan. To je jedan od najčešćih oblika istraživanja, i za neka pitanja, posebno opisnog tipa, je veoma jak dizajn.

2.1. Temeljni koncepti istraživačkih dizajna

Eksperimentalne dizajne možemo klasificirati postavljajući neka jednostavna pitanja kao što je to prikazano na **Error! Reference source not found.**. Dakle, prvo što se trebamo zapitati je da li postoji slučajan odabir. Ukoliko postoji zaključujemo da je dizajn pravi eksperiment ili eksperimentalni dizajn, a ukoliko ne postoji slučajan odabir, pitamo se da li postoji kontrolna grupa ili više provedenih mjerenja. Ukoliko je odgovor da, zaključujemo da se radi o kvazi-eksperimentalnom dizajnu, a ako je odgovor ne tada je riječ o ne-eksperimentu.

2.2. Eksperimentalni dizajn

Dakle, kao što smo već rekli, dizajn je eksperimentalan ako su subjekti dodijeljeni eksperimentalnoj i kontrolnoj grupi slučajnim odabirom. Cook i Campbell [COOK1979] su u svom radu naveli deset različitih tipova eksperimentalnog dizajna. Zabilježili su da se kontrolnu grupu može tretirati na dva načina, može biti grupa koja radi po standardnom postupku, ili se na nju uopće ne mora primijeniti nikakav tretman.

Klasični eksperimentalni dizajn

Slučajan odabir subjekata u kontrolnu i eksperimentalnu grupu u metodi klasičnog eksperimentalnog dizajna je pogodan za ANOVA analizu varijance. Kod klasičnog eksperimentalnog dizajna razlikujemo *dizajn između subjekata* (eng. between subjects design) i *dizajn unutar subjekata korištenjem ponavljanih mjerenja* (eng. within subjects (repeated measures) design).

Kod dizajna između subjekata, vrijednosti zavisne varijable za jednu grupu sudionika (na primjer, ako promatramo skupinu sudionika koja već ima iskustvo s određenim proizvodom) se uspoređuju s vrijednostima za drugu grupu sudionika (na primjer, grupa sudionika koja nikada prije nije koristila taj proizvod.).Ovaj dizajn karakteriziraju različiti subjekti u svakoj grupi u eksperimentu. Svaki subjekt je izložen samo jednom tretmanu i usporedbe se rade između

njihovih reakcija i učinaka. Istraživač se oslanja na slučajni odabir subjekata u eksperimentalnu i kontrolnu grupu za nemjerljive varijable, iako je ponekad slojevitost subjekata namještena da osigura proporcionalnost na određenoj ključnoj varijabli. Razlikujemo više vrsta dizajna između subjekata:

- 1) *Faktorski dizajni* (eng. factorial designs) koriste kategorične nezavisne varijable za uspostavljanje grupa. Na primjer, u dva faktorska oblikovanja nezavisne varijable mogu biti tip informacija (fikcija, ne fikcija) i tip medije (televizija, tisak, Internet), stvarajući šest kategorija. Svaki jednak broj subjekata bi bio određen slučajnim odabirom na svaki od šest mogućih grupa. Tada bi se subjekti mogli mjeriti na informacijskom nivou. Nulta posljedica će biti naznačena prosječnim rezultatom koji je isti za svih šest grupa faktorskog oblikovanja. Nejednaki rezultati bi označavali glavni učinak informacijskog ili medijskog tipa, i/ili međusobni utjecaj oba tipa.
- 2) *Dizajn potpunog-križanja* (eng. fully-crossed) ako postoji studijska grupa za svaku moguću kombinaciju faktora, odnosno nezavisnih varijabli. Ako dizajn izostavlja neke grupe nazivamo ga *nepotpunim faktorskim dizajnom* (eng. incomplete factorial design). Njega koristimo ako neke kombinacije vrijednosti faktora nisu od teoretskog interesa.
- 3) *Dizajn nasumice odabranih blokova* (eng. randomized block designs) razvrstava subjekte u slojeve i faktorski dizajn se provodi za svaki sloj. Ovaj se postupak provodi kad se istraživač boji parazitalnih faktora koji trebaju biti kontrolirani, ili ako postoje drugi olakotni strukturni faktori koji imaju prednost. Blokirajuće varijable koje primjerak dijele na slojeve su faktori koje ćemo razmotriti da budu kontrolne varijable, a ne nezavisne, kao što bi bile jednostavnom faktorskom dizajnu. Nasumice odabran blok dizajn nam kontrolira učinke glavnih faktora i njihova međudjelovanja, te blokirane varijable.

Za razliku od dizajna između subjekata, kod dizajna unutar subjekata je na istu grupu subjekata primijenjeno više od jednog tretmana. Kako su subjekti isti za svaki tretman, nezavisne varijable su same sebi kontrolne. Kao primjer dizajna između subjekata navesti ćemo istraživanje o tome kako različite vrste vježbi utječu na memoriju [HALL1998]. Upotrijebit ćemo dva tretmana: aerobnu vježbu i anaerobnu vježbu. U uvjetima aerobne vježbe ispitanici će trčati u mjestu pet minuta, a zatim će riješiti test iz memorije, dok će u uvjetima anaerobne vježbe ispitanici dizati utege pet minuta, nakon čega će im se dati test iz memorije jednake težine ali s drugim zadacima. Budući da se koristio dizajn unutar subjekata, svi sudionici su trčali u mjestu pet minuta i rješavali prvi test, nakon čega su svi dizali utege i zatim rješavali drugi test. Usporedba rezultata ta dva testa nam daje odgovor na pitanje koja vježba više pomaže memoriji.

Dvije su osnovne prednosti ovog dizajna, a to su statistička snaga i smanjenje pogreške koja proizlazi iz individualnih razlika. Temeljni nedostatak ovog dizajna nazivamo *efekt prenošenja* (eng. carryover effect). To općenito znači da sudjelovanje u jednim uvjetima može utjecati na rezultate u drugim uvjetima. Dvije osnovne vrste efekta prenošenja su vježba i umor.

Lutrijski dizajn

Lutrijski dizajni (eng. lottery design) se koriste kada se očekuje efekt lutrije, kao što u nekim zajednicama to slučaj kod dodjele učenika u stručne škole, ovaj dizajn eliminira zapreke

slučajnog odabira u društvenim znanostima, gdje se često razmatra neetični, ili čak ilegalni standardi da bi ponudili privilegije nekim građanima, ali ne i drugima.

Mandatni kontrolni dizajn

Mandatni kontrolni dizajn (eng. mandated control designs) se koristi kada se dogodi da su, kao u vojsci, kontrolni tipovi tretmana dovoljno visoki tako da se slučajno dodjeljivanje tretmanu i kontrolnom stanju prihvaća kao bit istraživanja.

Lista čekanja

Lista čekanja (eng. waiting list designs) se koristi kada potražnja nadmašuje opskrbu. Ovaj dizajn također dopušta primjenu tretmana samo dijelu građana. Budući da je davanje tretmana svima nemoguće, pribjegava se ograničenoj primjeni.

Ekvivalent dizajn vremenskog slijeda

Ekvivalent dizajn vremenskog slijeda (eng. equivalent time series designs) se koristi kada se tretman ne može primijeniti na sve ispitanike istovremeno. Tada se tretman privremeno primjeni na neke ispitanike, a zatim na druge. Kao na primjer kada svi zaposlenici neke firme trebaju proći obuku, ali kako ne mogu svi u isto vrijeme, podijele se u grupe pa u rotacijama različite grupe slušaju različita predavanja.

Dizajn prostornog odjeljivanja

Dizajn prostornog odjeljivanja (eng. spatial separation designs) se koristi kada su eksperimentalne grupe odvojene i nemaju međusobnu komunikaciju važnu za eksperiment. Na primjer, kada se sudjelovanje u menadžmentu pokušava sa eksperimentalnom grupom novih, slučajno dodijeljenih zaposlenika na jednoj lokaciji, a ne na drugoj.

Ovlaštena promjena/ dizajn nepoznatog rješenja

Ovlaštena promjena/ dizajn nepoznatog rješenja (eng. Mandated change/unknown solution designs) se koristi kada se zahtjeva promjena ali ne postoji jasno rješenje. Tada slučajna dodjela subjekata može pribjeći odobravanju. Npr. Kada želimo podijeliti studente u razrede sa alternativnim gradivom.

Dizajn razbijanja čvorova

Dizajn razbijanja čvorova (eng. Tie-breaking designs), u slučajevima kada je primanje tretmana zasnovano na temelju zasluga (eng. merit-based), kao na primjer u nekim akademskim slučajevima, dogodi se da postignuti rezultati u ispitima budu kriterij slučajne dodjele eksperimentalnoj i kontrolnoj grupi.

Dizajn indiferentne krivulje

Dizajn indiferentne krivulje (eng. Indifference curve designs). Atraktivnost tretmana se može ponekad prilagoditi do točke gdje će neki ljudi biti nepristrani, odnosno indiferentni, u izjavi da li žele primiti tretman ili ne, i upravo to može biti razlog slučajne podijele u kontrolnu i eksperimentalnu grupu.

Dizajn nove organizacije

Dizajn nove organizacije (eng. New organizations designs). Kada se uspostavi nova organizacija, kao na primjer obučavanje za određenu vrstu posla, prilike mogu omogućiti proizlaženje slučajne dodijele klijenata u kontrolnu i eksperimentalnu grupu.

2.3. Kvazi-eksperimentalni dizajn

Kvazi-eksperimentalni dizajn (eng. Quasi-experimental designs) je eksperimentalni dizajn koji ne susreće sve zahtjeve nužne za kontroliranje utjecaja stranih varijabli. Razvijen je da bi se uredio raznolik svijet polja istraživanja, gdje nije uvijek praktičan, etičan ili čak moguć postupak slučajnog odabira ljudi za izradu eksperimentalnih i kontrolnih grupa. Ako želimo dobiti uzročne izjave, moramo provesti pravi eksperiment. No, u slučaju da nismo u mogućnosti koristiti naizmjenično dodjeljivanje subjekata za dizajniranje pravog eksperimenta, kvazi-eksperiment je drugi najbolji izbor. "Usprkos tome, naizmjenično dodjeljivanje subjekata eksperimentalnoj i kontrolnoj grupi je ponekad nemoguće, posebno u terenskim istraživanjima. Takvi eksperimenti su nazvani kvazi-eksperimenti, kako ih nazivaju Campbell i Stanley da bi naznačili da naizmjenično dodjeljivanje subjekata eksperimentalnoj i kontrolnoj grupi nije postignuto. Svakako bi istraživač trebao biti svjestan posebnih problema koji mogu proizići iz toga što subjekti nisu naizmjenično podijeljeni grupama i trebali bi poduzeti određene korake za rješavanje tog problema." [BORG1998]. "Kvazi-eksperimenti su istraživački dizajni koji nemaju naizmjenično dodijeljen tretman i usporedne grupe. Umjesto usporedbi između tretmana i ne tretmana, uvjeti, odnosno okolnosti uvijek moraju biti napravljene s ne ekvivalentnim grupama ili s istim subjektima koji pristupaju tretmanu. Ako niste u mogućnosti naizmjenično dodijeliti ljude ili grupe uvjetima tretmana, gubite na sposobnosti kontrole što se kome dešava. Još uvijek možete promatrati što se kada dešava, odlučujući što i kada mjeriti, možete dizajnirati jedan ili nekoliko kvazi-eksperimenata. Gubite kontrolu, kada ne koristite naizmjenično dodjeljivanje, ali još možete provesti istraživanje i analizu efekta uzroka veza bez pravog eksperimenta." [KIDD1980]. Cook i Campbell [COOK1979] u svom radu o istraživačkim dizajnim dijelje kvazi-eksperimentalne dizajne u dvije velike skupine, na dizajne *ne ekvivalentnih kontrolnih grupa* (eng. nonequivalent control group designs) i na *isprekidani dizajn vremenskog slijeda* (eng. interrupted time series).

Dizajni ne ekvivalentnih kontrolnih grupa

Dizajn je kvazi-eksperimentalan ako subjekti nisu dodijeljeni grupama slučajnim odabirom, ali još uvijek postoji kontrolna ili usporedna grupa. Iako subjekti nisu slučajno dodijeljeni, oni su ili slučajno odabrani kao primjeri ili su svi mjerodavni slučajevi. Cook i Campbell [COOK1979] navode deset oblika dizajna s ne ekvivalentnim kontrolnim grupama. U svakom slučaju, zbog ne ekvivalencije kontrolnih grupa vjerojatnost da se valjanost poremeti je mnogo veća nego u slučajnom dizajnu i istraživač bi trebao razmotriti sve moguće tipove prijetnji za valjanost. Nabrojati ćemo nekoliko vrsta dizana ovoga tipa.

1. *Dizajn s jednom grupom i post-testom* (eng. one-group posttest-only design). Zbog ne postojanja pred testa ili usporedne grupe, nemoguće je doći do valjanih zaključaka o efektu tretmana, baziranih samo na rezultatima završnog testa. Promjene u zavisnoj varijabli mogu ovisiti o tretmanu ili pak mogu biti posljedica bilo koje prijetnje za valjanost, kao što su povijest, očekivanja provoditelja eksperimenta itd.
2. *Dizajn sa post-testom i ne ekvivalentnom usporednom grupom* (eng. posttest-only design with nonequivalent comparison groups design). Nemoguće je doći do valjanih zaključaka o učinku eksperimenta osnovano samo na informacijama post-testa na dvije ne ekvivalentne grupe s obzirom da učinci mogu biti posljedica zahvaljujući tretmanu ili ne ekvivalenciji između grupa.
3. *Dizajn sa post-testom i predvidljivim redosljedom međudjelovanja* (eng. posttest-only design with predicted higher-order interactions). Ponekad se očekivanja učinka eksperimenta pomiješaju sa trećom varijablom. Umjesto očekivanja da eksperimentalna grupa bude na višem stupnju zavisnosti, očekujemo da subjekti budu na višem stupnju ako su u gornjoj polovici treće varijable, a na nižem ako su u donjoj polovici. Na primjer, treniranje može voditi većoj produktivnosti za visoku izobrazbu zaposlenika, ali ne i za nisku izobrazbu zaposlenika na istim zadacima. Međudjelovanje stvara dva ili više očekivanja, usporedna sa jednostavnim jednim očekivanjem dizajna koji sadrži jednu grupu i post-test. Zbog toga što postoji više očekivanja, postoji i veća kontrola eksperimentalnog učinka. Svakako, ovaj dizajn je tema za moguće izazove za valjanost i zahvaljujući takvim faktorima kao što su povijest (subjekti koji su visoko u obrazovanju imaju različita očekivanja) protu-argumenti trebaju biti više uključeni u međudjelovanje što može donekle dovesti do manjka vjerojatnosti.
4. *Dizajn s jednom grupom, pred-testom i post-testom* (eng. one-group pretest-posttest design) Ovo je čest dizajn u socijalnim znanostima. Podvrgnut je takvim prijetnjama za valjanost kao što su povijest (događaji koji se dešavaju između pred testa i post testa), sazrijevanje (promjene kod subjekata), regresija prema značenju (težnja da se krajnosti povrate na uobičajeno), testiranje (efekt učenja na post test) itd.
5. *Dizajn s dvije grupe, pred testom, post testom koji koristi netretiranu kontrolnu grupu* (eng. two-group pretest-posttest design using an untreated control group). S obzirom da grupe nisu ekvivalentne još uvijek postoji mogućnost odabira (promatrane promjene su nastale zahvaljujući odabiru subjekata, kao što su rad sa više motiviranih volontera u grupi na koju se primjenjuje tretman). Mnogo ovisi o posljedici. Na primjer, ako eksperimentalna grupa počne ispod usporedne grupe, a nakon tretmana završi poviše, tada postoje jači zaključci o učinku tretmana nego u slučaju da se obje grupe uzdignu tijekom

eksperimenta. Za ovaj dizajn se preporuča da ima više od jednog pred-testa. Veći broj tih pred-testova pomaže u uspostavi pravca učinkovitosti u izvedbi u obje eksperimentalne i jednoj kontrolnoj grupi. Tretman bi trebao iznijeti na vidjelo promjene u tom pravcu za eksperimentalnu grupu ali ne i za kontrolnu grupu.

6. *Dizajn dva pred testa* (eng. Double pretest designs) Dizajn s pred-testom i post-testom se može pojačati tako da se uvedu dva ili više pred-testnih mjerenja. To nam može otkriti postoji li pravac u podacima koji ne ovisi o tretmanu čiji se učinak mjeri u završnom testu. Uočavajući da je učinak završnog testa poviše tog pravca kontroliramo opasnosti za proučavanje valjanosti.
7. *Dizajn uklonjenog tretmana s pred-testom i post-testom* (eng. removed-treatment pretest-posttest design). U nekim slučajevima je moguće, ne samo uvesti eksperiment, nego ga i poništiti. Ako zavisne varijable poprimaju veće vrijednosti poslije tretmana i onda se smanjuju nakon uklanjanja tretmana, moguće je izvesti neke zaključke o učinku tretmana. Naravno, ako se varijable povećavaju nakon tretmana, može se desiti da se one poslije smanje same od sebe, bez obzira na tretman. Zbog toga Cook i Campbell [COOK1979] preporučuju najmanje dva završna testa, jedan nakon tretmana a drugi nakon uklanjanja tretmana, u svrhu uspostavljanja linije učinka nakon tretmana.
8. *Dizajn ponavljanog tretmana* (eng. repeated-treatment design). U ovom dizajnu je očekivano da zavisne varijable poprime veće vrijednosti nakon tretmana, da se te vrijednosti smanje nakon uklanjanja tretmana, te opet povećaju sa vraćanjem tretmana. No u slučaju i da dođe do ovog efekta zaključak se ne može uzeti sa stopostotnom sigurnošću jer se smanjenje i povećavanje vrijednosti može pripisati i drugim čimbenicima. Također se može lakše desiti da subjekti postanu svjesni istraživačevih očekivanja u tom dizajnu pa mogu težiti ostvarenju tih očekivanja i tako ugroziti zaključke.
9. *Dizajn s uključenim replikacijama* (eng. switching replication design). U ovom dizajnu postoje dvije usporedne grupe i tri mjerenja. Na obje grupe se provede pred test. Tretman se provodi nad jednom grupom, ali ne kontrolnom i uzimaju se mjerenja prvog pred-testa. Tada se tretman provodi nad kontrolnom grupom i uzimaju se mjerenja drugog pred-testa.
10. *Dizajn sa suprotnim tretmanom, pred-testom post-testom i ne ekvivalentnim usporednim grupama* (eng. reversed-treatment pretest-posttest nonequivalent comparison groups). U ovom dizajnu ne ekvivalentna kontrolna grupa prima suprotan tretman od eksperimentalne grupe. Za očekivati je da će se postići povećanje učinka kod eksperimentalne grupe, a smanjiti učinak kod kontrolne grupe. Cook i Campbell [COOK1979] predlažu dodavanje i grupe na koju se neće primijeniti tretman. Višestruki pred-testovi bi unaprijedili ovaj dizajn ukazujući na postojanje pravca u eksperimentalnoj i kontrolnoj grupi.

Isprekidani dizajni vremenskog slijeda

U raspravi o dizajnim ne ekvivalentnih kontrolnih grupa smo predložili da se dizajni koji sadržavaju pred-test i post-test mogu usavršiti povećanjem broja pred-testova. Dizajn vremenskog slijeda podliježe pred-testiranju i post-testiranju grupe subjekata u različitim

intervalima. Serije promatranja prije i poslije mogu pružiti bogate informacije o razvoju subjekata. Zbog mjerenja na nekoliko točaka u određenom vremenskom periodu i slijedu će pružiti mnogo realniju sliku postignuća. Broj testova kojim se podvrgavaju subjekti varira od dizajna do dizajna. Kada dođe do prekida između testova u svrhu ojačanja tretmana kroz prošireni vremenski period govorimo o isprekidanom dizajnu vremenskog slijeda. Cook i Campbell [COOK1979] navode popis od šest isprekidanih dizajna vremenskog slijeda koji teže sugestiji da imaju više pred testova i post testova.

1. *Jednostavni isprekidani dizajn vremenskog slijeda* (eng. simple interrupted time series design) Ovo je dizajn s jednom grupom, pred testom i post testom proširen s više pred i post testova. Rezultati dobiveni višestrukim pred-testiranjem se mogu usporediti s rezultatima dobivenim višestrukim post-testovima ne bi li se ocijenilo da li je poboljšanje rezultata posljedica efekta sazrijevanja koji bi svakako vodio prema poboljšanju. Budući da ne postoji kontrolna grupa istraživač ne može procijeniti ostale faktore koji utječu na valjanost.
2. *Isprekidani dizajn vremenskog slijeda sa ne ekvivalentnom kontrolnom grupom koja ne prima tretman* (eng. interrupted time series with a nonequivalent no-treatment comparison group). Ovo je u biti dizajn s dvije grupe, pred testom i post testom koji koristi ne tretiranu kontrolnu grupu. Postojanje kontrolne grupe, čak i ne ekvivalentne može dovesti do opasnosti za valjanost, ali se obično te opasnosti pojavljuju u kompliciranjem obliku i stoga se lakše dokazuju.
3. *Isprekidani dizajn vremenskog slijeda sa ne ekvivalentnim zavisnim varijablama* (eng. interrupted time series with nonequivalent dependent variables). Ovo je dizajn ne ekvivalentnih zavisnih varijabli s višestrukim pred i post testovima. Svrha ovog dizajna je da se nađu zavisne varijable za koje se misli da bi mogle utjecati na tretman. Cook i Campbell [COOK1979] nam daju primjer utjecaja alkotesta kojeg policija provodi samo vikendom kada su barovi otvoreni. Zavisna varijabla koja nas ovdje zanima je učestalost nesreća vikendima. Povezane zavisne varijable su učestalost nesreća kroz noći tokom tjedna kada su barovi otvoreni i noći tokom tjedna kada nisu otvoreni. Za očekivati je da će efekt alkotesta biti veći vikendom nego ostalim noćima. Ostala objašnjenja za manji broj nesreća (npr. sigurnija auta, veće kazne) objašnjavaju smanjenje nesreća ne samo vikendom, nego i ostalim danima.
4. *Isprekidani dizajn vremenskog slijeda sa uklonjenim tretmanom* (eng. Interrupted time series with removed treatment) Ovaj dizajn je dizajn sa uklonjenim tretmanom, pred-testom i post-testom samo što ovaj ima više pred i post testova, uključujući i one između provođenja tretmana i ne provođenja tretmana. U ovom dizajnu je prijetnja od povijesti značajno smanjena.
5. *Isprekidani dizajn vremenskog slijeda s višestrukim replikacijama* (eng. Interrupted time series with multiple replications). Ovo je jednostavno isprekidani dizajn vremenskog slijeda sa uklonjenim tretmanom, s iznimkom da se tretman i uklanjanje tretmana ponavljaju više puta. Ovo je jak dizajn, ali ga okolnosti rijetko dopuštaju
6. *Isprekidani dizajn vremenskog slijeda sa uključenim replikacijama* (eng. Interrupted time series with switching replications) U ovom dizajnu se koriste dvije grupe, od kojih svaka služi ili kao kontrolna ili kao eksperimentalna grupa na koju se primjenjuje tretman, kroz

višestruka replikacija tretmana i njegovog uklanjanja. Ovaj dizajn zahtjeva veoma visoku kontrolu subjekata, ali je veoma jak dizajn s obzirom na prijetnje za valjanost.

2.4. Valjanost eksperimenta

Istraživači koji su razvili eksperimentalno istraživanje su opisali nekoliko faktora koji mogu ograničavati valjanost eksperimenta. Valjanost eksperimenta je povezana sa dizajnom eksperimenta i istraživač može utjecati na učinke faktora u eksperimentu mijenjajući eksperimentalni dizajn u onaj koji sprječava ugrožavanje valjanosti.

Opasnosti koje prijete valjanosti eksperimenta možemo podijeliti u dva opsežna područja: ugrožavanje unutarnje valjanosti i ugrožavanje vanjske valjanosti.

Istraživanja ili eksperimenti imaju unutarnju valjanost ako je posljedica dobivena kao funkcija varijabli koje su mjerene, kontrolirane ili manipulirane u istraživanju. [ENDE2008]

Vanjska valjanost, je stupanj u kojem rezultati istraživanja na uzorku populacije vrijede za cijelu populaciju.

Unutarnja valjanost

Unutarnja valjanost je približna istina o zaključcima u odnosu na uzročni efekt ili uzročni odnos.[TROC2008] Proučavanje je unutarnje valjano kada je rezultat ili učinak na zavisnoj varijabli pripisan nezavisnoj varijabli i nijednom drugom faktoru. Koliko su ti drugi faktori dobro kontrolirani uvelike utječe na unutarnju valjanost eksperimenta. Unutarnja valjanost je važna samo u istraživanjima gdje se pokušava uspostaviti uzročni odnos. U većini opažачkih ili opisnih istraživanja ona nije važna. Za studije koji ocjenjuju učinke socijalnih programa ili intervencija, unutarnja valjanost se prva uzima u obzir. U tim kontekstima poželjno je biti u mogućnosti da zaključimo da taj program ili postupak dovodi do neke razlike. No, postoji puno razloga, osim postupka, koji dovode do boljih ili lošijih rezultata testa. Zbog toga se postavlja ključno pitanje u unutarnjoj valjanosti da li promatrane promjene mogu biti pripisane tretmanu, a ne drugim mogućim uzrocima (koje ponekad opisujemo kao „alternativna objašnjenja“ za posljedicu). Unutarnja valjanost zapravo znači da imamo dokaze o tome da upravo taj postupak koji smo proveli u istraživanju uzrokuje pojavu posljedice. Ona nam ne govori o tome da li je ona posljedica koju promatramo upravo i posljedica koju smo željeli.

Dakle, istraživači moraju biti svjesni faktora koji mogu utjecati na smanjenje unutarnje valjanosti i trebaju učiniti što god mogu da kontroliraju te opasnosti za unutarnju valjanost. Te opasnosti, ako ostanu zanemarene, mogu smanjiti unutarnju valjanost do te točke da svi rezultati postanu beznačajni i cijelo istraživanje pogrešno. Postoji osam glavnih prijetnji [XXXX2008] za unutarnju valjanost:

1. Povijest

Povijest se odnosi na bilo koji događaj izvan istraživačkih studija koji može mijenjati ili utjecati na učinak djelovanja subjekta. S obzirom da se istraživanje ne provodi izolirano, subjekti često osjete utjecaj različitih događaja u okolišu. Ti događaji mogu odigrati veliku ulogu u izvedbi eksperimenta, stoga ih se treba navesti. Jedan od načina na koji možemo spriječiti utjecaj tih događaja je da ih kontroliramo. S obzirom da je to često nemoguće pribjegavamo slučajnim postupcima u svrhu minimalizacije tog rizika, osiguravajući da se vanjski događaji koji se pojave u jednoj grupi, pojave i u drugoj.

2. Sazrijevanje.

Sazrijevanje igra veliku ulogu u dugoročnim studijima. Odnosi se na prirodne fiziološke i psihološke promjene koje mogu nastupiti sa starenjem. Ovo je naročito važno u djetinjstvu i mora se adresirati kroz podudaranje subjekata ili slučajan odabir. Na primjer, velik broj epizoda depresije se značajno smanjuje u roku od šest mjeseci čak i bez liječenja. Zamislite da provodimo testiranje novih lijekova protiv depresije. Ukoliko rezultati pokazuju značajno smanjenje depresije kod subjekata kroz šest mjeseci, možemo li uistinu zaključiti da lijek djeluje? Vjerojatno ne bi jer bi se nakon sazrijevanja pojavili slični rezultati i bez uzimanja lijekova.

3. Testiranje

Što su dulje izloženi nekoj aktivnosti, to je ljudi i uspješnije obavljaju. Tako i testiranje nije iznimka. Kada subjekti, pogotovo u eksperimentima sa jednom grupom, dobiju pred-test i nakon toga isti taj test kao završni test, šanse da će biti uspješniji drugi put zahvaljujući samo vježbi su veoma upitne. Iz toga razloga se u ovakvim situacijama preporučuje eksperiment s dvije grupe.

4. Statička regresija

Statistička regresija se uglavnom odnosi na istraživanja s ekstremnim rezultatima. To se odnosi na subjekte koji postignu veoma visoke ili veoma niske rezultate u testiranjima. Dakle, ako netko postigne 99% na testu, šanse da će mu rezultat biti manji na idućim testovima su mnogo veće nego da će postići bolje rezultate.

5. Instrumentacija

Ako se uređaj za mjerenje koji se koristi u istraživanju promijeni tijekom provođenja eksperimenta, tada će promjene u rezultatima prije biti povezane s uređajem nego s nezavisnom varijablom. Na primjer ako su pred test i završni test različiti, može se desiti da su razlike u rezultatima posljedica lakšeg završnog testa, a ne metode učenja. Zbog toga se preporuča da pred test i završni test budu identični ili bar jako slični po težini.

6. Selekcija

Selekcija ili odabir se odnosi na način na koji su subjekti odabrani da sudjeluju u istraživanju i na način na koji su dodijeljeni grupama. Ako postoje razlike između grupa i prije nego se eksperiment pokrene, te razlike će se nastaviti kroz eksperiment i na kraju se mogu pojaviti kao uzrok promjene u statističkoj analizi. Preporučuje se rješavanje tih razlika preko izjednačavanja subjekata ili slučajne podjele subjekata u grupe.

7. Očekivanja provoditelja eksperimenta

Istraživanje provodimo u svrhu podupiranja neke teorije pa se lako može desiti da je provoditelj eksperimenta pristran prema rezultatima koje želi dobiti. Zbog toga se preporučuje da eksperiment provede osoba koja nije upoznata sa tim koji su rezultati poželjni. Ovakav način pokusa nazivamo dvostruko slijepi pokus.

8. Smrtnost

Smrtnost ili izostajanje subjekata je problem koji se često javlja. Razlika smrtnosti između grupa može drastično utjecati na rezultate.

2.1. Prijetnje unutarnjoj valjanosti

Prijetnja unutarnjoj valjanosti	Kontroliranje opasnosti
Povijest	slučajan odabir, slučajna dodjela
Sazrijevanja	podudaranje subjekata, randomizacija
Testiranje	kontrolna grupa
Statistička regresija	
Instrumentacija	dosljednost
Odabir	slučajan odabir, slučajna dodjela
Očekivanja provoditelja eksperimenta	dvostruko slijepi pokus
Smrtnost	i propusta

Vanjska valjanost

Vanjska valjanost se odnosi na generalizaciju istraživanja. Drugim riječima, možemo li biti u potpunosti sigurni da se rezultati našeg istraživanja sastoje od uzorka populacije koja uistinu predstavlja cijelu populaciju. Opasnosti za vanjsku valjanost mogu dovesti do značajnih rezultata unutar uzorka, ali i biti uzrok nesposobnosti da se rezultati generaliziraju na cijelu populaciju. Navesti ćemo tri prijetnje za vanjsku valjanost.

1. Zahtijevanje karakteristika

Subjekti su često dobivali napomene o poželjnim rezultatima. Kada, na primjer, pitate niz pitanja o depresiji subjekti mogu postati svjesni hipoteza da su neki tretmani učinkovitiji u liječenju duševnih bolesti. Kada subjekti postanu svjesni mogu pribjeći ponašanju za koje vjeruju da se očekuje od njih. Da bi se smanjila ova prijetnja potrebno je osigurati da subjekti nikako nisu u mogućnosti da postanu svjesni željenih posljedica.

2. Hawthorneov efekt

Sličan je placebo učinku, istraživanja su dokazala da nečija prisutnost i gledanje uzrokuje promjenu ponašanja kod subjekata. Ako je ta promjena značajna možemo li biti sigurni da će se posljedica pojaviti i ako nitko ne gleda? Rješavanje tog problema može biti veoma

problematično. Pri tom nam pomaže uvođenje kontrolne grupe za mjerenje Hawthorneovog efekta kod subjekata koji ne primaju tretman. Dakle, kontrolna grupa se promatra isto kao i eksperimentalna i tako se mogu uočiti slične promjene u ponašanju obje grupe, te se tako smanjiti Hawthorneov efekt.

3. Efekti prenošenja ili redosljeda

Efekti prenošenja se odnose na redosljed u kojem se provodi neko istraživanje i može sadržavati veliku opasnost za vanjsku valjanost, naročito ako se koristi više tretmana. Na primjer, ako subjektima prva dva mjeseca damo lijekove, druga dva mjeseca terapiju i sljedeća dva mjeseca ih uopće ne liječimo moguće je da subjekti osjete najveće olakšanje nakon posljednje faze-ne tretmana. Postavlja se pitanje da li to znači da je ne davanje tretmana uspješnije nego liječenje lijekovima ili terapijom? Naravno da ne, vjerojatnije je da su posljedice prva dva tretmana prenesene na zadnju fazu, prividno prikazujući je kao najučinkovitiju.

2.2. Prijetnje za vanjsku valjanost

Prijetnje vanjskoj valjanosti	
Zahtijevanje karakteristika	slijepi pokus, uvođenje kontrolne grupe
Hawthorneov efekt	kontrolne grupe

3. Eksperimentalni dizajn

Eksperimentalni dizajn (eng. experimental design) je najsnažniji od svih istraživačkih dizajna. Prihvaćeno je mišljenje da je on «zlatni standard» koji je mjerilo za sve ostale istraživačke dizajne. On osigurava najbolje rezultate zbog toga što su subjekti (studenti, učitelji, razredi, škole...) slučajnim odabirom dodijeljeni eksperimentalnoj i kontrolnoj grupi. Samo preko slučajnog dodjeljivanja osoba koja provodi eksperiment (u daljnjem tekstu, istraživač), može biti uvjerena da su grupe doista usporedive i da promatrane razlike u posljedicama nisu rezultat nekih vanjskih faktora ili već postojećih razlika. Bitno je uočiti da se tretman ne primjenjuje na kontrolnu grupu, iako to nije nužno jer se na nju može primijeniti i standardni tretman, dok se tretman čija se učinkovitost istražuje primjenjuje na eksperimentalnu grupu.

U eksperimentalnom dizajnu namjerno, u strogo kontroliranim uvjetima izazivamo neku pojavu u svrhu njenog opažanja i mjerenja. Upravo to namjerno izazivanje pojave, poznavanje točnog mjesta i vremena događanja nam omogućava da se unaprijed pripremimo za opažanje. Poznavanje uvjeta pod kojim se neka pojava događa, pruža mogućnost i drugim istraživačima da repliciraju opažanja i tako kontroliraju točnost rezultata. Kontrola nad uvjetima predstavlja i mogućnost sistematskog mijenjanja prilika u kojima se neka pojava događa. Postupno mijenjanje jednog po jednog faktora, uz održavanje drugih faktora konstantnima, omogućuje istraživaču da iz mnoštva prilika identificira one koje su uvjet za nastanak određene pojave. U kontroliranim uvjetima moguće je izazvanu pojavu u tolikoj mjeri stabilizirati da se ona može i izmjeriti.

Glavna svrha eksperimenta je da se provjeri djeluje li određeni faktor ili varijabla na jasno definirani fenomen. Upravo taj fenomen koji je predmet proučavanja nazivamo zavisnom varijablom, dok faktor koji se u eksperiment namjerno unosi i mijenja da bi se provjerilo utječe li on na zavisnu varijablu, nazivamo nezavisna varijabla ili eksperimentalna varijabla. Ako se eksperimentalni dizajn dobro provede, s velikim naglaskom na ako, onda je eksperiment vjerojatno najjači dizajn u odnosu na unutrašnju valjanost. Upravo ta unutrašnja valjanost je središte svih uzročnih ili uzročno-posljedičnih zaključaka. Ona je veoma važna u proučavanjima koja pokušavaju uspostaviti uzročno-posljedičnu vezu. Želimo li utvrditi da li neki postupak uzrokuje neku posljedicu ili posljedice, onda nam treba da ta unutrašnja valjanost bude veoma visoka. U osnovi želimo ocijeniti sljedeću tvrdnju:

IF X THEN Y

odnosno:

ako je proveden dani postupak, onda se javlja i posljedica.

Nažalost, nije dovoljno samo pokazati da ako se postupak provede dolazi do posljedice. To je upravo zbog toga što postoji mnogo razloga koji utječu na javljanje posljedice. Da bi pokazali da postoji uzročna veza, moramo obratiti pozornost na dvije tvrdnje:

IF X THEN Y

AND

IF not X THEN not Y

odnosno:

posljedica se javlja ako provedemo postupak, a ako postupak nije proveden posljedica se neće javiti.

Ako možemo dokazati obje tvrdnje onda imamo važeći izdvojeni postupak iz svih drugih potencijalnih uzroka posljedice. Tada smo pokazali da kad je postupak aktualan javlja se i posljedica, a kad nije aktualan posljedica se ne javlja. To ukazuje na uzročnu učinkovitost postupaka. S obzirom da postupak ne možemo i provesti i ne provesti postavlja se pitanje kako izvesti te dvije radnje odjednom. Idealno bi bilo da imamo dvije grupe s istim uvjetima: iste ljude, isti kontekst, vrijeme itd. te da na jednu grupu primijenimo postupak, a na drugu ne. Tada bi lako došli do zaključka da li gornja tvrdnja vrijedi. Odnosno da ako je postupak dan, imamo posljedicu, ili ako nije dan, nemamo posljedicu. Očito nikad nećemo postići ovu hipotetičku situaciju. Ako damo postupak grupi ljudi, ne možemo ga istodobno ne dati! Postavlja se pitanje kako riješiti ovu očitu dilemu?

Možda bi samo trebali problem sagledati malo drugačije. Što ako bi napravili dvije grupe ili okruženja koja su slična što je više moguće? Ako bismo mogli biti sigurni da su dvije situacije usporedive onda bi mogli provesti postupak u jednom okruženju, to jest grupi (da bi dobili posljedicu) i ne provesti postupak u drugoj grupi (da vidimo da ne dolazi do posljedice). A ako su ta dva okruženja usporediva to bi bilo isto kao da smo prošli oba puta istovremeno. Upravo je to ono što eksperimentalni dizajn pokušava postići.

U osnovnom obliku eksperimenta stvaramo dvije grupe koje su “jednake,, jedna drugoj. Pri tom pod pojmom “jednako,, smatramo da su te dvije grupe usporedive. Jedna grupa (eksperimentalna) radi po postupaku a druga (kontrolna) ne. U svim drugim pogledima grupe tretiramo jednako. Grupe sadrže slične ljude, koji žive u sličnim okruženjima imaju slične prošlosti itd., i ako usporedimo posljedice u tim grupama onda razlike moraju proizići iz jedine stvari koja je različita među njima, a to je da je na jednu primijenjen postupak a na drugu ne.

Dakle, kako stvoriti grupe koje su jednake? Pristup koji koristimo u eksperimentalnom dizajnu je da podijelimo ljude slučajnim odabirom iz zajedničke skupine ljudi u dvije grupe. Eksperiment se bazira na ideji slučajne podjele u grupe kao temelj za dobivanje sličnih grupa. Tada na jednu provedemo postupak, a na drugu ne. Na isti način promatramo posljedicu u obje grupe. Upravo taj slučajan odabir je ključ uspjeha eksperimenta. Ustvari, čak i sa tim slučajnim odabirom nikad ne očekujemo da će grupe koje smo stvorili biti točno iste. Kako bi i bile kad su sastavljene od različitih ljudi?! Pouzdajemo se u ideju o vjerojatnosti i pretpostavljamo da su te dvije grupe statistički jednake ili da su jednake unutar poznatog vjerojatnosnog raspona. Dakle ako nasumice podijelimo ljude u dvije grupe, i imamo dovoljno ljudi u istraživanju da bi postigli željenu vjerojatnosnu jednakost, onda možemo smatrati da će eksperiment biti čvrst u odnosu na unutrašnju valjanost, a tada vjerojatno imamo dobar zaključak da li postupak uzrokuje posljedicu.

No, mnogo toga može poći po zlu. Može se desiti da je uzorak nedovoljno velik, ili pak da nam ljudi odbijaju sudjelovati u istraživanju, ili možemo biti uspješno izazvani na etičkom polju (radi korištenja ovog pristupa moramo uskratiti postupak ljudima koji ga možda jednako zaslužuju kao i drugi), ili možemo naići na otpor uprave koji bi željeli imati svoje ljude u postupaku (na

primjer, ravnatelj možda inzistira na sudjelovanje svog djeteta u novom postupaku misleći da će tako povećati njegove ocjene).

Dakle, eksperimentalni dizajn je težak za provođenje u realnom svijetu. Eksperiment je često ometan, pa treba pribjeći proširenju na stvaranje umjetnih situacija tako da bi mogli procijeniti uzročnu vezu za visokom unutrašnjom valjanošću. U tom slučaju ograničavamo stupanj na kojem uopćavamo svoje rezultate na ukupno značenje gdje nismo namjestili eksperiment. Ako smo situaciju dobro namjestili, eksperiment može biti veoma jak dizajn za korištenje.

Da bi objasnili kako se odvija eksperimentalni dizajn, što je to eksperimentalna pogreška (eng. experimental error), odstupanje (eng. bias), ponavljanje (eng. replication) i onemogućavanje (eng. blocking), poslužiti ćemo se primjerom jednog trgovca. Pokazati ćemo čemu služe eksperimentalna i kontrolna grupa, te što se misli pod slijepi (eng. blind) i dvostruko slijepi pokus (eng. double blind trial).

Promatrati ćemo trgovca voćem i povrćem koji svoju robu nabavlja u Splitu. Zanima nas da li će, ako promjeni dobavljača, povećati svoja primanja. Stoga ćemo provesti eksperiment koji bi nam trebao dati odgovor na to pitanje. Trgovac je zabilježio svoja dnevna primanja kroz deset dana, kada je kupovao u Splitu, pa je onda narednih osam dana robu nabavljao u Šibeniku i također bilježio primanja. Ta su primanja dana u **Error! Reference source not found.**

Tablica 3.0.1. Dnevna primanja kroz deset dana

	primanja									
Split	323	274	269	552	435	391	208	529		
Šibenik	286	517	492	264	367	399	198	581	362	303

Prvo što uočavamo je da troškovi u oba dućana variraju. Da je trgovac, svaki put kupujući u Splitu potrošio 323 kn, a u Šibeniku 286 kn, bilo bi očito gdje je isplativije i ne bi trebalo praviti statističku analizu. Ipak, troškovi variraju i to uzrokuje eksperimentalnu pogrešku. No to ne znači i da je pogreška napravljena. To jednostavno znači da su na primanja utjecali i drugi faktori. U tom slučaju je moguće da je vrijeme, promet i vrsta dostupnog voća i povrća utjecala na primanja. Sigurno je da i dućan radi bolje vikendom nego u sredini tjedna i to ima bitan utjecaj na primanja. Problem sa eksperimentalnom pogreškom je da ako postoji razlika u prosječnim primanjima kod nabavke na dva različita mjesta, moglo bi biti teško, ili nemoguće reći da li je to posljedica faktora koji se istražuju ili je pak to posljedica eksperimentalne pogreške. Eksperimentalna bi pogreška trebala biti smanjena na minimum održavajući faktore koji se ne istražuju konstantne koliko je moguće i dalje od eksperimentalnog dizajna. U tom slučaju bilo bi nemoguće prilagoditi propisima vremenske prilike, ili stanje u prometu, ali radno vrijeme dućana i broj radnika bi trebao biti konstantan.

Ako uzastopce promatramo, pod naočigled identičnim okolnostima, onda možemo procijeniti veličinu eksperimentalne pogreške. Uzastopna promatranja su poznata kao ponavljanja. Veličina eksperimentalne pogreške se obično procjenjuje računajući standardnu devijaciju ponavljanja. U eksperimentu poviše ima osam ponavljanja prihoda iz Splita i deset ponavljanja prihoda iz Šibenika. Da bi smanjili veličinu eksperimentalne pogreške koristiti ćemo eksperimentalni dizajn. Najjednostavniji eksperimentalni dizajn je korištenje *dvojnog uspoređivanja* (eng. paired comparisons). Ovdje se dva postupka uspoređuju i oba se odnose na sličan neobrađen materijal. Na primjer, da bi usporedili gubitak težine uslijed dvije slične dijete idealan dizajn bi bio da

osiguramo suradnju više parova identičnih blizanaca. Jedan blizanac svakog para bi pratio jednu dijeteu a drugi drugu dijeteu. Na osnovu toga eksperimentalna greška određena fizičkim razlikama kod ljudi na dijeteu bila bi minimalna.

Vratimo se sada na našeg trgovca. Očit izvor eksperimentalne pogreške su dani u tjednu. Prihodi će vjerojatno biti veći vikendom. Zbog toga ako ispitamo razlike prihoda svakog dućana na isti dan u tjednu, jedan veliki izvor eksperimentalne pogreške biti će eliminiran. Eksperiment bi mogao biti izveden na način prikazan u **Error! Reference source not found.**

Tablica 3.0.2. Dnevna primanja kroz šest dana

	PON	UTO	SRI	ČET	PET	SUB
TJEDAN 1	DUĆAN U SPLITU					
PRIHODI	272	295	318	307	532	599
TJEDAN 2	DUĆAN U ŠIBENIKU					
PRIHODI	268	272	324	352	511	604

Podatke sada možemo analizirati koristeći t-test. Ovo bi bilo savršeno zadovoljavajuće za analiziranje podataka o gubitku težine kod blizanaca koji su pratili različite dijete. No u ovom slučaju, dizajn možemo još usavršavati. Pretpostavimo da je tjedan_1 bio vedar i suh a tjedan_2 kišan i vjetrovit. Ili pretpostavimo da su novi krumpiri bili dostupni u dućanu drugi tjedan a ne prvi tjedan. Ako se pojavila razlika između prihoda nećemo zanati da li je to određeno razlikom između dućana ili određeno različitim vremenskim uvjetima(ili kvalitetom krumpira). Ta dva učinka nećemo moći razlikovati. To nas dovodi do zaključka da bi grupiranje bilo bolje kad bi se napravilo na način prikazan u **Error! Reference source not found.**

Tablica 3.0.3. Kvalitetniji prikaz primanja kroz šest dana

TJEDAN 1	PON	UTO	SRI	ČET	PET	SUB
DUĆAN	A	B	A	B	A	B
PRIHODI	248	296	333	376	494	517
TJEDAN 2	PON	UTO	SRI	ČET	PET	SUB
DUĆAN	B	A	B	A	B	A
PRIHODI	276	308	307	400	482	512

A i B predstavljaju dva dućana. Posljednja odluka bi bila da li A predstavlja Split, a B Šibenik ili obratno. Tamo gdje ne postoji očit razlog za odabir jednog, a ne drugog izbor bi trebao biti napravljen odabirom kao što je bacanje novčića. Ta je metoda poznata kao *slučajan odabir* (eng. randomisation) i služi za sprečavanje nesvjesnih i neočekivanih odstupanja koja utječu na rezultat. Nema više očitih daljnjih poboljšanja koja bi mogli napraviti ovom dizajnu. Pretpostavimo da je A izabran, i neka bude Split. Tada bi analiza izgledala ovako:

Tablica 3.0.4. Razlika primanja kroz šest dana

	PON	UTO	SRI	ČET	PET	SUB
SPLIT	248	308	333	400	494	512
ŠIBENIK	276	296	307	376	482	517
RAZLIKA	12	12	26	24	12	-5

Označimo sa H0 hipotezu da je *razlika aritmetičkih sredina* (eng. mean difference) beznačajna, a sa H1 hipotezu koja govori da je razlika aritmetičkih sredina značajna. Ove podatke možemo testirati koristeći t-test. Prosjek uzorka (eng. sample mean) će biti približno normalno

raspodijeljen čak i ako pojedinačne razlike ne prate točno normalnu raspodjelu. Prema tome je sigurno primijeniti ovaj test osim ako je jako očito da je raspodjela krajnje asimetrična. Beznačajna razlika je $d=13.5$, a standardna devijacija razlika je $\sigma=11.095$, pa iz toga dobivamo da je

$$t = \frac{13.5}{\frac{11.095}{\sqrt{6}}} = 2.98 \quad (1)$$

Kritične vrijednosti od t za 5% dvostranog rizika od *pogreške prvog stupnja* (eng. Type 1 error) su ± 2.571 . Hipotezu H_0 smo ovim postupkom odbacili, i s obzirom da je beznačajna razlika aritmetičkih sredina Split-Šibenik pozitivna, zaključujemo da su prihodi viši kada trgovac kupuje u Splitu. Nemoguće je potpuno isključiti mogućnost da je razlika posljedica nekog drugog faktora osim dućana. Mogućnost uvijek postoji. Ali dobro dizajniran eksperiment eliminira sva moguća druga objašnjenja.

Ako prodavač želi usporediti djelovanje i drugih dućana na prihode mogao bi posjetiti svaki dućan na slučajno odabrane dane i bilježiti primanja.

Tablica 3.0.5. Djelovanje drugih dućana na primanja

	PRIHODI					
MAKARSKA	344	479	503	290	207	
SPLIT	459	234	602	222	598	479
ZADAR	322	600	308	344	506	
ŠIBENIK	292	588	347	399	544	406

Ovo je znano kao *potpuno slučajan dizajn* (eng. completely randomized design) i analizira se koristeći jedno faktorsku analizu varijance (eng. one factor analysis of variance) Ovdje se mogu primijeniti ideje dvojnog uspoređivanja (eng. paired comparison). Eksperiment se može izvesti na sljedeći način:

Tablica 3.0.6. Prihodi kroz šest dana

	prihodi					
	PON	UTO	SRI	ČET	PET	SUB
MAKARSKA	294	306	343	386	494	527
SPLIT	277	318	399	360	524	566
ZADAR	299	265	302	410	488	530
ŠIBENIK	260	289	299	391	460	488

Ovo je poznato kao *slučajni blok dizajn* (eng. randomised block design). Dućani su podijeljeni na blokove po danima u tjednu. Ovaj dizajn se analizira koristeći dvofaktorsku analizu varijance. Prvi faktor, u ovom slučaju dućan, je objekt istraživanja. Drugi faktor, u ovom slučaju dan u tjednu, je uveden zato što se misli da bi to moglo imati bitan učinak na rezultat. Kao kod dvojnog uspoređivanja, uvođenje dana u tjednu kao faktora će, ako je dizajn uspješan, smanjiti eksperimentalnu pogrešku. Biti će lakše odrediti razliku između dućana, ako ta razlika uopće postoji.

U primjerima poviše, uspoređivali smo dva ili više dućana. Ponekad postoji samo jedna obrada, a ne dvije i više. Na primjer, želimo promatrati učinak detaljnog medicinskog tretmana na upalu zglobova ili učinak trenerskog tečaja na učenike tenisa. Njihov učinak se ne može mjeriti u izolaciji. Upala zgloba se može poboljšati ili pogoršati bez ikakvog tretmana. I slično tenisač se može usavršiti i bez da ide na trenerski tečaj.

Nužno je imati kontrolnu i eksperimentalnu grupu. Ove dvije grupe bi trebale biti što je moguće sličnije. To jest ljudi u jednoj grupi bi trebali biti slični ljudima u drugoj grupi za obilježja koja su važna za istraživanje. To ne znači da svaki čovjek mora biti sličan onom drugom, nego da grupa kao cjelina mora biti slična drugoj grupi. Npr. u slučaju bolesnika koji pate od upale zgloba, dvije grupe trebaju sadržavati ljude sličnog godišta, spola, općenitog zdravlja i stupnja bolesti.

U slučaju tenisača grupe moraju sadržavati učenike slične dobi, spola, kondicije i teniske vještine. Kada odaberemo grupe, slučajnim odabirom se odluči koja će se grupa tretirati kao eksperimentalna, a druga ostaje kao kontrolna grupa i na nju nećemo provoditi nikakav postupak, odnosno tretman, ili ćemo je tretirati kao i do sada. Tada možemo usporediti učinak na tim grupama. U slučaju medicinskog liječenja nekad se mislilo da će se stanje pacijenta poboljšati, ili da će ozdraviti bez liječenja i da će u nekim slučajevima poboljšanje biti veće i brže, ako im se kaže da ih se liječi a u biti nije. Pa je standardna praksa u testovima za lijekove da se kontrolnoj grupi da placebo. To je bezopasna supstancija koja izgleda kao lijek ali ne sadrži nikakve kemikalije. Mnogim će se pacijentima poboljšati stanje nakon uzimanja placeba. Da bi pokazali da je lijek učinkovit, značajan broj pacijenata koji uzimaju lijek moraju pokazati poboljšanje od onih koji uzimaju placebo.

Ako pacijenti koji uzimaju placebo znaju da uzimaju to, učinak će naravno biti izgubljen. Bitno je da pacijenti ne znaju da li uzimaju, ili ne uzimaju placebo i to se zove slijepi pokus. Može se desiti i da, čak i ako pacijenti ne znaju da li uzimaju placebo, doktori mogu očekivati od pacijenata koji uzimaju lijek da će se oni oporavljati brže nego oni koji uzimaju placebo. Ovo se očekivanje može nekako prenijeti i na pacijente i za rezultat dobijemo njegovo brže poboljšanje zdravlja. Zato je nužno da doktor ne zna koji pacijent uzima placebo a koji lijek. Naravno netko to mora znati, inače bi bilo nemoguće analizirati podatke. No to bi trebao biti netko tko nema direktnu vezu s pacijentima. Pokus gdje ni doktor ni pacijenti ne znaju tko prima lijek zove se dvostruko slijepi pokus. Predlaže se da bi i osoba koja radi statističku analizu trebala ne znati koji pacijenti koriste lijekove da bi se spriječio utjecaj na analizu. To bi opisali kao trostruki slijepi pokus. Testiranje lijekova predstavlja velik etički problem. Prvo bitno je da pacijenti koji sudjeluju budu potpuno informirani o mogućim rizicima eksperimenta. Drugo jednom kad se liječenje pokaže djelotvorno, bilo bi krivo ne dopustiti svim pacijentima da sudjeluju u liječenju. Treće, bilo bi loše nastaviti provoditi eksperiment ako nije dobro dizajniran zbog svih neugodnosti i pobuđivanja lažnih nadi.

3.1. Postavljanje i etape eksperimenta

Proces od zamisli do realizacije nekog eksperimenta prolazi kroz različite međusobno povezane faze. Ukoliko želimo proizvesti smislene i korisne rezultate, formalni eksperimenti traže mnogo

brige i planiranja. U ovom poglavlju ćemo razmotriti planiranje koje je potrebno da bi definirali i pokrenuli službeni eksperiment. Također ćemo razmatrati nekoliko ključnih karakteristika eksperimenta. Službeni eksperiment se provodi u nekoliko koraka:

- koncepcija
- dizajn
- priprema
- izvršenje
- analiza
- oglašavanje i donošenje zaključaka

Na početku i u osnovi svakog eksperimenta nalazi se neki problem, to jest neko pitanje na koje ne možemo pouzdano odgovoriti. Takva pitanja su na primjer: da li takmičenje među skupinama koje se znatno razlikuju po svojim sposobnostima daje bolje rezultate, nego takmičenje među izjednačenim skupinama; da li broj informacija koje nastavnik prikuplja o znanju učenika u toku ispita utječe na valjanost ocjena.... Dakle, upravo u prvom koraku, to jest koncepciji, donose se odluke o tome što istraživati, te se definiraju ciljevi eksperimenta. Cilj istraživanja se mora izreći jasno i precizno. Taj cilj može uključivati i da je određena metoda bolja od neke druge metode. Stoga u fazu koncepcije uključujemo analizu koja će nam osigurati da je službeni eksperiment najprikladnija tehnika za određeno istraživanje.

U drugom koraku, to jest u dizajnu želimo pokazati da za određenu metodu ili alat, razlike u uvjetima okoliša ili razlike u kvaliteti sredstava mogu utjecati na upotrebu ili krajnji rezultat metode ili oruđa. Bez obzira koju metodu odaberemo, naš cilj mora biti naznačen tako da može biti jasno ocijenjen na kraju eksperimenta, to jest mora biti izražen u obliku pitanja na koje želimo dati odgovor. Već i sa samim pojavljivanjem problema javlja se jedna ili više hipoteza.

Hipotezu definiramo kao “proizvodni, prema mišljenju onoga koji je čini, jedan od mogućih odgovora na problem, ali koji odgovor tek daljnje istraživanje treba da potvrdi ili odbaci“, [BUJA1967]. Već i samo pitanje uključuje neku hipotezu, ali su hipoteze koje definira istraživač u vezi sa nekim problemom, mnogo određenije jer pretpostavljaju, ne samo da bi neki faktor mogao utjecati na neki fenomen, nego uključuju i određeno predviđanje na koji bi način nezavisna varijabla mogla utjecati na zavisnu varijablu. Kada istraživač postavlja svoju hipotezu, tako da predviđa, ne samo mogućnost postojanja neke veze među varijablama, nego i vrstu učinka te veze, tada govorimo o afirmativnim hipotezama. Naprotiv, kada je hipoteza postavljena u indiferentnom obliku, tada govorimo o takozvanoj nultoj-hipotezi. Nulta-hipoteza je ona koja pretpostavlja da nema razlike između dva tretmana (to jest između dvije metode, oruđa, tehnike ili dva uvjeta koji utječu na ono što mjerimo) u odnosu na ovisnu varijablu koju mjerimo (kao na primjer produktivnost, kvaliteta, trošak). Alternativna hipoteza pretpostavlja da postoji važna razlika između dva tretmana. Zbog korektnog planiranja eksperimenta preporučuje se da istraživač polazi od hipoteza koje su postavljene u indiferentnom obliku, te da ispitanici, odnosno subjekti, ne znaju za hipoteze jer bi mogli nehotice zauzeti određene stavove koji se odražavaju u njihovim reakcijama. Na osnovu hipoteze formulira se cilj eksperimenta i izrađuje dizajn.

Navesti ćemo jedan primjer. Charl Lawrence Pfleeger je istraživao utjecaj *tehnike čiste sobe* (eng. clean room) na razvoj programske podrške. Želio je saznati da li ta tehnika čiste sobe, u razvoju programske podrške, proizvodi kod više kvalitete od tekućeg procesa razvoja. Hipoteze je formulirao na sljedeći način: nulta-hipoteza bi bila: «nema razlike u kvaliteti koda između koda koji je proizveden upotrebljavajući čistu sobu i koda koji je proizveden upotrebljavajući tekući

proces», [PFLE1995] dok bi alternativna hipoteza glasila: «kvaliteta koda koji je proizveden upotrebom čiste sobe je viša nego kvaliteta koda koji je proizveden upotrebom tekućeg procesa», [PFLE1995]. Razlikovna karakteristika uključuje statističke pretpostavke, i za nultu hipotezu se pretpostavlja da je istinita, osim ako podaci dobiveni istraživanjem pokazuju suprotno. Upravo zbog toga eksperiment se usredotočuje na polaznje od nulte hipoteze radije nego od alternativne hipoteze. U tom smislu testiranje hipoteze znači određivanje da li su podaci dovoljno uvjerljivi da odbiju nultu hipotezu i prihvate alternativnu hipotezu kao točnu. Nakon definiranja hipoteze slijedi stvaranje formalnog dizajna, kako bi se testirale hipoteze. Eksperimentalni dizajn je kompletan plan za primjenu razlikovnih eksperimentalnih uvjeta na eksperimentalnim subjektima, kako bismo odredili na koji način i u kojoj mjeri uvjeti utječu na ponašanje rezultata neke aktivnosti. Posebno bi trebalo isplanirati kako će primjena danih uvjeta pomoći da testiramo hipoteze i dati odgovor na ciljano pitanje.

Nužnost službenog plana ili dizajna opisati ćemo upravo u primjeru koji S.L. Pfleeger navodi u svom radu o eksperimentalnom dizajnu i analizi u inženjeringu programske podrške. Pfleeger je želio odrediti učinak upotrebe Ada jezika na kvalitetu rada. Kako je problem preopćenit da bi bio koristan trebalo bi se zapitati nekoliko specijalnih pitanja poput: kako se mjeri kvaliteta, kako se mjeri upotreba Ada jezika i koji faktori utječu na karakteristike koje se mjere. Na primjer, hoće li iskustvo, alat, tehnika dizajna ili tehnika testiranja nešto promijeniti, koji će od faktora biti proučavani u istrazi i koliko bi se eksperiment trebao izvoditi i pod kojim uvjetima i u kojem okolišu će se istraživati upotreba koda, zatim, kako bi se rezultati trebali analizirati i kolika će se veličina razlike u kvaliteti smatrati važna. On napominje da su ovo samo neka od pitanja koja moraju biti odgovorena prije nego eksperiment započne. Postoji službena tehnika za opisivanje komponenti eksperimenta. Ova terminologija traži od nas da razmotrimo sve aspekte eksperimenta, tako da je on potpuno isplaniran u svrhu osiguravanja korisnih rezultata. Nova metoda ili alat koje želimo procijeniti (u usporedbi sa postojećom metodom ili alatom) se zove tretman. Cilj nam je odrediti da li je tretman koristan u određenim okolnostima, to jest želimo utvrditi da li tretman proizvodi rezultate koji su na neki način različiti. Dakle, želimo saznati povećava li novi alat produktivnost u usporedbi sa postojećim alatom i njegovom produktivnošću, ili želimo odabrati između dvaju tehnika ovisno o njihovom učinku na kvalitetu učinka.

Eksperiment se sastoji od serije testiranja odabranih metoda i alata i eksperimentalnog dizajna koji opisuje kako će testiranje biti organizirano i kojim redoslijedom će se pokretati. U eksperimentalnom dizajnu je potrebno odrediti i razraditi nezavisnu i zavisnu varijablu, subjekte koji su uključeni u istraživanje, kako će se eksperiment provesti i kojim postupcima će se vršiti mjerenja, na koji način će se vršiti kontrola i koje statističke metode koristiti U svakom individualnom pokretanju testa koristi se samo jedan tretman. Individualni test ove vrste ponekad nazivamo pokusni test, a eksperiment formalno definiramo kao set pokušaja. Eksperiment može uključivati i više od jednog tretmana. Tada uspoređujemo samo one rezultate tretmana koji se razlikuju.

Ciljevi eksperimenta ili eksperimentalne jedinice su ciljevi zbog kojih se tretman primjenjuje. U svrhu toga razvoj ili briga o projektu može biti cilj eksperimenta i pogledi procesa projekta ili organizacije mogu biti promijenjeni kako bi utjecali na rezultat. Bitno je da u isto vrijeme identificiramo tko će sve koristiti tretman; te ljude nazivamo eksperimentalni subjekti. Karakteristike eksperimentalnih subjekata moraju biti jasno definirane tako da učinci razlika između subjekata mogu biti ocijenjeni u promatranim rezultatima. Kada uspoređujemo rezultate

upotrebe tretmana sa rezultatima gdje tretman nije upotrijebljen, potrebno je uspostaviti kontrolni cilj koji se dobije ne upotrebom tretmana. Kontrola nam daje baznu liniju informacije koja omogućava da radimo usporedbu. Ona je okoliš u kojem se pokreće istraživanje i već postoji u vrijeme kada istraživanje počinje. U formalnom eksperimentu kontrolna situacija mora biti definirana eksplicitno i oprezno tako da se razlike između kontrolnog cilja i eksperimentalnog cilja razumiju.

Povratne ili ovisne varijable su oni faktori od kojih se očekuje da promijene rezultat primjene tretmana. Suprotno od njih su statične ili neovisne varijable. To su one koje mogu utjecati na primjenu tretmana i stoga na rezultat eksperimenta. Statične varijable obično opisuju karakter ljudi, proizvoda ili procesa korištenih da bi proizveli u eksperimentu. Važno je definirati i okarakterizirati statične varijable tako da njihov utjecaj na povratne varijable može biti istražen. Štoviše statične varijable mogu biti korisne u definiranju cilja eksperimenta i u upotrebi projekta u kojem će sudjelovati. Statične varijable i kontrola koju imamo nad njima pomažu razlikovati istraživanje slučajnosti od običnih eksperimenata. Broj i veze među subjektima, ciljevima i varijablama može biti detaljno opisan u planu eksperimenta. Što više subjekata, ciljeva i varijabla, to više eksperimentalni dizajn postaje kompliciraniji i često teži za analizu. Upravo zbog toga je veoma važno uložiti veliku količinu vremena i brige u dizajniranju eksperimenta, a ne brzati s primjenom pokušaja i prikupljanja podataka. Jednom kada je dizajn gotov, moguće je uočiti koji su eksperimentalni faktori uključeni (to jest, krajnje i statične varijable), koliko subjekata će biti potrebno, od koje populacije ćemo ih uzeti i kojim uvjetima će svaki subjekt biti izložen. Ako je uključeno više od jednog tretmana, treba odrediti red izlaganja kao i kriterij mjerenja učinka, treba definirati na koji način donijeti zaključak, kao i metode za nastavljavanje mjerenja.

Treći korak, to jest priprema uključuje pripremanje subjekata za primjenu tretmana. Pribor koji je potreban pri administraciji nezavisne varijable i pri bilježenju ponašanja zavisi o vrsti eksperimenta. Očito je da tehnička pomagala moraju besprijekorno funkcionirati i biti pod stalnom kontrolom. Osim toga, bitno je da uređaji koje upotrebljavamo u eksperimentu budu tako konstruirani i raspoređeni da što manje smetaju subjektima ispitivanja. Potrebno je obučiti osoblje koje će sudjelovati u eksperimentu, te napisati upute razumljivo i propisno. Veoma korisno bi bilo izvesti, naravno ako je to moguće, probni pokus na maloj skupini ljudi, da bi bili sigurni da je plan potpun i upute razumljive. Unatoč svemu tome, pouzdanost i upotrebljivost dobivenih rezultata mnogo više ovisi o načinu na koji istraživač kontrolira izvođenje eksperimenta, nego o složenosti uređaja kojima se pri tom služi.

Glavno pravilo kojeg se istraživač treba pridržavati je to da uvijek na isti način provodi eksperiment. To znači da bi uputa subjektima trebala uvijek biti ista, kao i ponašanje istraživača s ispitanicima prije, za vrijeme i nakon izvođenja eksperimenta. Ako se istraživač toga ne pridržava, onda konstantnost prilika pod kojima se eksperiment provodi neće biti ni približno ostvarena. Ako istraživač, zbog određenih iskustava koja je stekao u provođenju eksperimenta, odluči nešto promijeniti u daljnjem izvođenju eksperimenta, onda on u stvari počinje drugi eksperiment i ne može uspoređivati rezultate koje je dobio u staroj i novoj situaciji.

Istraživač koji izvodi eksperiment je dužan organizirati svoj eksperiment na način da pomoću njega može, ne samo utvrditi neki učinak na zavisnu varijablu, nego i dokazati da su te posljedice rezultat djelovanja baš određene varijable, a ne nekog drugog relevantnog faktora ili slučajnog stjecaja okolnosti. Postupci kojima utvrđuje različite relevantne faktore u eksperimentu, pomoću

kojih ih uklanja ili stabilizira, i pomoću kojih dokazuje njihovu stabilizaciju zovu se kontrola eksperimenta. Kontrola eksperimenta počinje s pažljivom analizom i određivanjem koji bi sve faktori mogli utjecati na rezultate eksperimenta. Nije moguće poznavati sve faktore koji imaju utjecaj na predmet ispitivanja, ali je u mogućnosti da razvrsta bar one koji su u tolikoj mjeri bitni da bi mogli prekriti pravo djelovanje nezavisne varijable. Velik broj parazitalnih faktora se može neutralizirati eliminiranjem. Tako na primjer, ako su svjetlo i buka takvi faktori, eksperiment bi trebalo provesti u mračnoj i tihoj prostoriji. One faktore koje nije moguće ukloniti treba održavati konstantnima, tako da oni u istoj mjeri sudjeluju u svim situacijama eksperimenta. Stabilizacija različitih vanjskih prilika obično ne zadaje poteškoće. Na primjer, utjecaj temperature, doba dana i vlage stabiliziramo tako da eksperiment vršimo u konstantnoj klimi, u isto doba dana.

Na posljetku, kada napravimo sve pripreme, eksperiment možemo početi izvoditi. Slijedeći korake koji su izloženi u planu i mjereći attribute koji su predodređeni u planu, primjenjujemo tretman na eksperimentalne subjekte. Potreban je veliki oprez da su predmeti koji se mjere i tretmani koji se primjenjuju stalni tako da usporedba rezultata bude osjetljiva.

Fazu analize dijelimo u dva koraka. Prvo moramo provjeriti sva mjerenja kako bismo osigurali da se mogu upotrijebiti i biti korisna, zatim organiziramo mjerenja u skupu podataka koji će biti pregledani kao dio procesa testiranja hipoteze. U drugom koraku analiziramo skup podataka prema određenim statističkim principima. Statistički testovi, kada su propisno primijenjeni, kazuju na to da li je nulta hipoteza podržana ili odbačena od rezultata eksperimenta. Na samom kraju ove faze dolazimo do zaključka o tome koliko različite karakteristike koje pregledavamo utječu na krajnji rezultat. Zaključke je važno zabilježiti na način koji će dopustiti ostalim kolegama da ponovo upotrijebe isti eksperiment i da potvrde naše zaključke u sličnom okruženju. U toj svrsi trebalo bi bilježiti sve ključne aspekte istraživanja; ciljeve, hipoteze, eksperimentalne subjekte, povratne i statične ciljeve, varijable, tretmane i rezultate istraživanja. Također bi i svaka druga važna dokumentacija trebala biti zabilježena; upute i karakteristike korištenih alata ili metoda. Zaključci moraju biti izraženi što jasnije, trudeći se navesti sve probleme koji su se pojavili tijekom provođenja eksperimenta.

3.2. Principi eksperimentalnih dizajna

Korisni rezultati ovise o pažljivom, rigoroznom i cjelovitom eksperimentalnom dizajnu. Nadalje se istražuju principi koji se trebaju razmotriti u dizajniranju eksperimenta. Svaki princip se obraća potrebi za pojednostavljuvanjem i povećanjem informacija. Samostalni dizajni pomažu učiniti eksperiment praktičnim umanjujući upotrebu vremena, novca, osoblja i sredstava za eksperiment. Dodatna korist je da su jednostavni dizajni lakši za analizu (i stoga su ekonomičniji) nego komplicirani dizajni. Povećanje informacija daje vam kompletno razumijevanje uvjeta eksperimenta i rezultata omogućujući generalizaciju rezultata u svim mogućim situacijama. Eksperimentalni dizajn uključuje dvije važne koncepcije, eksperimentalnu jedinicu i eksperimentalnu pogrešku.

Eksperimentalna jedinica je eksperimentalni cilj na kojem se primjenjuje jedan tretman. [PFLE1995] Uobičajeno je da se tretman primjenjuje više od jedanput. U najmanjem slučaju

primjenjuje se na kontrolnu grupu, kao i na bar jednu drugu grupu koja se razlikuje od kontrolne grupe statističkom varijablom. U mnogim slučajevima tretman se primjenjuje mnogo puta na više grupa. U svim tim slučajevima rezultati se pregledavaju kako bi se uočile razlike u primjeni tretmana. Međutim, čak i kad se uvjeti održe istima može doći do neznatno drugačijih rezultata.

Eksperimentalna pogreška opisuje neuspjeh dvije jednako tretirane eksperimentalne jedinice kako bi dale identične rezultate. [PFLE1995] Na pogrešku može utjecati mnogo problema kao što su pogreška provođenja eksperimenta, pogreška promatranja, pogreška mjerenja, varijacija u eksperimentalnom sredstvu, kombinirani efekti svih vanjskih faktora koji mogu utjecati na karakteristike koje se proučavaju ali nisu izdvojene kako bi se istraga njima bavila. Cilj dobro eksperimentalnog dizajna je kontrolirati što je moguće više varijabla, kako bi se maksimalno smanjila varijabilnost među sudionicima i smanjili efekti nevažnih varijabli. Idealno bi bilo eliminirati efekte drugih varijabli, tako da se samo efekti nezavisnih varijabli odražavaju u vrijednosti zavisne varijable, to jest željeli bi eliminirati eksperimentalnu pogrešku. Realno, potpuna eliminacija eksperimentalne pogreške nije uvijek moguća, pa se umjesto tog pokušava dizajnirati eksperiment tako da se efekti nevažnih varijabli jednako raspodijele preko svih eksperimentalnih uvjeta, radije nego dopustiti da povećaju odstupanje rezultata određenih uvjeta. Više ćemo o eksperimentalnoj pogreški reći u poglavlju **Error! Reference source not found.**
Error! Reference source not found.

Tri su osnovna principa eksperimentalnog dizajna:

1. *replikacija* (eng. replication)
2. *randomizacija* (eng. randomization)
3. *lokalna kontrola* (eng. local control)

1. REPLIKACIJA

Replikacija je ponavljanje osnovnog eksperimenta. Ona u biti prije uključuje ponavljanje eksperimenta pod identičnim okolnostima, nego ponavljanje mjerenja iste eksperimentalne jedinice. Poželjna je iz više razloga. Prvo, u kombinaciji sa odgovarajućim statističkim tehnikama pruža nam procjenu eksperimentalne pogreške koja djeluje kao osnova za određivanje važnosti promatranih razlika kod nezavisnih varijabli. Ona nam omogućuje da saznamo koliko povjerenja možemo imati u rezultate eksperimenta. Drugo, omogućava nam uspostavljanje beznačajnog učinka bilo kojeg eksperimentalnog faktora.

Bitno je osigurati da nas replikacija ne dovede do *zbunjujućih* (eng. confound) efekata. Kažemo da su dvije ili više varijabli zbujujuće ako je nemoguće odvojiti njihove učinke kada se provede sljedeća analiza. Objasniti ćemo to na sljedećim primjerima [PFLE1995].

Pretpostavimo da želimo usporediti učinak novog alata s već postojećim alatom. Postavili bi eksperiment gdje programer A koristi novi alat u razvojnom okruženju, dok programer B koristi već postojeći alat. Kada usporedimo mjerenja kvalitete u novonastalom kodu, nismo u mogućnosti kazati kolika je razlika između ta dva alata iz razloga što nismo uračunali razliku u programerskim vještinama ta dva ispitanika. Zaključujemo da su učinak novog alata (jedna varijabla) i programerska vještina (druga varijabla) zbujujuće. Ovo zbujuvanje možemo riješiti

s replikacijom kada ponavljanje eksperimenta ne kontrolira ostale varijable, kao na primjer tu programersku vještinu.

Slično, možemo promotriti uspoređivanje dvaju tehnika za testiranje. Prvo grupu obučimo za rad na tehnici za testiranje X i tražimo od njih da testiraju set modula. Otkriven je određen broj nepravilnosti kod odabranog mjerenja učinka te tehnike. Zatim, testni tim obučimo za rad na tehnici za testiranje Y, nakon koje testiraju iste module. Usporedba broja nepravilnosti kod testiranja X i testiranja Y može biti pobrkana sa sličnostima između dviju tehnika ili procesom učenja. U ovom slučaju je slijed ponavljanja izvor zbunjivanja.

Iz ovog razloga eksperimentalni dizajn treba u detalje opisati sve vrste replikacija u eksperimentu. Treba identificirati uvijete pod kojima je svaki od ponavljanih eksperimenata pokrenut, uključujući i njihov poredak, i mjerenja koja će se provesti za svaku replikaciju.

2. RANDOMIZACIJA

Replikacija omogućava statistički test značajnosti rezultata, ali nam ne osigurava njihovu valjanost. Treba biti siguran da eksperimentalni rezultati slijede iz tretmana koji je primijenjen, a ne iz drugih varijabli. Neki aspekti eksperimentalnog dizajna trebaju organizirati eksperimentalne pokuse na način da distribuiraju očitovanja neovisno, kako bi osigurali valjanost eksperimenta. Randomizacija označava slučajnu raspodjelu subjekata grupama ili različitim tretmanima kako bi osigurali neovisnost (pa tako i valjanost) rezultata. Randomizacija ne osigurava neovisnost, ali dopušta pretpostavku da je odnos bilo koje usporedbe tretmana najmanji moguć. Drugim riječima, slučajnom raspodjelom tretmana eksperimentalnim jedinicama pokušavamo onemogućiti odstupanje rezultata tretmana koje nastaje pod utjecajem efekata nad kojima nemamo kontrolu.

Na primjer, ponekada na rezultate može utjecati vrijeme, mjesto ili nepoznate karakteristike sudionika. Ti nekontrolirani faktori mogu uvelike utjecati na rezultate kontroliranih varijabli. Da bi uklonili efekte tih nekontroliranih ili nepoznatih faktora možemo slučajnom raspodjelom odrediti redoslijed tretmana, ili možemo slučajnom raspodjelom dodijeliti subjekte svakom tretmanu ili slučajnim odabirom možemo odrediti lokaciju za svaki tretman posebno, kada god je to moguće.

Ključni aspekt randomizacije je dodjela subjekata grupama i tretmanima. Ako koristimo iste subjekte u svim eksperimentalnim uvjetima, govorimo o dizajnu unutar subjekata. Ako koristimo različite subjekte u različitim eksperimentalnim uvjetima dobijemo nepovezani dizajn između subjekata. Eksperimentalni dizajn bi trebao uključivati detalje o tome kako je planirana slučajna raspodjela subjekata grupama, ili tretmana eksperimentalnim jedinicama. U slučajevima gdje potpuna randomizacija nije moguća treba dokumentirati područja u kojima manjak randomizacije može utjecati na valjanost rezultata.

3. LOKALNA KONTROLA

Jedan od ključnih faktora pomoću kojeg razlikujemo formalni eksperiment od proučavanja slučaja je stupanj kontrole. Lokalna kontrola je aspekt eksperimentalnog dizajna koji pokazuje koliku kontrolu imate nad smještanjem subjekata u eksperimentalnim jedinicama, te koliku

kontrolu imate nad njihovom organizacijom unutar tih jedinica. S obzirom da replikacija i randomizacija osiguravaju valjani test značajnosti, lokalna kontrola čini dizajn učinkovitijim tako što reducira veličinu eksperimentalne pogreške. Obično se lokalna kontrola ispituje u odnosima dviju karakteristika dizajna: blokiranju i balansiranju jedinica

Pod blokiranjem mislimo na premještanje eksperimentalnih jedinica u blokove ili grupe tako da su jedinice unutar bloka relativno homogene (ujednačene). Blokovi su dizajnirani tako da se predvidljiva varijacija jedinica pomiješa sa efektom blokova. Eksperimentalni dizajn zadržava predviđenu varijaciju u blokovima tako da grupira po sortama, kako bi se izbjeglo njeno pridonšenje eksperimentalnoj pogreški. Na primjer pretpostavimo da istražujemo uspoređujuće učinke triju dizajnerskih tehnika na kvaliteti rezultirajućeg koda. [PFLE1995] Eksperiment uključuje učenje tehnika za dvanaest programera i mjerenje broja nepravilnosti koje su nađene po tisuću linja koda u svrhu dobivanja kvalitete koda. Može se desiti da su programeri diplomirali na tri različita sveučilišta. Zbog toga je moguće da oni posjeduju drugačija znanja koja utječu na to kako će svaki od njih razumjeti i koristiti dizajnersku tehniku. Da bi se ova mogućnost eliminirala definiraju se tri bloka, od kojih prvi sadrži programere sa sveučilišta X, drugi sa sveučilišta Y, a treći sa sveučilišta Z. Zatim tretmane raspodijelimo slučajno programerima u svakom bloku. Ako u prvom bloku ima šest programera, može se očekivati da će dvojici dopasti dizajnerska metoda A, dvojici dizajnerska metoda B i dvojici C.

Balansiranje je blokiranje i dodjeljivanje tretmana na način da je podjednak broj subjekata dodijeljen svakom tretmanu, gdje god je to moguće. Balansiranje je poželjno zbog toga što pojednostavljuje statističku analizu, ali nije nužno. Dizajni mogu biti napravljeni tako da su u potpunosti balansirani pa sve do toga da imaju vrlo malo ili čak nimalo balansiranja.

U eksperimentalnim istraživanjima samo jedan faktor, blokiranje i balansiranje igra važnu ulogu. Ako dizajn ne uključuje blokove onda mora biti u potpunosti randomiziran. To jest, subjekti trebaju biti dodijeljeni slučajnim odabirom svakom tretmanu. Balansirani dizajn, sa jednakim brojem subjekata po tretmanu je poželjan ali ne i nužan. Ako je korišten jedan blokirajući faktor, subjekti su podijeljeni u blokove i tada slučajnim odabirom dodijeljeni svakom tretmanu. U takvom dizajnu, kojeg nazivamo slučajni blok dizajn, balansiranje je ključno za analizu. Zbog toga ovaj tip dizajna ponekad nazivamo potpuno balansirani blok dizajn. Ponekad su jedinice podijeljene u blokove u odnosu na dvije različite varijable i tada slučajno dodijeljene tretmanima tako da je svaka kombinacija blokirajućih varijabli dodijeljena svakom tretmanu jednak broj puta. U ovom slučaju, koji se naziva Latinski kvadrat, balansiranje je obavezno za točnu analizu.

Eksperimentalni dizajn bi trebao uključivati opis definiranih blokova i raspoređivanja tretmana svakom bloku.

3.3. Eksperimentalna pogreška

Greška (ili nesigurnost) je definirana kao razlika između izmjerene ili procijenjene vrijednosti za količinu i njene prave vrijednosti.[CORN2007] Ona je stvarna vrijednost i svojstvena je svim mjerenjima. Znanje o vrsti i stupnju pogreške koja bi mogla biti prisutna je bitno ako se podaci koriste mudro, bilo da su dobiveni osobnim mjerenjem, ili pročitani iz priručnika o proizvodu. U

medicini, biologiji i društvenim znanostima plan za skupljanje podataka i analizu je glavni dio eksperimenta. Inženjeri također moraju biti oprezni; iako su neka njihova mjerenja napravljena savršenom točnosti (npr. brzina svjetlosti je 299 792 458 m/sec), većina pogrešaka manjih od 1% smatra se prihvatljivima, a za nekolicinu se mora koristiti unaprjeđeni eksperimentalni dizajn tehničke analize da bi se dobili korisni podaci. Njihovo mjerenje i analiziranje je ključni dio inženjerskog procesa, od početne karakterizacije materijala i dijelova potrebnih za dizajn, testiranja prototipa, kontrole kvalitete prilikom proizvodnje pa do krajnjeg proizvoda.

Eksperimentalni rezultati bi uvijek trebali uključivati realnu procjenu pogreške, ili jasno određenu kao pozitivnu/negativnu vrijednost, ili koristeći prikladan broj značajnih figura. Nadalje, morali bi uključiti razmišljanja i izračune koji su ušli u procjenu pogreške, ako je to prihvatljivo i drugima. Direktna procjena pogreške može biti dana ili kao pozitivno/negativna veličina apsolutne pogreške, u jedinicama mjerenja, ili kao djelomična ili relativna pogreška (eng. *fractional or relative error*), izražena kao pozitivno/negativni razlomak postotka mjerenja. Prednost oblika djelomične pogreške je ta što nam daje uvid na važnost pogreške.

Na primjer pogreška od 10-grama je samo 0.0125% težine čovjeka od 80kg, ali je 33.3% težine miša od 30g.

Kod izvještaja eksperimentalnih rezultata mora se napraviti razlika između «preciznost» i «točnost» Točnost mjeri koliko je vrijednost koju smo mjerili blizu istinitoj vrijednosti. Tako je visoko preciznom mjerenju pridružena veoma mala pogreška. Uočimo da je u eksperimentalnom radu istinita vrijednost često nepoznata, i zato ono što izvještavamo mora biti procijenjeno točno. Preciznost je mjera ponavljanja i razlučivosti mjerenja, to jest, najmanja šansa u količini mjerenja koja se može pouzdano otkriti. Visoko precizna eksperimentalna oprema može neprekidno mjeriti veoma male razlike u fizičkim količinama. Iako visoko precizno mjerenje može biti nepouzdan, visoka preciznost u mjerenju je nužna ali nije dovoljan uvjet za točnost.

Pogreške se mogu podijeliti u dvije kategorije: Sustavna pogreška (eng. *systematic error*) u mjerenju je dosljedna i tvori ponavljajuće odstupanje ili pomak od istinite vrijednosti. Ona je tipični rezultat opreme, ili problema sa eksperimentalnim postupkom. [CORN2007] S druge strane varijacije između uspješnih mjerenja pod naizgled sličnim okolnostima su nazvane slučajne pogreške (eng. *random errors*). Slučajna odstupanja se mogu pojaviti ili u fizičkim veličinama koje su mjerene, ili u procesu mjerenja ili u oboje. Uz ova dva generalna tipa pogreške postoji i pogreška izazvana ljudskim faktorom i nju nazivamo ljudska pogreška.

Ljudske pogreške (propusti) se javljaju kada osoba koja provodi eksperiment napravi propust. Na primjer kada postavite eksperiment pogrešno, kada mjerite nešto ili kada napravite pogrešku u izračunu. Ljudske pogreške nisu izvor eksperimentalnih pogreški, nego one baš spadaju u osobne pogreške provoditelja eksperimenta.

Sustavna pogreška je pogreška nastala u postavljanju eksperimenta koja uzrokuje krive rezultate svaki put, to jest uzrokuje da rezultati uvijek budu preveliki ili premali. Jedan primjer sustavne pogreške bio bi pokušati izmjeriti vrijeme pada ping pong loptice, ne bi li odredili ubrzanje uslijed gravitacije. Otpor zraka bi smanjio mjerenje ubrzanja uzrokujući sistematsku pogrešku. Neke sustavne pogreške se mogu lako ispraviti. Na primjer, ako vaga očitava 0.25 kada nema

mase na njoj, to bi uzrokovalo sistematsku pogrešku pa bi za svaku masu koju mjerimo svi rezultati iznosili 0.25g više. To se može ispraviti tako da vagu postavimo na nulu. Druge se sustavne pogreške mogu jedino eliminirati koristeći druge eksperimentalne postavke. Mnogi jednostavni eksperimenti koje provodite imati će neke sustavne pogreške.

Svi eksperimenti imaju slučajnu pogrešku, koja se javlja zbog toga što se nijedno mjerenje ne može napraviti s potpunom točnošću. Slučajne pogreške će uzrokovati seriju mjerenja koja su nekada prevelika, a nekada premala. Primjer slučajne pogreške bio bi kada mjerimo štopericom. Nekada će se desiti da je zaustavimo prerano, a nekada prekasno. Oba slučaja uzrokuju slučajnu pogrešku u vašem mjerenju. (Primijetimo da kad čovjek sudjeluje u procesu mjerenja, on uvijek uzrokuje važeću eksperimentalnu pogrešku koja nije u definiciji ljudske pogreške. Dakle, vrijeme čovjekova reagiranja nije pogreška, to je u biti ograničenje dijela eksperimentalnog procesa kad čovjek vrši mjerenje). Slučajna pogreška se može smanjiti tako da učinimo nekoliko mjerenja.

3.2.1. Analiza pogreške

Jedan način da analiziramo pogrešku je tako da izračunamo postotak pogreške (eng. *%error calculation*). Postotak pogreške je koristan kada imamo jedan rezultat koji želimo usporediti sa standardnom vrijednošću, ili kada imamo dvije vrijednosti eksperimenta postignute različitim značenjima koje želimo usporediti. Postotak pogreške računamo prema sljedećoj formuli:

$$\% \text{ pogreška} = \frac{\text{expt.}\# - \text{std.}\#}{\text{std.}\#} \times 100\% \quad (2)$$

Eksperimentalnu vrijednost smo označili s «expt.#», a standardnu ili nagovještenu vrijednost s «std.#». Koristeći ovu formulu, pozitivni postotak pogreške govori nam da su rezultati veći od standardnih, dok negativni postotak pogreške govori da su rezultati manji od standardnih. Iako nam postotak pogreške govori o odnosnoj veličini naše pogreške, ne daje nam ništa o vrsti te pogreške (da li je slučajna ili sustavna). Za određivanje vrste pogreške u nekim slučajevima možemo koristiti standardnu devijaciju (eng. *standard deviation*), koja se obično označava malim grčkim slovom sigma (σ), ili kraticom stand.dev. Ona nam ukazuje na to da li je pogreška sistematska ili slučajna. U biti nam standardna devijacija govori kako su podaci slučajni ili raspršeni. Ona se ne primjenjuje na populaciju, već na uzorak koji predstavlja dio te iste populacije. S obzirom da uzorak sadržava sve karakteristike populacije iz koje je on odabran, zaključak se može generalizirati, odnosno poopćiti i na populaciju. Pri tom valja imati na umu da, kako je uzorak manji dio populacije, zaključivanje o populaciji na osnovi podataka o uzorku, sadrži pogrešku. Mala standardna devijacija, povezana s mjerenim brojem, ukazuje na mali stupanj slučajne pogreške, dok velika standardna devijacija ukazuje na veliki stupanj slučajne pogreške. Relativnu veličinu (postotak) standardne devijacije računamo sljedećom formulom:

$$\% = \frac{\sigma}{\text{expt.}\#} \times 100\% \quad (3)$$

Standardnu devijaciju možemo koristiti i za usporedbu sa standardnom vrijednošću. Ako standardna vrijednost leži između jedne ili dvije standardne devijacije eksperimentalnog rezultata, vaša eksperimentalna pogreška je većim dijelom slučajna pogreška. Ako je standardna vrijednost veća od dvije ili tri standardne devijacije nego vaš rezultat, pogreška je uglavnom sistematska. Konkretno, iz teorije statistike, ako je standardna vrijednost 1σ od vaše vrijednosti, postoji samo 68.3% šanse da je pogreška sistematska a ne slučajna, a ako je 3σ šansa je 99.7%. Količinu σ razlike između eksperimentalne i standardne vrijednosti možemo odrediti koristeći sljedeću formulu:

$$\#of\sigma = \frac{\text{expt.}\# - \text{std.}\#}{\sigma} \quad (4)$$

Dakle, kada imamo niz mjerenja i možemo izvesti analizu pogreške uključujući izračun postotka pogreške i standardnu devijaciju, možemo dobiti uvid u prirodu eksperimentalnih pogreški:

- Mali postotak pogreške, standard od jedan do dva sigma, ukazuje na male, uglavnom slučajne pogreške
- Veliki postotak pogreške, standard od dva ili tri sigma, ukazuje na male, uglavnom sustavne pogreške.
- Veliki postotak pogreške, standard od jedan ili dva sigma ukazuje na velike uglavnom slučajne pogreške
- Veliki postotak pogreške, standard koji nije između dva ili tri sigma ukazuje na velike uglavnom sustavne pogreške

4. Vrednovanje učinka učenja i poučavanja uz pomoć sustava xTEx-Sys

U ovom poglavlju ćemo govoriti o eksperimentu kojeg smo proveli u svrhu vrednovanja učinka učenja i poučavanja uz pomoć sustava xTEx-Sys. Nakon kratkog opisa i uvođenja u sustav xTEx-Sys navesti ćemo kako je izgledala sama priprema eksperimenta. Naglasimo da su u sljedećim poglavljima opisana dva eksperimenta, jedan sa šestim razredom i jedan sa osmim razredom osnovne škole “Gradac“, Gradac. Zatim ćemo detaljno opisati tijek istraživanja eksperimenta za oba razreda i način na koji smo učenike upoznali sa sustavom. Analize podataka ćemo razdvojiti na dva poglavlja, posebno za šesti, posebno za osmi razred. Na samom kraju obradit ćemo rezultate anketa danih eksperimentalnim grupama. Te rezultate ćemo prikazati grafički. Za razliku od analize podataka, rezultate anketa ćemo prikazati za oba razreda zajedno. Glavni razlog tome je mali broj učenika koji su sudjelovali u eksperimentu. Svi testovi (inicijalni testovi, među testovi, završni testovi i ankete) koji su dani učenicima neće biti prikazati u ovom poglavlju, već ih se može vidjeti u 7. poglavlju.

4.1. Sustav xTEx-Sys

Web orijentirana inteligentna autorska ljuska eXtended Tutor-Expert System (xTEx-Sys) je računalni sustav čija je svrha poboljšati proces učenja i poučavanja učenika, te povećati efikasnost rada učitelja [STAN2005]. 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. Glavna razlika u odnosu na ostale načine učenja i poučavanja je da se ovom sustavu može pristupiti s bilo koje lokacije i u bilo koje vrijeme. Sudionici ovog programskog sustava su učenici, radi stjecanja znanja i vještina, učitelji, radi didaktičkog oblikovanja nastavnih sadržaja, stručnjaci za područna znanja, radi izgradnje baza znanja, administrator sustava, radi nadzora funkcija sustava, korisnika sustava kao i načina korištenja sustava. Temeljne funkcije autorske ljuske su oblikovanje baza znanja za različita područna znanja, oblikovanje i pristup nastavnim sadržajima temeljenim na bazama područnih znanja, učenje i poučavanje, testiranje i ocjenjivanje znanja i nadzor sustava.

4.2. Priprema eksperimenta

Testiranje na sustavu xTEx-Sys je provedeno za područno znanje iz matematike za šesti i osmi razred. Cilj testiranja je bilo utvrđivanje učinka korištenja sustava xTEx-Sys, te postignuća učenika u procesu učenja i podučavanja.

Prije provedbe eksperimenta, razvili smo baze znanja iz područja geometrije za šesti i osmi razred, te izgradili nastavne sadržaje u sustavu xTEx-Sys. Baza znanja za šesti razred (četverokut) sadrži 64 čvora, a povezana je vezama *dio od*, *vrsta od*, *primjerak od* i *svojstvo*, dok baza znanja za osmi razred (geometrijska tijela) sadrži 71 čvor koji su povezani vezama *vrsta od*, *primjerak od*, *dio od*, *svojstvo* i *prikaz*. Većina čvorova imaju pridruženu sliku ili prezentaciju, osnovne značajke, te kratak opis.

U eksperimentu je sudjelovalo 18 učenika šestog razreda i 18 učenika osmog razreda OŠ "Gradac", Gradac. Na temelju rezultata inicijalnog testa, učenike svakog razreda smo podijelili u dvije grupe, eksperimentalnu i kontrolnu.

Eksperimentalna grupa je za učenje i podučavanje koristila sustav xTEx-Sys, dok je kontrolna grupa sudjelovala u tradicionalnom procesu učenja i podučavanja. Učenici šestog razreda su bili uključeni u odgovarajuće oblike učenja i poučavanja po dva sata tjedno u razdoblju od 26.04.2007.g. do 08.06.2007.g., odnosno 14 sati, dok je s osmim razredom bilo više problema u organizaciji zbog ekskurzije i raznih obaveza vezanih za kraj njihovog osnovnoškolskog obrazovanja. Za njih je eksperiment trajao od 23.04.2007.g. do 14.06.2007.g., sveukupno 12 sati.

U oba razreda su eksperimentalna i kontrolna grupa pisale inicijalni test (**prilog a i b**), jedan među test (**prilog c i d**), i završni test (**prilog e i f**), s tim da su učenici eksperimentalne grupe u oba razreda ispunjavali anketu očekivanja prije učenja i podučavanja i anketu mišljena o sustavu nakon učenja i podučavanja. Bilo je predviđeno pisati i drugi među test, no to nije realizirano zbog gore navedenih problema u organizaciji, zbog loše opremljenosti računalnom opremom u školi, te učestalog pada mreže koji je uvelike otežavao pristup učenicima na sustav xTEx-Sys. Inicijalni test nam je dao informaciju o postojanju statistički značajne razlike u predznanju učenika, a završni test informaciju o postojanju značajnih razlika između grupa s obzirom na različit način podučavanja.

4.3. Tijek istraživanja

Istraživanje je provedeno od 23.04.2007.g. do 14.06.2007.g., a način na koji je organizirano je prikazan u Tablici 4.1. i u Tablici 4.2.

U prvom tjednu je proveden inicijalni test koji nam je omogućio podjelu učenika u dvije ekvivalentne grupe, eksperimentalnu i kontrolnu. Kontrolne grupe u oba razreda su se do kraja eksperimenta podučavale na tradicionalni način.

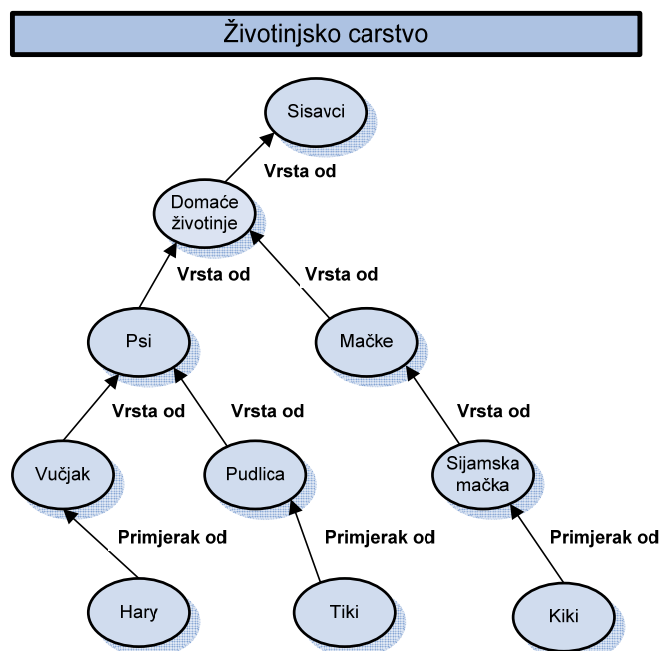
Tablica 0.1. Tijek istraživanja za šesti razred

Aktivnost 6.razred	Datum
Inicijalni test	26.04.2007.
Upoznavanje eksperimentalne grupe sa sustavom xTeX-Sys	04.05.2007.
Učenje i poučavanje uz pomoć sustava xTeX-Sys	od 11.05.2007.do 18.05.2007.
Međutest	18.05.2007.
Učenje i poučavanje uz pomoć sustava xTeX-Sys	od 25.05.2007 do 08.06.2007.
Završni test	08.06.2007.

Tablica 0.2. Tijek istraživanja za osmi razred

Aktivnost 8.razred	Datum
Inicijalni test	23.04.2007.
Upoznavanje eksperimentalne grupe sa sustavom xTeX-Sys	10.05.2007.
Učenje i poučavanje uz pomoć sustava xTeX-Sys	od 17.05.2007.do 24.05.2007.
Međutest	24.05.2007.
Učenje i poučavanje uz pomoć sustava xTeX-Sys	od 31.05.2007 do 14.06.2007.
Završni test	14.06.2007.

Eksperimentalna grupa šestog razreda se u drugom tjednu upoznala sa sustavom xTeX-Sys, dok je eksperimentalna grupa osmog razreda kasnila jedan tjedan za njima jer su taj tjedan bili na ekskurziji. Prikazana im je prezentacija o različitim načinima učenja i podučavanja, od klasičnog načina učenja, pa do učenja pomoću računala. Također im je objašnjen pojam semantičke mreže pomoću Slike 4.1., koja se odnosi na područno znanje o životinjama. To im je bilo važno objasniti jer je gradivo u sustavu xTeX-Sys prikazano skupom međusobno povezanih pojmova, nazvanih čvorovima.



Slika 0.1. Primjer semantičke mreže

Nakon što su učenici usvojili ovakav način povezivanja čvorova, demonstriran im način učenja i poučavanja u sustavu xTex-Sys na sadržaju iz geometrije iz prethodnih razreda. Nakon toga su ispunili anketu očekivanja čije ćemo rezultate analizirati u poglavlju 4.6.1.4.6.1. Analiza ankete očekivanja.

Treći tjedan je svaki učenik dobio svoje korisničko ime i lozinku, te im je objašnjen način na koji će se prijaviti na sustav, pojedini odnosi među pojmovima iz nastavnog sadržaja geometrija za šesti razred, te geometrijska tijela za osmi razred. Učenici su se samostalno prijavili na sustav i započeli proces individualnog učenja i poučavanja. Učenicima je naglašeno da se mogu prijaviti na sustav od kuće tako da im učenje i poučavanje nije bilo ograničeno na samo tri školska sata, koliko smo izdvojili za prvi dio nastavnog saržaja, ne računajući sate upoznavanja sa sustavom.

Idući tjedan su učenici jedan školski sat proveli u ponavljanju gradiva, što je bilo važno zbog velike vremenske razlike između blok-satova, da bi nakon toga pristupili prvom međutestu.

Učenici šestog razreda su sljedeća dva tjedna učili i poučavali se na sustavu, te nakon sveukupno osam školskih sati učenja i poučavanja pisali završni test. Kod učenika osmog razreda se stalno javljao problem što ih je bilo teško sve okupiti pa su oni imali dva sata manje za učenje i poučavanje prije završnog testa, što će se očitovati i u rezultatima koje ćemo obraditi u poglavlju 4.5.. Nakon svega toga učenici su ispunili anketu o njihovom mišljenju o sustavu xTeX-Sys, te načinu učenja i testiranju na taj način. Rezultati ove ankete će biti analizirani u poglavlju 4.6.2..

4.4. Analiza rezultata za šesti razred

Ukupno je 18 učenika šestog razreda pisalo sve testove. U Tablici 4.3., su dani rezultati inicijalnog testa, međutesta i završnog testa koje su pisali učenici šestog razreda. Najveći broj bodova koji je učenik mogao postići je 100.

Prije početka eksperimenta učenicima je dan inicijalni test koji nam je omogućio podjelu učenika u dvije grupe između kojih ne postoji statistički značajna razlika.

Vrijednosti dobivene analizom inicijalnog testa su prikazane u tablici 4.4. Broj stupnjeva slobode je 16. Kako je granična vrijednost 1.746 veća od apsolutne vrijednosti dobivene t-vrijednosti koja iznosi -0,0731, zaključujemo da su grupe statistički ekvivalentne, tj. ne postoji statistički značajna razlika između grupa s obzirom na rezultate inicijalnog testa.

Definirali smo dvije nul-hipoteze, jednu za međutest i jednu za završni test. Da bi hipoteze mogli prihvatiti ili odbaciti, potrebno je izračunati razliku između rezultata testova i inicijalnog testa za kontrolnu i eksperimentalnu grupu, te t-vrijednost da bi se moglo potvrditi ili opovrgnuti odgovarajuća nul-hipoteza. Račun proveden nad razlikama u rezultatima prikazan je za svaku hipotezu posebno.

Tablica 0.3. Rezultati testova za šesti razred

	Inicijalni test Bodovi	Međutest bodovi	Završni test Bodovi
Eksperimentalna grupa			
Učenik1	34	40	20
Učenik2	73	52	70
Učenik3	73	43	70
Učenik4	74	40	30
Učenik5	75	47	55
Učenik6	77	45	30
Učenik7	83	97	30
Učenik8	79	41	50
Učenik9	93	84	90
Kontrolna grupa			
Učenik10	60	23	20
Učenik11	62	22	15
Učenik12	65	45	35
Učenik13	82	49	40
Učenik14	82	30	30
Učenik15	74	23	20
Učenik16	75	46	15
Učenik17	78	51	30
Učenik18	79	59	65

Tablica 4.0.4. Podaci potrebni za ispitivanje statistički značajne razlike

	Kontrolna grupa	Eksperimentalna grupa
\bar{X}	73	73.44
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}}$	8.529	16.109
$s_{\bar{X}_c - \bar{X}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	6.076	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{X}_c - \bar{X}_e}}$	-0.0731	

Hipoteza H_1 je glasila: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata međutesta“

Tablica 0.5 Podaci potrebni za prihvaćanje ili odbacivanje nul-hipoteze H1

	Kontrolna grupa	Eksperimentalna grupa
\bar{X}	-34.3333	-19.1111
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}}$	11.8532	18.6242
$s_{\bar{X}_c - \bar{X}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	7.3588	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{X}_c - \bar{X}_e}}$	-2.0686	

S obzirom da je apsolutna vrijednost dobivene t-vrijednosti veća od granične t-vrijednosti koja iznosi 1.746, dolazimo do zaključka da se nul-hipoteza odbacuje, tj. razlika je statistički značajna. S obzirom da je predznak t-vrijednosti negativan zaključujemo da je *eksperimentalna grupa statistički značajno bolja od kontrolne grupe u među testu*.

Hipoteza H2 je glasila: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata završnog testa“

Tablica 0.6 Podaci potrebni za prihvaćanje ili odbacivanje nul-hipoteze H2

	Kontrolna grupa	Eksperimentalna grupa
\bar{X}	-43	-24
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}}$	13.9284	20.1059
$s_{\bar{X}_c - \bar{X}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	8.1530	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{X}_c - \bar{X}_e}}$	-2.3304	

Kao je apsolutna vrijednost dobivene t-vrijednosti veća od granične t-vrijednosti koja iznosi 1.746, dolazimo do zaključka da se nul-hipoteza odbacuje, tj. razlika je statistički značajna. S obzirom da je predznak t-vrijednosti negativan zaključujemo da je *eksperimentalna grupa statistički značajno bolja od kontrolne grupe u završnom testu*.

Veličinu učinka možemo izračunati na više načina, u našem eksperimentu koristit ćemo razliku između aritmetičkih sredina kontrolne i eksperimentalne grupe koju dijelimo standardnom devijacijom kontrolne grupe:

$$\Delta = \frac{X_c - X_e}{\sigma_c} \quad (5)$$

Tako smo dobili parcijalne veličine učinka za među test 1.284 , a za završni test 1.364. Ukupnu veličinu učinka dobijemo kao prosječnu vrijednost parcijalnih veličina učinka

$$\Delta = \frac{\Delta_{mt} + \Delta_{zt}}{2} \quad (6)$$

U našem slučaju za šesti razred ona iznosi 1.324

4.5. Analiza rezultata za osmi razred

Ukupno je 18 učenika osmog razreda pisalo sve testove. U Tablica 0.7. Rezultati testova za osmi razred su dani rezultati inicijalnog testa, međutesta i završnog testa koje su pisali učenici osmog razreda. Najveći broj bodova koji je učenik mogao postići je 100.

Tablica 0.7. Rezultati testova za osmi razred

	Inicijalni test Bodovi	Međutest bodovi	Završni test Bodovi
Eksperimentalna grupa			
Učenik19	42	50	35
Učenik20	48	75	80
Učenik21	83	100	100
Učenik22	70	83	100
Učenik23	56	60	90
Učenik24	50	45	65
Učenik25	40	45	30
Učenik26	35	27	30
Učenik27	30	45	100
Kontrolna grupa			
Učenik28	85	95	100
Učenik29	70	65	90
Učenik30	56	60	80
Učenik31	51	75	80
Učenik32	48	60	45
Učenik33	44	75	75
Učenik34	38	30	30
Učenik35	33	30	50
Učenik36	30	25	60

Prije početka eksperimenta učenicima je dan inicijalni test koji nam je omogućio podjelu učenika u dvije grupe između kojih ne postoji statistički značajna razlika.

Tablica 4.0.8. Podaci potrebni za ispitivanje statistički značajne razlike

	Kontrolna grupa	Eksperimentalna grupa
\bar{X}	50.556	50.444
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}}$	17.791	17.016
$s_{\bar{X}_c - \bar{X}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	8.206	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{X}_c - \bar{X}_e}}$	0.013	

Vrijednosti dobivene analizom inicijalnog testa su prikazane u tablici 4.8. Granična F-vrijednost iznosi 3.14 za alfa 0.05., a broj stupnjeva slobode je 16. Kako je granična t-vrijednost 1.746 veća od apsolutne vrijednosti dobivene t-vrijednosti koja iznosi 0,013, zaključujemo da su grupe statistički ekvivalentne, tj. ne postoji statistički značajna razlika između grupa s obzirom na rezultate inicijalnog testa.

Definirali smo dvije nul-hipoteze, jednu za međutest i jednu za završni test. Račun proveden nad razlikama u rezultatima prikazan je za svaku hipotezu posebno.

Hipoteza H3 je glasila: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata međutesta“

Tablica 0.9 Podaci potrebni za prihvatanje ili odbacivanje nul-hipoteze H3

	Kontrolna grupa	Eksperimentalna grupa
\bar{X}	6.6667	8.4444
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}}$	13.7840	10.9785
$s_{\bar{X}_c - \bar{X}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	5.8739	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{X}_c - \bar{X}_e}}$	-0.3027	

S obzirom da je apsolutna vrijednost dobivene t-vrijednosti manja od granične t-vrijednosti koja iznosi 1.746, dolazimo do zaključka da se nul-hipoteza prihvaća, tj. razlika nije statistički

značajna. S obzirom da je predznak t-vrijednosti negativan zaključujemo da je eksperimentalna grupa statistički neznačajno bolja od kontrolne grupe u međutestu.

Hipoteza H4 je glasila: „Ne postoji statistički značajna razlika između kontrolne i eksperimentalne grupe kod rezultata završnog testa“

Tablica 0.10 Podaci potrebni za prihvaćanje ili odbacivanje nul-hipoteze H4

	Kontrolna grupa	Eksperimentalna grupa
\bar{X}	17.2222	19.5556
$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}}$	14.1225	25.5592
$s_{\bar{X}_c - \bar{X}_e} = \sqrt{\frac{\sigma_c^2}{N_c} + \frac{\sigma_e^2}{N_e}}$	9.7338	
$t = \frac{\bar{X}_c - \bar{X}_e}{s_{\bar{X}_c - \bar{X}_e}}$	-0.2397	

Kako je apsolutna vrijednost dobivene t-vrijednosti veća od granične t-vrijednosti koja iznosi 1.746, dolazimo do zaključka da se nul-hipoteza prihvaća, tj. razlika nije statistički značajna. S obzirom da je predznak t-vrijednosti negativan zaključujemo da je eksperimentalna grupa statistički neznačajno bolja od kontrolne grupe u međutestu.

Parcijalne veličine učinka računamo po formuli:

$$\Delta = \frac{X_c - X_e}{\sigma_c} \quad (7)$$

i one iznose 0.129 za međutest i 0.165 za završni test. Ukupnu veličinu učinka računamo po formuli:

$$\Delta = \frac{\Delta_{mt} + \Delta_{zt}}{2} \quad (8)$$

i ona iznosi 0.147 za osmi razred..

Primijetimo da je u šestom razredu eksperimentalna grupa bila značajno bolja od kontrolne u oba dva testa, dok u osmom razredu to nije slučaj. Takvi rezultati su se mogli i očekivati zbog toga što su učenici šestog razreda mnogo ozbiljnije shvatili svoj zadatak, trudili su se što više učiti pomoću sustava xTeX-Sys i natjecali se međusobno tko će postići bolje rezultate. Nadalje, bilo ih je lako okupiti i rado su ostajali nakon nastave, dok s učenicima osmog razreda to nije bio slučaj. Njima je bio kraj školske godine i koncentracija im je sve više popuštala. Manje su učili, što se očitovalo ne samo u ovom eksperimentu, nego i u drugim predmetima, te su često izostajali iz škole. Isto tako su u manjoj mjeri pristupali sustavu sa kućnih računala.

4.6. Analiza anketa danih eksperimentalnim grupama

Anketne upitnike ispunilo je ukupno 18 učenika, od čega 9 učenika šestog razreda i 9 učenika osmog razreda. U prvoj anketi, koju smo nazvali *anketa očekivanja*, cilj je bio prikupiti njihova očekivanja o primjeni računala u nastavi, odnosno, o primjeni sustava xTeX-Sys, a u drugoj, *završnoj anketi*, njihovo mišljenje o sustavu. Analize anketa nisu rađene posebno za djevojčice i posebno za dječake zbog malog broja učenika.

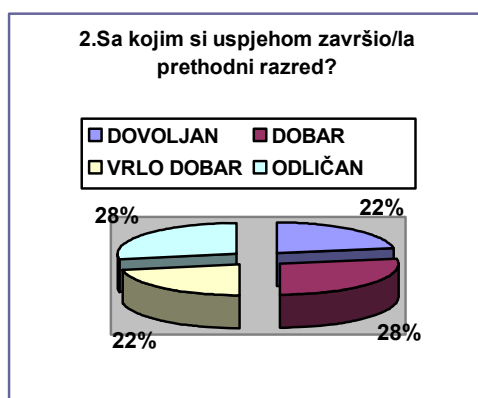
4.6.1. Analiza ankete očekivanja

Svrha ove ankete bila je saznati posjeduju li učenici računalo kod kuće, koliko se služe s njim i u koji svrhu, te kakva su njihova očekivanja o učenju i poučavanju uz pomoć sustava xTeX-Sys. Rezultati ankete biti će prikazani grafički, a s obzirom na mali broj učenika prikazati ćemo ih zajedno za učenike osmog i za učenike šestog razreda.



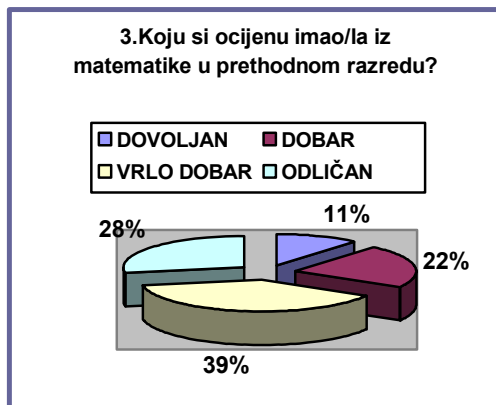
Graf 4. 1

Što se tiče uspjeha na kraju prošle školske godine, **Error! Reference source not found.** nam pokazuje da je podjednak broj učenika prošao s odličnim i dobrim uspjehom, te podjednak broj učenika s dovoljnim i vrlo dobrim uspjehom. Dakle 28% učenika je imalo odličan uspjeh, 22% vrlo dobar, 28% dobar i 22% dovoljan.



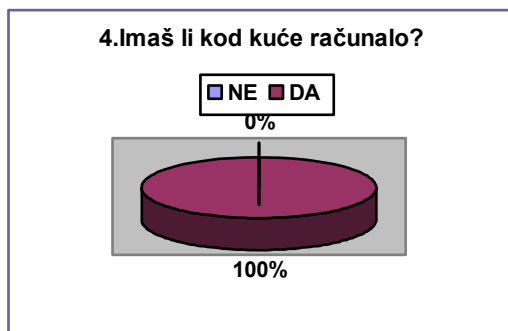
Graf 4. 2

Iz **Error! Reference source not found.** je vidljivo da je većina učenika u prethodnom razredu imala matematiku vrlo dobar, čak 39% učenika, ocijenu odličan imalo je 28% učenika, dobar 22%, a dovoljan samo 11% učenika.



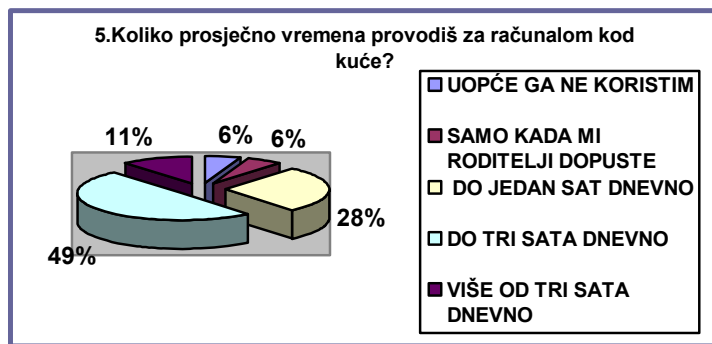
Graf 4. 3

Iz **Error! Reference source not found.** je vidljivo da svi učenici imaju kod kuće računalo, dakle svi su bili u mogućnosti od kuće pristupiti sustavu i poučavati se.



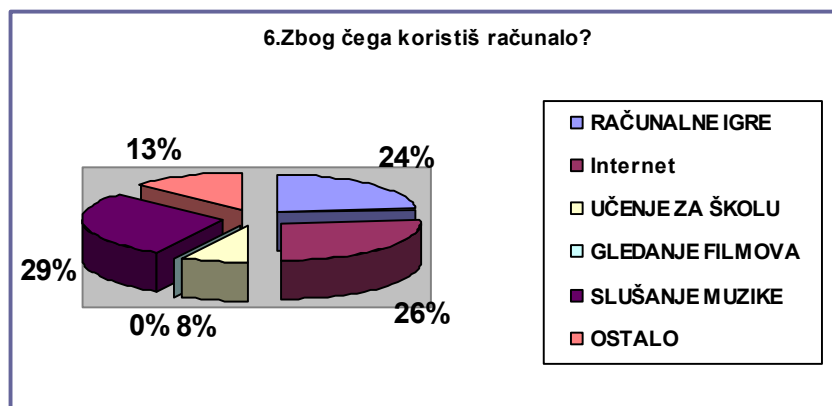
Graf 4. 4

Error! Reference source not found. nam pokazuje da učenici jako često koriste računalo, čak 49% učenika koristi računalo do tri sata dnevno, 11% ga koristi više od tri sata dnevno, 28% do jedan sat dnevno, a po 6% ga uopće ne koriste ili ga koriste samo kada im roditelji dopuste.



Graf 4. 5

Na 6. pitanje učenicima je bilo dopušteno odabrati više odgovora. Iz **Error! Reference source not found.** je lako odmah uočiti da učenici najčešće koriste računalo za slušanje muzike (29%), interneta (26%), te zbog igrica (24%).



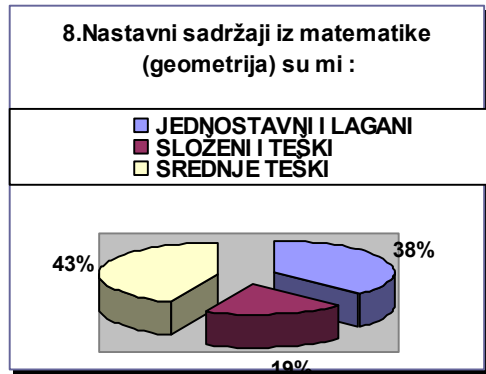
Graf 4. 6

Iz **Error! Reference source not found.** je vidljivo, da što se tiče njihove procjene samih sebe o sposobnostima i znanju korištenja računalom većina učenika je dalo odgovor -dobrim, što je i bilo za očekivati. Isto tako je za očekivati da nema odgovora -lošim zbog toga što svi učenici imaju računalo kod kuće pa je za pretpostaviti da imaju bar osnovna znanja korištenja računala.



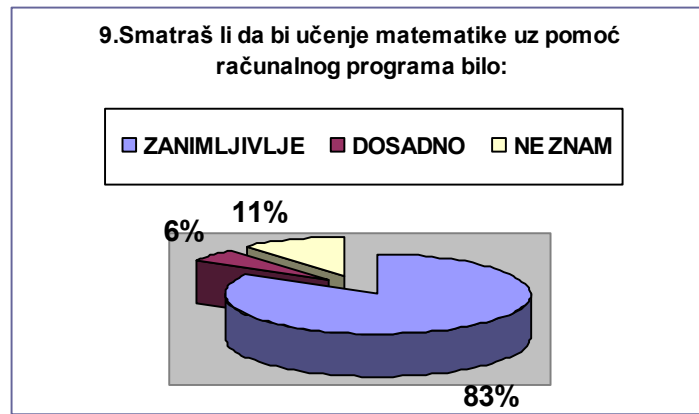
Graf 4. 7

Nastavne sadržaje iz matematike, **Error! Reference source not found.**, je većina, 43% učenika ocijenilo srednje teškima, 38% jednostavnim i laganim, a 19% složenim i teškim. Ako posebno pogledamo rezultate osmog i šestog razreda uočavamo da su učenici osmog razreda većinom ocijenili gradivo geometrije težim, dok su ga učenici šestog razreda ocijenili kao jednostavnim i laganim, što je bilo za očekivati jer im se gradivo uvelike razlikuje. Učenici šestog razreda nadograđuju znanje iz geometrije iz prethodnih razreda, dok se učenici osmog razreda prvi put susreću s geometrijskim tijelima, što mnogima predstavlja veliki problem.

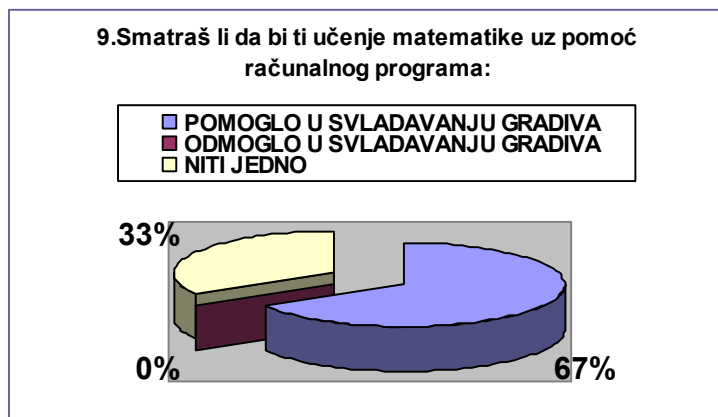


Graf 4. 8

Zadnja dva pitanja su dala jako pozitivne rezultate, naime velik broj učenika, čak 83% njih smatra da bi učenje matematike uz pomoć računalnog programa bilo zanimljivlje nego tradicionalan način učenja, dok 67% učenika smatra da bi im određeni računalni program čak i pomogao u svladavanju gradiva iz matematike.



Graf 4. 9



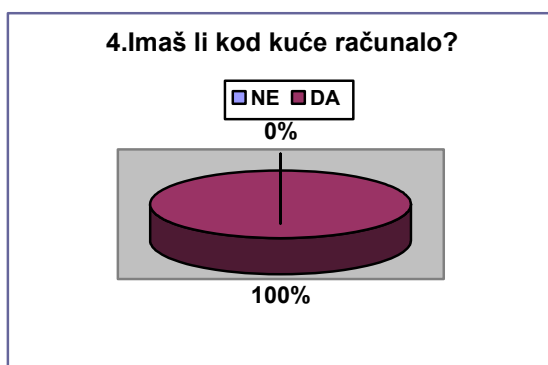
Graf 4. 10

4.6.2. Analiza završne ankete

Način na koji su učili učenici eksperimentalne skupine bio je bitno drugačiji od klasičnog načina učenja i poučavanja kojim se koristila kontrolna skupina. Upravo iz toga razloga dana im je anketa i nakon završetka eksperimenta. S tom anketom smo željeli saznati koliko se učenikovo mišljenje o sustavu xTeX-Sys promijenilo nakon završetka eksperimenta. Rezultate ankete ćemo prikazati zajedno za šesti i osmi razred, zbog malog broja učenika. Anketu su ispunili svi učenici koji su i pisali anketu očekivanja.

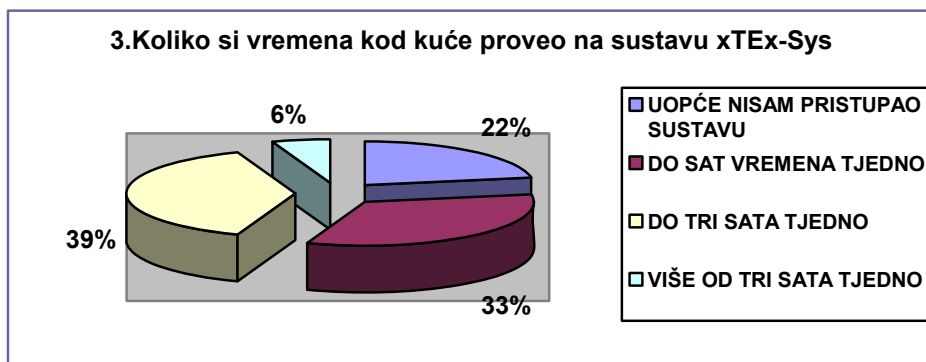


Graf 4. 11



Graf 4. 12

Iz **Error! Reference source not found.** vidimo da je veliki broj učenika pristupao sustavu od kuće, no upravo zbog velike razlike u odgovorima između šestog i osmog razreda, razbili smo ga u dva grafa, za svaki razred posebno. Dakle, iz **Error! Reference source not found.** i **Error! Reference source not found.** je vidljivo da je više od pola učenika šestog razreda, čak 56%, pristupalo sustavu do tri sata tjedno, dok 45% učenika osmog razreda uopće nije pristupilo sustavu od kuće. Smatram da je upravo taj nedostatak zainteresiranosti kod učenika osmog razreda uvelike utjecao na nepostojanje značajne razlike u rezultatima završnog testa između eksperimentalne i kontrolne grupe.



Graf 4. 13

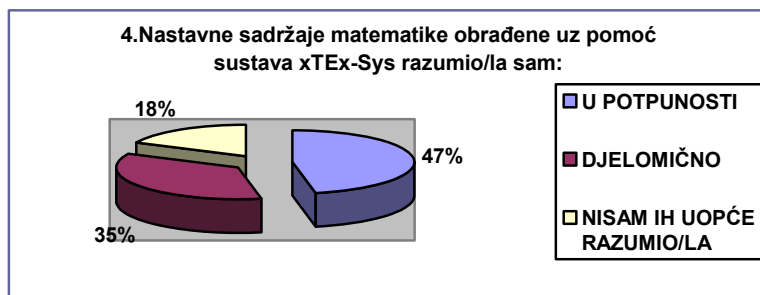


Graf 4. 14



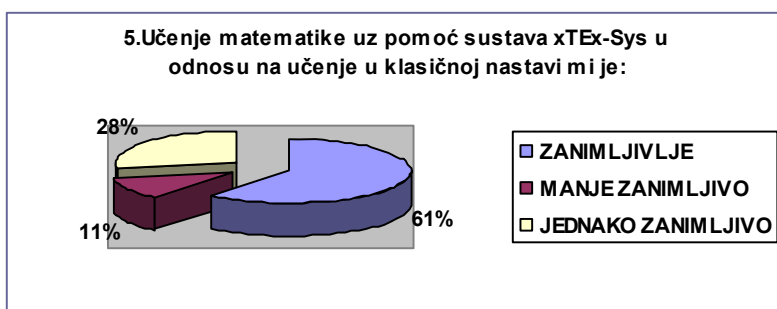
Graf 4. 15

Iz **Error! Reference source not found.** se vidi da je 47% učenika u potpunosti razumjelo sadržaje iz matematike, 35% ih je razumjelo djelomično, a 18% ih nije uopće razumjelo. Bez obzira na to što ih je 47% u potpunosti razumjelo gradivo, mnogo veći broj učenika, 61%, se izjasnio da im je učenje uz pomoć sustava xTeX-Sys zanimljivlje od učenja u klasičnoj nastavi.



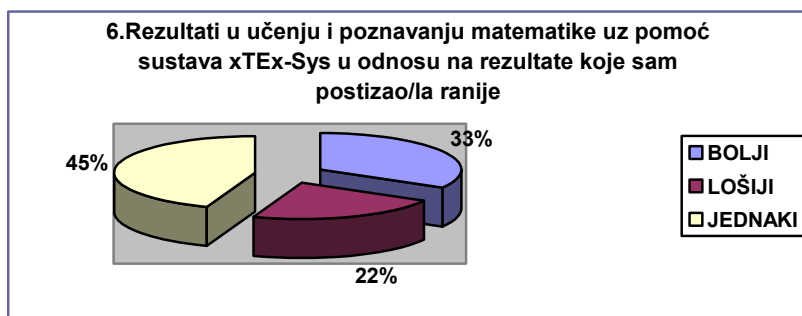
Graf 4. 16

Iako se čak 61% učenika izjasnilo da im je bilo zanimljivlje učiti uz pomoć sustava xTeX-Sys, samo ih 33% misli da su rezultati u odnosu na one ranije bolji. 45% smatra da su rezultati jednaki, dok 22% smatra da ima lošije rezultate.



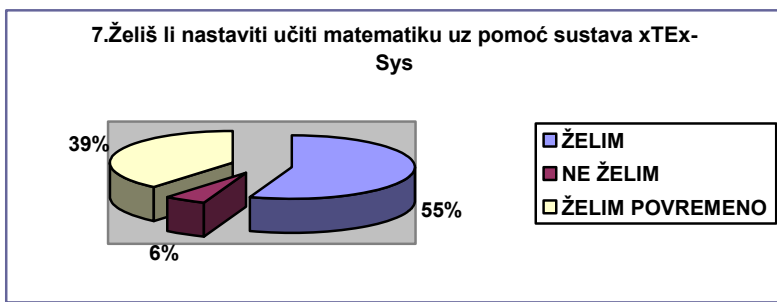
Graf 4. 17

Mislím da je i za **Error! Reference source not found.** važno napomenuti da od učenika šestog razreda 45% smatra da su rezultati bolji, 44% da su rezultati jednaki, a samo 11% smatra da su rezultati lošiji od prijašnjih. Jednak broj učenika osmog razreda, 45%, smatra da su rezultati bolji, dok 22% smatra rezultate istima, a čak 33% smatra da su rezultati lošiji.



Graf 4. 18

Iz **Error! Reference source not found.** se vidi da većina učenika, 55% želi nastaviti učiti matematiku uz pomoć sustava xTeX-Sys, 39% bi to željelo povremeno, kao dodatak tradicionalnoj vrsti nastave, a samo 6% uopće ne bi htjelo nastaviti učiti na ovakav način.



Graf 4. 19

Posljednja dva pitanja u anketi su bila pitanja otvorenog tipa. Od učenika se tražilo da izlože svoje mišljenje o sustavu, te da nabroje neke loše strane i dobre strane učenja matematike pomoću sustava xTeX-Sys. Ovdje su navedeni samo neki od njihovih odgovora:

- „Zanimljivo mi je, mogu učiti s prijateljicom.“
- „Loša strana je ta da u školi kompjuteri često zastaju, inače je super!“
- „Sustav xTeX-Sys je dobar, mogu bolje naučiti matematiku i zanimljiv je.“
- „Lakše je učiti“
- „Loše je što moramo dolaziti kad nema nastave, ali je lakše učiti nego onako u učionici.“
- „Radio sam zato što sam morao, nije mi uopće zanimljivo i ne volim matematiku!“
- „U početku mi je bilo teško, ali poslije kad sam skužila postalo je lako i zanimljivo, bilo bi dobro više ovako učiti.“
- „Ne bih voljela učiti na ovaj način, jer mi je teško bez papira i olovke, ali bi bilo super koristiti program kada nastavnica objasni cjelinu pa mi ovako ponavljamo gradivo.“
- „Ovo je dobro ali mi je teško ovako učiti“
- „Nisam mogao kod kuće učiti jer mi je Internet prespor i u školi Internet često ne valja pa mi je to smetalo.“
- „Ovako učenje mi se sviđa, trebalo bi matematiku zamijeniti s ovim pa bi bilo bolje.“

5. Zaključak

Iako se inteligentni tutorski sustavi sve više koriste u nastavnom procesu, mali broj tih sustava je vrednovan. Upravo zbog toga javlja se sve veća potreba za njihovim vrednovanjem. Eksperimentalni dizajn 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.

Postoje različiti eksperimentalni dizajni kao što su eksperiment, kvazi-eksperimentalni dizajn i ne-eksperiment. Osnovna razlika između ova tri dizajna je postojanje, odnosno ne postojanje slučajne raspodjele subjekata u grupe i postojanje, odnosno ne postojanje kontrolne grupe. Na temelju tekstova raznih autora, napravili smo pregled i klasifikaciju određenog broja različitih eksperimentalnih i kvazi-eksperimentalnih dizajna. Posebnu pozornost smo posvetili eksperimentalnom dizajnu, budući da se on ponajviše koristi u psihološkim i pedagoškim istraživanjima.

Na samom kraju, proveli smo eksperimentalno istraživanje u svrhu vrednovanja učinka učenja i poučavanja uz pomoć sustava xTEx-Sys. Cilj nam je bio ispitati da li postoji, i ako postoji kolika je, statistički značajna razlika između učenja i poučavanja klasičnim putem i putem sustava xTEx-Sys. Za šesti razred smo zaključili da je eksperimentalna grupa statistički značajno bolja od kontrolne grupe u među testu i u završnom testu, dok smo za osmi razred dobili da je eksperimentalna grupa statistički neznačajno bolja od kontrolne grupe u međutestu, dok kod završnog testa ne postoji statistički značajna razlika.

Sudjelovanjem u ovom istraživanju, kao budući učitelj, uvidjela sam koliko ovakav način organizacije može pomoći u radu. Smatram da bi i učenike i učitelje trebalo upoznati s novim tehnologijama i načinima održavanja nastave jer vidim da su učenici jako brzo savladali rad na sustavu i način na koji će se pomoću njega poučavati, te su jako pozitivno reagirali na drugačiji način nastave. Iz anketa i analize rezultata testova se može zaključiti se xTEx-Sys može uspješno koristiti u izgradnji baza znanja, te da su učenici postigli bolje rezultate korištenjem ovog sustava i što je najbitnije da radije prihvaćaju ovakav način poučavanja nego tradicionalni.

U daljnjem radu bih željela koristiti sustav xTEx-Sys kao nadopunu u nastavi. Sustav bi se jako dobro mogao iskoristiti u interakciji satova matematike i informatike, jer uz učenje matematike na taj način, djeca razvijaju i svoje vještine u korištenju računala. S obzirom na moje iskustvo, da su djeca nižeg razreda bolje prihvatila rad na sustavu, počela bih koristiti sustav od petog razreda tako da učenici to prihvate odmah u početku i da im do kraja školovanja takav način rada postane uobičajan. Posebno bih ga koristila na dopunskoj nastavi sa učenicima kojima slabije leži matematika jer sam primjetila da je ovakav način učenja matematike jedan od rijetkih kojima se oni mogu zainteresirati za gradivo.

6. Literatura

- [MUŽI1977] Mužić, V. (1977) Metodologija pedagoških istraživanja, Svjetlost, Sarajevo
- [CORN2007] Cornell: Virtual labs, Real data, URL: http://instruct1.cit.cornell.edu/courses/virtual_lab/index.html, posjećeno 20.12.2007.
- [ENDE2008] Ender, P. (2008) Education 230A, Introduction to Research Design and Statistics Research Validity, URL: <http://gseis.ucla.edu/courses/ed230a2/internal.html> posjećeno 21.08.2008.
- [TROC2008] Trochim W. M. K. (2008) Research methods, URL: <http://www.socialresearchmethods.net>, 10.07.2008.
- [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
- [BUJA1967] Bujas, Z. (1967) Uvod u metode eksperimentalne psihologije, Školska knjiga, Zagreb
- [COOK1979] Cook, T. D., Campbell, D. T. (1979) Quasi-experimentation: Design and analysis issues for field settings, Boston: Houghton-Mifflin
- [HALL2008] Hall R. (2008) Psychology world, URL: http://web.mst.edu/~psyworld/within_subjects.htm, 16.08.2008.
- [BORG1988] Borg, W., Borg, M. (1988) Educational Research: An Introduction, p.680, New York
- [KIDD1980] Kidder, L. H. (1980) Research Methods in Social Relations, p.43
- [PFLE1995] Pfleeger, S.L. Experimental Design and Analysis in Software Engineering ACM SIGSOFT, 1995,p.22-26.

7. Prilozi

- A. Inicijalni test 6.r.
- B. Inicijalni test 8r.
- C. Međutest 6.r.
- D. Međutest 8.r.
- E. Završni test 6.r.
- F. Završni test 8.r.
- G. Anketa očekivanja
- H. Završna anketa

A. Inicijalni test 6.r.

- 1) Nacrtaj kvadrat stranice 3cm.
- 2) Koliko dijagonala ima četverokut?
- 3) Što je romb?
- 4) Koliki je zbroj kutova u trokutu?
- 5) Dijagonale pravokutnika su jednake.
 - (a) točno
 - (b) netočno
- 6) Koliki je zbroj unutarnjih kutova u četverokutu?
- 7) Izračunaj opseg pravokutnika kojem su stranice $a=80\text{mm}$ i $b=20\text{mm}$
- 8) Izračunaj površinu pravokutnika kojem su dane stranica $a=9\text{cm}$ i visina $b=3\text{cm}$.
- 9) Izračunaj površinu kvadrata kojem je opseg 6.8cm .
- 10) U trokutu ABC je $\alpha=\beta=1/2\gamma$. Koliko stupnjeva ima γ ?
- 11) Jedan šiljasti kut pravokutnog trokuta ima 75° . Koliko stupnjeva ima drugi šiljasti kut?
- 12) Kako dijelimo trokute s obzirom na kutove?

B. Inicijalni test 8.r.

- 1) Nacrtaj kvadar ABCDEFGH i na slici istakni pravce AD, BH i GC.
- 2) Nacrtaj kvadar ABCDEFGH. Odredi sve pravce određene vrhovima kvadra koji sadrže neku plošnu dijagonalu kvadra.
- 3) Nacrtaj kocku ABCDEFGH i odredi međusobne položaje pravaca AG i GH.
- 4) Kakav sve može biti međusobni položaj dviju ravnina?
- 5) Koliko pravaca prolazi kroz dvije točke? Nacrtaj!
- 6) Što je to ortogonalna projekcija?
- 7) Na koji način određujemo udaljenost točke od ravnine?
- 8) Duljine bridova kocke su 12cm. Odredi udaljenost točke A od ravnine EBC.
- 9) Kada će se desiti da je ortogonalna projekcija dužine na ravninu točka?
- 10) Kakve su ravnine koje sadrže dvije susjedne strane kvadra?

C. Međutest 6.r.

ZAOKRUŽI JEDAN ILI VIŠE PONUĐENIH ODGOVORA

1. **Kako zovemo dio ravnine omeđen četrima dužinama, uključujući i točke tih dužina?**
 - a) četverokut
 - b) trokut
 - c) pravac
 - d) kvadrat

2. **Kako zovemo dužinu koja spaja nasuprotne vrhove četverokuta?**
 - a) vrh
 - b) stranica
 - c) dijagonala
 - d) točka

3. **Romb je:**
 - a) paralelogram
 - b) trokut
 - c) trapez
 - d) pravac

4. **Kome su sve stranice jednake duljine?**
 - a) kvadratu
 - b) rombu
 - c) paralelogramu
 - d) ništa od navedenog

5. **Što je temeljna značajka pravokutnika?**
 - a) Ima dva para paralelnih stranica
 - b) Dijagonale su jednake duljine
 - c) Sve stranice su jednake duljine
 - d) Dijagonale su međusobno okomite

6. **Koja je formula za opseg pravokutnika?**

7. **Što je temeljna značajka dijagonala romba?**
 - a) Međusobno su okomite
 - b) Protežu se u beskonačnost
 - c) Jednake su duljine
 - d) Ništa od navedenog

8. Kome su kutevi uz osnovicu međusobno jednaki?

- a) pravokutniku
- b) rombu
- c) jednakokračnom trapezu
- d) paralelogramu

9. Zbroj veličina kuteva u četverokutu je?

- a) 90°
- b) 360°
- c) 160°
- d) ništa od navedenog

10. Vrhovi četverokuta su:

- a) pravci
- b) točke
- c) dužine

D. Međutest 8.r.

ZAOKRUŽI JEDAN ILI VIŠE PONUĐENIH ODGOVORA

1. **Kako zovemo geometrijsko tijelo omeđeno s dva sukladna i usporedna n-terokuta i s n paralelograma?**
 - a) četverokut
 - b) n-terostrana prizma
 - c) piramida
 - d) pravilna prizma

2. **Trokut je baza:**
 - a) Peterostrane prizme
 - b) Trostrane prizme
 - c) Piramide
 - d) Pravilne trostrane prizme

3. **Što je od navedenog dio kocke?**
 - a) Brid
 - b) Pravac
 - c) Baza
 - d) Prostorna dijagonala

4. **Kome je baza kvadrat?**
 - a) Trostranoj prizmi
 - b) Pravilnoj četverostranoj prizmi
 - c) Piramidi
 - d) ništa od navedenog

5. **Što je temeljna značajka kocke?**
 - a) Ima dva para paralelnih stranica
 - b) Zadana je duljinom brida
 - c) Uspravna prizma kojoj je baza pravokutnik
 - d) Omeđena je bazom i pobočkama koje su trokuti

6. **Svi bridovi su jednake duljine je temeljna značajka?**
 - a) kocke
 - b) pravilnog tetraedra
 - c) pobočke
 - d) prostorne dijagonale

7. Oplošje je svojstvo:

- a) piramidi
- b) n-terokutu
- c) bazi
- d) prizmi

8. Ako je broj vrhova baze pravilne prizme 3, onda je duljina visine prizme?

- a) 2
- b) 4
- c) 5
- d) ništa od navedenog

9. Koja je formula za oplošje prizme?

10. Koja je formula za volumen piramide?

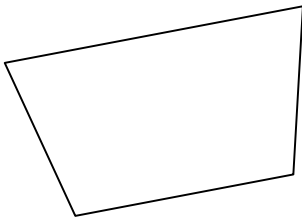
E. Završni test 6.r.

ZAOKRUŽI JEDAN ILI VIŠE PONUĐENIH ODGOVORA

1. Četverokut je:

- a. Geometrijski lik
- b. Trokut
- c. Paralelogram

2. Odredi da li je četverokut na slici paralelogram ili trapez?

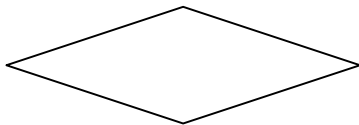


- a. Paralelogram
- b. Trapez

3. Kako zovemo dužinu koja spaja nasuprotne vrhove četverokuta?

- a. vrh
- b. stranica
- c. dijagonala
- d. točka

4. Što se nalazi na slici:

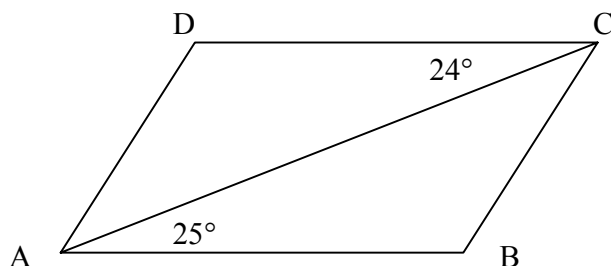


- a. Trapez
- b. Romb
- c. Paralelogram
- d. Ništa od navedenog

5. Koliki je zbroj unutarnjih kutova uz isti krak trapeza?

- a. 160°
- b. 360°
- c. 90°
- d. 180°

6. Jeli četverokut ABCD paralelogram? Obrazloži odgovor.



7. Kolika je veličina četvrtog kuta četverokuta ako su veličine ostalih kutova 53° , 61° , 78° ?

- 168°
- 188°
- 12°
- 360°

8. Što je temeljna značajka pravokutnika?

- Ima dva para paralelnih stranica
- Dijagonale su jednake duljine
- Sve stranice su jednake duljine
- Dijagonale su međusobno okomite

9. Izračunaj površinu paralelograma kojem je $a = 15 \text{ cm}$, $v = 130 \text{ mm}$?

10. Izračunaj površinu paralelograma kojem je $O = 36.2 \text{ cm}$, $b = 9.6 \text{ cm}$ i visina na stranicu a je $v = 4.7 \text{ cm}$?

11. Formula za površinu trapeza je.....

12. Koristimo li se pri računanju površine trapeza duljinama krakova trapeza?

- Da
- Ne

13. Dijagonala trapeza dijeli trapez na dva:

- Trokuta
- Četverokuta
- Pravokutnika
- ništa od navedenog

14. Što je temeljna značajka dijagonala romba?

- a. Međusobno su okomite
- b. Protežu se u beskonačnost
- c. Jednake su duljine
- d. Ništa od navedenog

15. Što nam je potrebno za konstrukciju ?

- a. Kutomjer
- b. Šestar
- c. Ravnalo
- d. Dva trokuta
- e. Ništa od navedenog

F. Završni test 8.r.

1. Preračunaj:

- a) $120\text{dm}^2 = \underline{\hspace{2cm}} \text{cm}^2$
- b) $73\text{dm}^2 = \underline{\hspace{2cm}} \text{m}^2$

2. Kome je baza kvadrat?

- a) Trostranoj prizmi
- b) Pravilnoj četverostranoj prizmi
- c) Piramidi
- d) Ništa od navedenog

3. Koliko prostornih dijagonala ima kvadar?

- a) Jednu
- b) Dvije
- c) Tri
- d) Ništa od navedenog

4. Što su strane kocke?

- a) Pravokutnici
- b) Kvadrati
- c) Trokuti
- d) Krug

5. Mreža geometrijskog tijela je:

- a) Geometrijsko tijelo
- b) Geometrijski lik

6. Plašt je dio:

- a) Piramide
- b) n-terokuta
- c) Valjka
- d) Stošca

7. Osni presjek valjka je?

- a) Trokut
- b) Pravac
- c) Dužina
- d) Pravokutnik

8. Izračunaj površinu osnovnog presjeka valjka ako je zadan radijus baze $r=2\text{cm}$ i visina $h=3\text{cm}$.

9. Mrežu valjka sačinjavaju:

- a) Dva kruga i pravokutnik
- b) Jedan krug i pravokutnik
- c) Jedan krug i dva pravokutnika

10. Izračunaj obujam valjka radijusa baze 5cm i duljine visine 12cm

11. Primjeri valjka su:

- a) Sunce
- b) Bunar
- c) Stablo
- d) Limenka

12. Što je potrebno da bi stožac bio zadan?

- a) Polumjer baze
- b) Visina stošca
- c) Vrh stošca
- d) Os stošca

13. Kugla je dio prostora omeđen _____

14. Obujam kugle radijusa r dan je formulom _____

15. Plastelin je oblikovan kao valjak visine 10cm i polumjera 3cm. Može li se preoblikovati tako da stane u uglatu posudu dimenzija 12cm x 4cm x 5cm?

G. Anketa očekivanja

ANKETNI UPITNIK

PRED TOBOM JE ANKETNI UPITNIK.

MOLIMO TE DA ISKRENO ODGOVORIŠ NA PITANJA, JER NITKO NEĆE ZNATI DA SI UPRAVO TI ISPUNJAVAO/LA OVAJ UPITNIK!

AKO NEKO PITANJE NIJE JASNO, SVAKAKO ZATRAŽI POMOĆ!

ZAHVALJUJEMO SE NA SURADNJI I STRPLJENJU!

1. Kojeg si spola?
 - a) Muško
 - b) Žensko

2. S kojim si uspjehom završio/la prethodni razred?
 - a) Dovoljan
 - b) Dobar
 - c) Vrlo dobar
 - d) Odličan

3. Koju si ocjenu imao/la iz matematike u prethodnom razredu?
 - a) Dovoljan
 - b) Dobar
 - c) Vrlo dobar
 - d) Odličan

4. Imaš li kod kuće računalo?
 - a) Da
 - b) Ne

5. Koliko prosječno vremena provedeš za računalom kod kuće?
 - a) Uopće ga ne koristim
 - b) Do sat vremena dnevno
 - c) Do tri sata dnevno
 - d) Više od tri sata dnevno
 - e) Samo kada mi odrasli dopuste

6. Zbog čega koristiš računalo
 - a) Zanimljivih računalnih igara
 - b) Mogućnosti korištenja Interneta
 - c) Učenje za školu
 - d) Gledanje filmova
 - e) Slušanje muzike
 - f) Ostalo

7. Kako bi ocijenio svoje znanje (sposobnost) korištenja računala?
 - a) Izvrsnim
 - b) Dobrim
 - c) Lošim

8. Nastavni sadržaji iz matematike (geometrija) su mi:
 - a) Jednostavni i lagani
 - b) Složeni i teški
 - c) Srednje teški

9. Smatraš li da bi učenje matematike (geometrije) uz pomoć posebnog računalnog programa bilo:
 - a) Zanimljivije
 - b) Dosadno
 - c) Ne znam

10. Smatraš li da bi ti učenje matematike (geometrije) uz pomoć posebnog računalnog programa:
 - a) Pomoglo u svladavanju gradiva
 - b) Odmoglo u svladavanju gradiva
 - c) Niti jedno

H. Završna anketa

ANKETNI UPITNIK

PRED TOBOM JE ANKETNI UPITNIK.

MOLIMO TE DA ISKRENO ODGOVORIŠ NA PITANJA, JER NITKO NEĆE ZNATI DA SI UPRAVO TI ISPUNJAVAO/LA OVAJ UPITNIK!

AKO NEKO PITANJE NIJE JASNO, SVAKAKO ZATRAŽI POMOĆ!

ZAHVALJUJEMO SE NA SURADNJI I STRPLJENJU!

1. Kojeg si spola?
 - a. Muško
 - b. Žensko

2. Imaš li kod kuće računalo?
 - a. Da
 - b. Ne

3. Koliko si vremena kod kuće proveo na sustavu xTeX-Sys?
 - a. Uopće nisam pristupio sustavu
 - b. Do sat vremena tjedno
 - c. Do tri sata tjedno
 - d. Više od tri sata tjedno

4. Nastavne sadržaje matematike obrađene uz pomoć xTeX-Sys-a razumio/la sam:
 - a. u potpunosti
 - b. djelomično
 - c. nisam ih uopće razumio/la

5. Učenje matematike uz pomoć sustava xTeX-Sys u odnosu na učenje u klasičnoj nastavi mi je bilo:
 - a. Zanimljivije
 - b. Manje zanimljivo
 - c. Jednako zanimljivo

6. Rezultati u učenju i poznavanju matematike uz pomoć sustava xTeX-Sys u odnosu na rezultate koje sam postizao/la ranije su:
 - a. Bolji
 - b. Lošiji
 - c. Jednaki

7. Želiš li nastaviti učiti matematiku uz pomoć sustava xTeX-Sys?
 - a. želim
 - b. ne želim
 - c. želim povremeno

8. Svojim riječima opiši što misliš o učenju matematike uz pomoć xTeX-Sys-a, te navedi neke dobre i neke loše strane?