

FAKULTET PRIRODOSLOVNO-MATEMATIČKIH ZNANOSTI I
KINEZILOGIJE
SVEUČILIŠTE U SPLITU

Marija Vidov

**META-ANALIZA U SUSTAVIMA
E-UČENJA**

DIPLOMSKI RAD

Mentor: prof.dr.sc. Slavomir Stankov

Split, 2006.

SADRŽAJ

1. UVOD	1
2. META-ANALIZA	3
2.1. SPECIFIKACIJA TEME I PRONALAZENJE ISTRAŽIVANJA	4
2.1.1. <i>Vrste rezultata istraživanja</i>	4
2.1.2. <i>Kriterij za odabir prikladnih istraživanja</i>	6
2.1.3. <i>Pronalaženje izvještaja istraživanja</i>	7
2.2. ODABIR METODE ZA IZRAČUN VELIČINE UČINKA	8
2.2.1. <i>Veličina učinka</i>	8
2.2.2. <i>Metode za izračun veličine učinka</i>	9
2.3. PRIKUPLJANJE VAŽNIH INFORMACIJA IZ POJEDINOG ISTRAŽIVANJA	18
2.3.1. <i>Izrada upitnika</i>	18
2.3.2. <i>Popunjavanje upitnika</i>	19
2.3.3. <i>Stvaranje podatkovnih datoteka</i>	19
2.4. ANALIZA PODATAKA	20
2.4.1. <i>Prilagodbe veličine učinka</i>	20
2.4.2. <i>Analiza skupa veličina učinka i njihove distribucije</i>	21
2.4.2.1. <i>Stvaranje nezavisnog skupa veličina učinka</i>	21
2.4.2.2. <i>Izračun ukupne veličine učinka</i>	21
2.4.2.3. <i>Interval pouzdanosti oko prosječne veličine učinka</i>	22
2.4.2.4. <i>Analiza homogenosti</i>	22
2.5. INTERPRETACIJA I KORIŠTENJE REZULTATA META-ANALIZE	23
2.5.1. <i>Prevođenje ukupne veličine učinka u druge metrike</i>	24
2.5.2. <i>Interpretacija ukupne veličine učinka u terminima preklapanja distribucija</i>	24
2.5.3. <i>Kvaliteta rezultata meta-analize</i>	27
2.5.4. <i>Grafičko prikazivanje rezultata</i>	28
2.6. PREDNOSTI I NEDOSTACI META-ANALIZE	30
3. PRIMJERI VREDNOVANIH SUSTAVA E-UČENJA	32
3.1. PRIMJERI SUSTAVA E-UČENJA	33
3.1.1. <i>Learning by asking - LBA</i>	33
3.1.2. <i>Conceptual Helper</i>	34
3.1.3. <i>SQL-Tutor</i>	35
3.1.4. <i>PsyCAL</i>	35
3.1.5. <i>Computer training of phonological awareness in kindergarten</i>	36
3.1.6. <i>Knowledge Mapping</i>	37
3.1.7. <i>Sherlock</i>	38
3.1.8. <i>Online general college chemistry course</i>	39
3.1.9. <i>Distance learning: Business communications class</i>	40
3.1.10. <i>Computer based instruction in an Engineering Class</i>	40
4. META-ANALIZA REZULTATA VREDNOVANIH SUSTAVA	42
4.1. TEMA META-ANALIZE	42
4.2. ODABIR PRIKLADNIH ISTRAŽIVANJA	43
4.3. ODABIR METODE ZA IZRAČUN VELIČINE UČINKA	43
4.4. IZRAČUN POJEDINAČNIH VELIČINA UČINKA I NJIHOVIH TEŽINA	43
4.4.1. <i>Learning by asking - LBA</i>	44
4.4.2. <i>Conceptual Helper</i>	45
4.4.3. <i>SQL-Tutor</i>	45
4.4.4. <i>PsyCAL</i>	46
4.4.5. <i>Computer training of phonological awareness in kindergarten</i>	47
4.4.6. <i>Knowledge Mapping</i>	48
4.4.7. <i>Sherlock</i>	49
4.4.8. <i>Online general college chemistry course</i>	50
4.4.9. <i>Distance learning: Business communications class</i>	51
4.4.10. <i>Computer based instruction in an Engineering Class</i>	52

4.5.	IZRAČUN UKUPNE VELIČINE UČINKA.....	53
4.5.1.	<i>Prosječna veličina učinka</i>	54
4.5.2.	<i>Interval pouzdanosti</i>	54
4.5.3.	<i>Analiza homogenosti</i>	55
4.5.4.	<i>Analiza rezultata meta-analize</i>	55
5.	ZAKLJUČAK	57
6.	LITERATURA	59

1. UVOD

Sredinom prošlog stoljeća, Eysenck (prema [LIPS2001]) je započeo raspravu o učincima psihoterapije na pacijente u kliničkoj psihologiji. Da bi se riješila započeta rasprava, analitičari su do sredine sedamdesetih godina proveli više stotina istraživanja o učincima psihoterapije na pacijente, i dobili niz različitih rezultata (pozitivnih, neutralnih, negativnih). Dvije strane koje su zastupale različita stajališta u toj raspravi, nisu se nikako mogle složiti oko zaključka, jer je svaka strana kao dokaz koristila dobivena istraživanja koja su govorila u prilog svojoj tezi.

Da bi spomenuta rasprava dobila zaključak, Glass je statistički standardizirao razlike u tih 375 istraživanja, između pacijenata koji su bili podvrgnuti psihoterapiji i onih koji nisu. Zatim je našao njihovu aritmetičku sredinu, i taj postupak nazvao *meta-analiza* (eng. meta-analysis). Objavio je svoje rezultate 1977. godine i donio zaključak da je psihoterapija uistinu učinkovita [SMIT1977].

U proteklih 30-tak godina, koliko je prošlo od tog pionirskog rada, provedeno je više tisuća meta-analiza i napravljena su brojna poboljšanja u njezinoj metodologiji.

Meta-analiza je, dakle, oblik istraživanja u kojem se pregledavaju rezultati nekih drugih istraživanja, i jedan je od načina za sumiranje, integriranje i interpretiranje odabranih radova iz različitih disciplina. Meta-analiza se primjenjuje samo na empirijskim (iskustvenim) rezultatima istraživanja, ne može se primjenjivati na teoretskim radovima. Primjenjuje se samo na ona istraživanja koja daju kvantitativne rezultate, odnosno na ona istraživanja koja koriste kvantitativno mjerenje varijabli [LIPS2001].

Važno je da istraživanja u istoj meta-analizi, osim zajedničke teme, imaju i sličnu strukturu istraživanja. Nije jednostavno rezultate skupa takvih istraživanja međusobno uspoređivati i smisleno analizirati, jer je obično slučaj da istraživanja koriste različite mjerne postupke za ključne varijable. Zato se u meta-analizi rezultati pojedinog istraživanja prikazuju u obliku *veličine učinka* (eng. effect size). Veličina učinka proizvodi statističku standardizaciju rezultata istraživanja, tako da se rezultirajuće numeričke vrijednosti mogu uspoređivati na dosljedan način kroz sve korištene varijable i mjerenja.

Veličina učinka, kao numerička vrijednost, otkriva da li je eksperimentalni faktor koji se istražuje učinkovit i kolika je točno ta učinkovitost u jedinicama standardne devijacije. Dobivena pozitivna vrijednost pokazuje učinkovitost eksperimenta, dok negativna vrijednost veličine učinka pokazuje suprotno, odnosno, u provedenom eksperimentu kontrolna grupa je ostvarila bolje rezultate.

Postoji mnogo različitih metoda za izračun veličine učinka, ovisno o strukturi istraživanja i vrsti dobivenih rezultata. Istraživanja koja daju podatke o korelaciji između dvije varijable, zahtjevaju različit izračun veličine učinka nego istraživanja eksperimentalnog dizajna. Zato je potrebno da sva istraživanja u istoj meta-analizi budu iste strukture, jer je jedino na taj način moguće izračunati veličine učinka svih istraživanja na isti način (koristeći istu metodu).

Rezultat svakog istraživanja se temelji na određenom uzorku ispitanika. Veličina učinka koja se temelji na velikom uzorku je preciznija od one temeljene na malom uzorku. Stoga se u

meta-analizi koristi vrijednost, kojom se „mjeri“ preciznost svake dobivene veličine učinka po veličini uzorka na kojem se temelji, kako bi njezin doprinos statističkoj analizi bio proporcionalan s preciznošću. Hedges (prema [LIPS2001]) je dokazao da je najbolja mjera temeljena na standardnoj pogrešci veličine učinka. Veća standardna greška odgovara manje preciznoj vrijednosti veličine učinka, zato se to mjerenje preciznosti računa kao inverzna vrijednost kvadrata standardne pogreške i naziva se *inverzna težina varijance* (eng. inverse variance weight) u meta-analizi.

Proces učenja i poučavanja u posljednjih se dvadesetak godina nastoji unaprijediti na različite načine zahvaljujući istraživanju koje je proveo Bloom [BLOO1984]. On je svojim istraživanjem pokazao da je veličina učinka individualnog poučavanja 2 standardne devijacije u odnosu na tradicionalno. Stoga, da bi se postigla veća učinkovitost u tom procesu učenja i poučavanja, nastoji ga se što više prilagoditi svakom pojedinom učeniku. To se nastoji postići na takav način koji ne zahtijeva nužno prisutnost učitelja za svakog učenika posebice, već primjenom informacijske i komunikacijske tehnologije. Svakodnevno se kreiraju i oblikuju sustavi koji integriraju različite pogodnosti koje navedena tehnologija pruža, od *web-temeljenog poučavanja* (eng. web-based learning), *računalno-temeljenog učenja* (eng. computer based learning), *digitalnih medija* (eng. digital collaboration) do *virtualnih učionica* (eng. virtual classrooms). Svi navedeni oblici poučavanja čine *e-učenje* (eng. e-learning) [LEAR2006], koje u novije vrijeme uključuje i primjenu mobilne tehnologije (mobilnih uređaja, PDA uređaja te MP3 player-a) [WIKI2006].

Sustavi e-učenja zamišljeni su kao osobni učitelji za svakog učenika, a poučavanje podupire spomenuta tehnologija. Iako takvi sustavi još nisu dostigli učinkovitost individualnog tutorskog poučavanja, stručnjaci s tog područja se slažu da suvremeno doba obećava značajno povećanje njihove učinkovitosti. Informacijska i komunikacijska tehnologija je postala sve prisutna i lako dostupna, svakodnevno se sve više škola oprema kvalitetnom računalnom opremom, a internetu se može pristupiti s bilo kojeg mjesta čime su zadovoljene pretpostavke za uspješno provođenje e-učenja. Isto se tako i programska podrška unaprijeđuje velikom brzinom, pa se dizajniraju sve bolji i učinkovitiji sustavi koji sve kvalitetnije podupiru taj proces.

Osim toga, e-učenje je od iznimne važnosti za proces poučavanja zbog potrebe permanentnog obrazovanja. Moderno doba ne dopušta završetak obrazovanja izlaskom iz škole ili fakulteta. Napredak u svim granama znanosti i u tehnologiji, zahtjeva od pojedinca praćenje napretka, novih spoznaja i svakodnevno usavršavanje. Zbog činjenice da e-učenje omogućava pristup izvorima za učenje u bilo koje vrijeme i bilo gdje, učenje je omogućeno svima.

U drugom poglavlju je opisan način provedbe meta-analize. Pošto meta-analiza predstavlja istraživački proces, opisane su pojedine faze u tom postupku: određivanje teme, odabir i pronalazak odgovarajućih istraživanja koja će se koristiti, određivanje metode pomoću koje će se izračunati veličine učinka danih istraživanja, zatim izdvajanje najvažnijih informacija i njihova analiza te izračun ukupne veličine učinka i njezina interpretacija. Na kraju je dan kratak pregled prednosti i nedostataka meta-analize.

U trećem poglavlju su opisani primjeri vrednovanih sustava e-učenja. Ukratko je opisan svaki sustav i način na koji se provodilo poučavanje. Dan je i kratak opis eksperimenta koji je proveden za vrednovanje pojedinog sustava, te rezultati jednog ili više završnih testova.

U četvrtom poglavlju provedena je meta-analiza rezultata opisanih vrednovanih sustava e-učenja. Tema te meta-analize je utvrđivanje učinkovitosti takvih sustava u procesu učenja i poučavanja u odnosu na tradicionalni način poučavanja. Nakon što je određena metoda za izračun veličine učinka, izračunate su veličine učinka svakog sustava posebno. Na kraju je provedena analiza tih rezultata i dobivena je ukupna veličina učinka.

2. META-ANALIZA

Meta-analiza je nastala kao odgovor na pitanje koje analitičari do tad nisu znali odgovoriti. Iz kolekcije istraživanja, koja su davala nizove pozitivnih, negativnih i neutralnih rezultata analize učinkovitosti nekog eksperimentalnog faktora (metode, sustava, tretmana, i sl.), nisu znali izvući zaključak o tome da li je on u konačnici učinkovit. I ako jest, koliki je taj njegov učinak. Meta-analiza kao svoj rezultat prikazuje upravo to, učinkovitost eksperimentalnog faktora kojeg istražuje.

Meta-analiza je prema [LIPS2001] oblik istraživanja u kojem se pregledavaju rezultati nekih drugih istraživanja, i jedan je od načina za sumiranje, integriranje i interpretiranje odabranih radova iz različitih disciplina. Meta-analiza se primjenjuje samo na empirijskim (iskustvenim) rezultatima istraživanja, ne može se primjenjivati na teoretskim radovima. Primjenjuje se samo na ona istraživanja koja daju kvantitativne rezultate, odnosno na ona istraživanja koja koriste kvantitativno mjerenje varijabli.

Ako je dan skup kvantitativnih rezultata istraživanja koji se bave istom temom i jednake su strukture, ostaje veliki problem kako te rezultate uvrstiti u bazu podataka koju možemo smisleno analizirati. Različita istraživanja obično ne koriste iste mjerne procedure za ključne varijable. Postavlja se pitanje kako se njihovi rezultati mogu kombinirati i uspoređivati. Odgovor leži u glavnoj karakteristici meta-analize. Meta-analiza prezentira rezultate svakog istraživanja u obliku veličine učinka. Veličina učinka je pojam koji otkriva kvantitativnu informaciju iz svakog rezultata istraživanja.

Veličina učinka uzima u obzir veličinu i smjer svake bitne statističke veze u istraživanju. Zato predstavlja značajni doprinos statističkoj analizi rezultata, jer osim što pokazuje da li je neka metoda ili tretman učinkovit, pokazuje i kolika je točno ta učinkovitost u jedinicama standardne devijacije. To, konkretno, znači da je kod eksperimentalnog istraživanja dobivena veličina učinka kao pozitivna numerička vrijednost, pokazatelj utjecaja eksperimentalnog faktora, tj. koliko standardnih devijacija je uspješnija eksperimentalna grupa. Dok je negativna znak neuspješnosti, tj. koliko je uspješnija kontrolna grupa. Dakle, veličina učinka meta-analize tvori varijablu koja je osjetljiva na rezultate različite jačine kroz sva istraživanja.

Načini izračunavanja veličine učinka, koji se koriste za obradu različitih kvantitativnih rezultata istraživanja u meta-analizi, temelje se na standardizaciji. Izračuni veličine učinka proizvode statističku standardizaciju rezultata istraživanja, tako da se rezultirajuće numeričke vrijednosti mogu interpretirati na dosljedan način kroz sve korištene varijable i mjerenja. Standardizacija u ovom kontekstu znači isto kao i standardizacija ispita znanja. Dakle, da bi usporedili uspjeh učenika npr. iz matematike, dobiveni rezultati se mogu pretvoriti u percentile ili u z-vrijednosti i tako smisleno uspoređivati. Na taj se isti način koristi i veličina učinka. Da bi usporedili rezultate više različitih istraživanja (iste teme i strukture), najprije izračunamo veličine učinka za svako istraživanje posebno i dobijemo rezultate izražene u jedinicama standardne devijacije. Tako se dobiveni rezultati mogu smisleno uspoređivati, ali i obrađivati u smislu pronalaženja aritmetičke sredine, varijance i sličnih statističkih izračuna.

Ključ meta-analize je, dakle, definiranje metode izračuna veličine učinka, koja omogućuje prikaz kvantitativnih rezultata skupa određenih istraživanja u standardiziranoj formi i smislenu numeričku usporedbu i analizu kroz sva istraživanja.

2.1. SPECIFIKACIJA TEME I PRONALAZENJE ISTRAŽIVANJA

Kao svako istraživanje, meta-analiza bi također trebala započeti sa opreznim izvješćem o temi koja se istražuje. Tema se obično definira na način da se oblikuju pitanja na koja je potrebno pronaći odgovore. Tako definirana, tema određuje odabir važnih istraživanja, koji daju odgovore na ta pitanja. Određuje i odabir najvažnijih informacija iz tih istraživanja, i na samom kraju usmjerava analizu dobivenih podataka.

2.1.1. VRSTE REZULTATA ISTRAŽIVANJA

Kvantitativni rezultati istraživanja mogu biti prikazani u mnogo različitih oblika. U meta-analizi se uglavnom koriste razlike između aritmetičkih sredina grupa, korelacija između varijabli, udio itd. Rezultati vezani za temu ili pitanje koje meta-analiza proučava obrađuju se koristeći isti izračun veličine učinka. Znači, da bi bili međusobno usporedivi i tako smisleni za analizu, isti izračun veličine učinka mora biti korišten za obradu svih rezultata u danoj meta-analizi. Zato analitičar mora pronaći oblike rezultata istraživanja koji su vezani za temu meta-analize i pobrinuti se da se ti rezultati mogu obrađivati zajedničkim izračunom veličine učinka. Ako to nije slučaj, rezultati koji su važni za meta-analizu moraju biti posloženi tako, da se oni rezultati koji zahtjevaju različit izračun veličine učinka, mogu meta-analizirati posebno.

Neki oblici rezultata se mogu obrađivati već poznatim izračunom veličine učinka koji se koristi u meta-analizi npr. standardna razlika aritmetičkih sredina, koeficijent korelacije, omjer šansi itd., dok drugi zahtjevaju specijalizirani izračun veličine učinka, koji se ponekad mora razviti "na licu mjesta" i nemoguće ga je unaprijed odrediti.

Slijede neki oblici rezultata istraživanja iz kojih se može izračunati veličina učinka na načine koji su analitičari već razvili u meta-analizi.

OPIS CENTRALNE TENDENCIJE (ENG. CENTRAL TENDENCY DESCRIPTION)

Rezultati ovog tipa dobiveni su od jednog uzorka ispitanika i oni se obrađuju nekom od mjera centralne tendencije kao što su aritmetička sredina, medijan, mod ili udio. Tako npr. istraživanje može davati za rezultate, postignuća učenika na nekom testu iz određenog predmeta. Da bi se provela meta-analiza, potrebno je da događaj ili varijabla koju se mjeri bude analizirana istom metodom u svim istraživanjima. To u ovom konkretnom primjeru znači testiranje standardiziranim testom kroz sva istraživanja, ili standardizacija dobivenih

rezultata (z-vrijednosti). Međutim, valja napomenuti da se rezultati ovog tipa rijetko meta-analiziraju.

KONTRASTI INICIJALNOG I ZAVRŠNOG MJERENJA (ENG. PRE-POST CONTRASTS)

Rezultati ovog tipa dobiveni su također od jednog uzorka ispitanika tako da se uspoređuje centralna tendencija (npr. aritmetička sredina ili udio) nekog događaja ili varijable mjerene u neko vrijeme, sa centralnom tendencijom istog događaja (varijable) mjerene nakon određenog vremena. Ovo se često provodi za provjeru promjene. Uzima se direktna razlika vrijednosti dviju centralnih tendencija, ili centralna tendencija razlike vrijednosti inicijalnog i završnog mjerenja za svakog ispitanika posebno. Veličinu učinka za inicijalno-završno mjerenje prezentirao je Becker (prema [LIPS2001]) kao standardnu razliku aritmetičkih sredina.

KONTRASTI U GRUPI (ENG. GROUP CONTRASTS)

Ovaj tip rezultata istraživanja uključuje jednu ili više varijabli mjenjenih na dvije ili više grupa ispitanika koji se tada uspoređuju, najčešće metodom centralne tendencije (aritmetička sredina ili udio). Dva su oblika kontrasta u grupi u meta-analizi:

- a) Istraživanje eksperimentalnih ili kliničkih pokusa (eng. Experimental or clinical trials research)

U ovom slučaju grupe ispitanika se uspoređuju i tako pokazuju stanje eksperimenta npr. grupa koja je podvrgnuta nekom tretmanu – eksperimentalna grupa s kontrolnom grupom. Razlika između tih grupa, vezano za varijablu koja se mjeri, interpretira se kao učinak tog tretmana.

- b) Istraživanje različitosti grupa (eng. Group differences research)

U ovom slučaju grupe se razlikuju po karakteristikama koje nisu uzrok djelovanja nekog tretmana ili eksperimenta, već su prirodnog karaktera, npr. istraživanje različitosti spolova, ili nekih demografskih grupa.

Kada se govori o usporedbi tri ili više grupa ispitanika, te se grupe moraju uspoređivati dvije po dvije, dakle, ne postoje neke osnovne tehnike uspoređivanja tri ili više grupa istovremeno.

POVEZANOST IZMEĐU VARIJABLI (ENG. ASSOCIATION BETWEEN VARIABLES)

Ovaj tip rezultata obrađuje dvije varijable istog ispitanika da bi se utvrdila eventualna povezanost između njih, npr. socijalno-ekonomski status obitelji nekog učenika i njegov uspjeh u savladavanju nastavog sadržaja. Rezultati se mogu prikazati u obliku koeficijenta korelacije ili nekog indeksa povezanosti.

Dok se iz prethodno navedenih oblika rezultata istraživanja u meta-analizi može izračunati veličina učinka, i to pomoću već razvijenih metoda, postoje i drugi oblici istraživanja koji su složeniji za meta-analizu. To su u prvom redu rezultati dobiveni od više-varijablenih analiza npr. multipla regresija i sl. Meta-analitičari još nisu pronašli odgovarajući način izračuna veličine učinka koji bi obrađivao tako dobivene rezultate, jer njihova raznovrsnost i složenost otežava taj proces.

2.1.2. KRITERIJ ZA ODABIR PRIKLADNIH ISTRAŽIVANJA

Kad je određena tema meta-analize i kad je jasno koji tip rezultata će se koristiti vezano za tu temu, treba prijeći na pronalaženje istraživanja koja su prikladna za tu meta-analizu. Karakteristika dobre meta-analize je ta, da istraživač jasno odredi populaciju istraživanja čije će rezultate obrađivati i sumirati. Tako se lakše provodi selekcija ili odbacivanje istraživanja za tu meta-analizu. Zato se preporuča napraviti detaljnu pisanu specifikaciju kriterija koje istraživanje mora zadovoljiti kako bi se njegovi rezultati uključili u meta-analizu. Ti kriteriji prikladnosti uglavnom ovise o temi meta-analize, međutim postoje određene općenite kategorije koje se mogu uzeti u obzir za većinu aplikacija.

a) *Istaknuti detalj*

Koji su prepoznatljivi detalji istraživanja koji ga čine primjerenim za temu meta-analize? Ako meta-analiza ispituje učinak nekog tretmana, tada bi npr. važan detalj bio priroda tog tretmana. Tako kriterij prikladnosti može specificirati karakteristike koje neki tretman mora imati da bude važan, da pruži informacije koje su potrebne. Npr. ako se meta-analiza bavi uspoređivanjem grupa (razlike spolova), kriterij bi morao specificirati vrstu grupa i usporedbi koje se istražuju.

b) *Ispitanici istraživanja*

Koje su primjerene karakteristike ispitanika (uzorka) koji daju podatke u istraživanju, i čine ga prikladnim za meta-analizu? Istraživanja koja su bitna za neku meta-analizu, ponekad mogu uključivati samo one, koji za ispitanike imaju maloljetnike, ili su važna samo ona istraživanja sa ispitanicima koji imaju neke druge demografske karakteristike. Ponekad su bitni ispitanici neke određene kulturne ili religijske pripadnosti npr. samo ispitanici engleskog govornog područja, samo rimokatolici itd.

c) *Ključne varijable*

Koje su to ključne varijable nekog istraživanja koje će biti prikladno za meta-analizu? Kod istraživanja koji se bavi proučavanjem učinka nekog tretmana, to bi bile specifične izlazne varijable koje odgovaraju toj temi. Kod istraživanja koje uspoređuje grupe, to bi bile osnovne varijable po kojima se te grupe razlikuju itd. Zbog činjenice da okosnicu meta-analize čini izračun veličine učinka da bi se prezentirali rezultati istraživanja, važno je naglasiti da se statistička informacija nekog istraživanja prezentira tako da omogućava izračun odgovarajuće veličine učinka ili neke druge informacije koje su vezane uz ključne varijable.

d) *Dizajni istraživanja*

Koji istraživački dizajni i metodičke karakteristike kvalificiraju istraživanje da bude prikladno za meta-analizu, a koji diskvalificiraju? Primarno je, naravno, da se specificira koji su oblici rezultata istraživanja bitni za meta-analizu. Kod istraživanja koje se bavi proučavanjem učinaka nekog tretmana, to bi značilo korištenje eksperimentalnog dizajna. Analitičar odlučuje koje će oblike istraživanja uključiti u svoju meta-analizu npr. samo one koji su s placebo kontrolom, ili samo one koji koriste pouzdane i već provjerene mjerne instrumente i sl.

Kod ovog kriterija je važno napomenuti, ako se uzima previše restriktivno, tako da se meta-analiza temelji na najpovoljnijim istraživanjima, to može ograničiti broj i okvir takvih prikladnih istraživanja, pa se može dogoditi da ih ima nedovoljno. S druge

strane ako ih se bira opušteno, bez puno restrikcija da bi bolje iskoristili provedena istraživanja, može dovesti do grešaka i zabluda u meta-analizi.

e) *Kulturni i jezični okvir*

Koji je kulturni i jezični okvir istraživanja prikladnog za neku meta-analizu? Istraživanja se provode u mnogo zemalja i na mnogo jezika. Analitičar treba odrediti koji su važni za temu te meta-analize. Nije neuobičajeno da meta-analitičari engleskog govornog područja isključe istraživanja koja nisu pisana na tom jeziku, upravo zbog praktičnih poteškoća prevođenja.

f) *Vremenski okvir*

Koliki je vremenski period u kojem se moraju provesti istraživanja, ako on postoji? Ponekad su autoru zanimljiva samo novija istraživanja, npr. koja su provedena od dana kad je neka metoda ili neko sredstvo postalo dostupno.

g) *Tip objavljivanja (publikacije)*

U kojem obliku su istraživački izvještaji prikladni za meta-analizu? Oblici izvještaja mogu biti raznoliki; od objavljenih novinskih članaka, knjiga, disertacija, tehničkih izvještaja do neobjavljenih rukopisa, konferencijskih prezentacija i tome slično.

Ako meta-analitičar odluči koje će oblike koristiti a koje isključiti, treba ih unaprijed specificirati i obrazložiti odabir ili eventualno isključivanje. Tako pojedini analitičari koriste samo formalno objavljene materijale jer su najlakše dostupni. Taj se odabir ponekad opravdava time da ako su objavljeni, znači da su kvalitetniji, što nije uvijek točno. Poznato je i to da su učinci koji su prikazani u objavljenim materijalima generalno veći nego učinci u neobjavljenim istraživanjima. Dakle, isključivanjem neobjavljenih istraživanja, meta-analitičar može tako utjecati na završni rezultat meta-analize i dobiti veći učinak od stvarnog. Zato se ovaj kriterij prikladnosti koji odabire istraživanja po obliku objavljivanja, mora vrlo oprezno definirati i odabir dobro obrazložiti.

Pojedinosti svake od ovih kategorija su različite za svaku meta-analizu, ovise o fokusu i namjeni sinteze, i mogu biti relativno restriktivni ili inkluzivni. Traženje prikladnih istraživanja prema kriteriju prikladnosti predstavlja iterativni postupak, jer se u većini slučajeva dogodi da se kriteriji određeni na početku meta-analize, tijekom obrade unaprijede i prodube pa je potrebno ponovno tražiti istraživanja koja odgovaraju tim "promjenjenim" kriterijima.

2.1.3. PRONALAZENJE IZVJEŠTAJA ISTRAŽIVANJA

Prvi korak u potrazi za odgovarajućim istraživanjima je razvoj opširnog sustava za bilježenje napretka u pronalazenju potencijalnih izvještaja. Taj sustav bi trebao biti bibliografski, u obliku baze podataka sa poljima, koja sadrže informacije o stanju potrage za svakog kandidata posebno. Neki analitičari u tu bazu unose pregled izvještaja istraživanja, ako je dostupan, ili pohrane kopiju u neki povezani dokument. Dodatna polja mogu sadržavati korisne dostupne informacije npr. tip izvještaja (novinski članak, disertacija). Trebalo bi postojati i polje koje će ukazivati na datum kad je izvještaj bio pronađen ili ako je potraga obustavljena i razlog. Pažljivo održavanje ovakvog sustava omogućava autoru provedbu detaljne potrage, ali isto

tako daje uvid u potencijalna istraživanja koja će se koristiti, i ona koja su odbačena zbog nekog razloga.

Nakon toga analitičar prelazi na sam proces traženja koji se sastoji od dva dijela. Prvo, traženje bibliografskih referenci na potencijalno korisna istraživanja, a zatim slijedi stvarna nabava kopija tih istraživanja. Pri pronalaženju referenci, koriste se različiti izvori: pregledi članaka, pregledi referenci u istraživanjima, bibliografske baze podataka, kontaktiranje eksperata i autora u zadanom području itd.

Ako se sadržaji istraživanja nalaze u novinskim člancima i knjigama, njih se najlakše pronalazi u knjižnicama. Doktorske disertacije se osim u knjižnici mogu dobiti kontaktiranjem autora ili fakulteta na kojem su nastale. Zato analitičar treba biti spreman pisati pisma ili na drugačiji način kontaktirati bilo autore, institucije ili agencije koje su sudjelovale u istraživanju ili su sponzorirale određena istraživanja.

Jednom kad je lista potencijalnih istraživanja formirana i kad su pronađene kopije sadržaja tih istraživanja, tada se cijeli sadržaj pregleda i još jednom utvrdi da li je to istraživanje prikladno za zadanu meta-analizu.

Na kraju se preporučuje zadržavanje kopija svih odgovarajućih istraživanja u nekoj arhivi koja će biti dio te meta-analize, jer se često tijekom same analize jave pitanja i problemi čije rješavanje zahtjeva provjeru originalnih materijala nekog istraživanja. Ponekad se dogodi da meta-analiza izazove kritike drugih istraživača i analitičara koji tada najčešće pregledavaju originalne materijale, što je najlakše izvesti ako se ti originalni izvještaji drže "pri ruci".

2.2. ODABIR METODE ZA IZRAČUN VELIČINE UČINKA

Najvažniji korak u meta-analizi je usporedba odnosno "mjerenje" odabranih rezultata istraživanja nekom vrstom numeričke skale vrijednosti, tako da se krajnje dobiveni rezultati mogu međusobno smisleno analizirati. Kao što je već spomenuto, vrijednost ili izraz koji se koristi za predstavljanje rezultata istraživanja u meta-analizi, nazivamo veličinom učinka. Jednom kad se utvrdi skup kvantitativnih rezultata istraživanja koji predstavljaju određenu vezu, ti rezultati se pretvaraju u vrijednosti za izračun veličine učinka i prosljeđuju dalje na analizu.

Koji će oblik izračuna veličine učinka biti najprikladniji, ovisi o prirodi rezultata, statističkom obliku u kojem su izraženi, te o pitanjima na koja odgovara ta meta-analiza.

2.2.1. VELIČINA UČINKA

Veličina učinka u meta-analizi sadrži informaciju o smjeru i veličini kvantitativnih rezultata istraživanja. To je numerička vrijednost koja pokazuje koliki je utjecaj eksperimentalnog faktora kojeg istraživanje proučava, tj. određuje razliku eksperimentalne i kontrolne grupe u jedinicama standardne devijacije.

U meta-analizi jedan rezultat istraživanja je statistički prikaz jedne empirijske veze neke važne varijable koja je mjerena na nekom uzorku. Npr. razlika između aritmetičke sredine izlazne varijable za dva slučaja nekog eksperimenta čini rezultat istraživanja. Jedno istraživanje tako može davati više rezultata, bilo da se radi o različitim varijablama, različitim uzorcima (muški, ženske), različitom vremenu mjerenja (odmah nakon nekog tretmana ili nešto kasnije) itd.

Izračun veličine učinka mora biti isti kroz sva istraživanja da bi se omogućila smisljena analiza, i taj izračun mora odgovarati prirodi i statističkoj formi rezultata. Veličina učinka se izračunava iz kvantitativnih informacija dobivenih iz istraživačkih izvještaja.

Osim veličine učinka, postoji još jedan pojam koji se mora uzeti u obzir u meta-analizi. Svaki rezultat se temelji na određenom uzorku ispitanika, iako oni u nekim istraživanjima (npr. eksperimentu) mogu biti podijeljeni u nekoliko podgrupa koje se međusobno uspoređuju. Broj ispitanika u uzorku gotovo uvijek varira od istraživanja do istraživanja, tako da su različite vrijednosti veličine učinka temeljene na različitim veličinama uzorka. Sa statističkog stajališta, veličina učinka koja se temelji na većim uzorcima ispitanika, mnogo preciznije reprezentira odgovarajuću populaciju nego ona temeljena na manjim uzorcima. To znači da ima manju *grešku vezanu uz uzorak* (eng. sampling error).

Ova karakteristika veličine učinka komplicira statističku analizu, čak i najjednostavniju kao što je izračun aritmetičke sredine veličine učinka. Problem je u tome, što je svaka dobivena veličina učinka različita po vrijednosti zbog pouzdanosti informacija iz kojih je izvedena.

Npr. veličina učinka temeljena na uzorku od 500 ispitanika preciznija je od one temeljene na 5 ispitanika. Ako jednostavno izračunamo njihovu aritmetičku sredinu, veličina učinka temeljena na manjem uzorku isto doprinosi rezultatu kao i ona temeljena na većem iako ima veću grešku u procjeni. Tako kombinirana procjena ispadne lošija, nego da smo samo uzeli onu vrijednost koja se temelji na većem uzorku.

Ovaj problem se u meta-analizi rješava "mjerenjem" vrijednosti veličine učinka izrazom koji će određivati njezinu preciznost. To bi značilo mjerenje točnosti svake veličine učinka po njezinoj veličini uzorka. Hedges (prema [LIPS2001]) je dokazao da je najbolja mjera temeljena na standardnoj pogrešci veličine učinka. Standardna pogreška je standardna devijacija distribucije uzorka. Veća standardna greška odgovara manje preciznoj vrijednosti veličine učinka, zato se "mjerenje" preciznosti računa kao inverzna vrijednost kvadrata standardne pogreške i to se naziva *inverzna težina varijance* (eng. inverse variance weight) u meta-analizi.

2.2.2. METODE ZA IZRAČUN VELIČINE UČINKA

Prethodno je bilo riječi o rezultatu istraživanja, kao statističkom prikazu jedne empirijske veze neke važne varijable koja je mjerena na jednom uzorku. Sada će za svaki rezultat istraživanja biti prikazan odgovarajući izračun veličine učinka. Općenito se razlikuju veze koje uključuju jednu varijablu, dvije varijable i više varijabli i u svaku od ovih kategorija spada nekoliko različitih vrsta rezultata. Za svaki tip rezultata slijedi opis prikladnog načina izračunavanja veličine učinka.

VEZE KOJE UKLJUČUJU JEDNU VARIJABLU (ENG. ONE-VARIABLE RELATIONSHIPS)

Izraz "veze" koja uključuje samo jednu varijablu je neobičan, ali se tu podrazumijeva niz promatranja kategorija i vrijednosti jedne varijable. Takve su vrijednosti tipa mjera centralne tendencije, znači aritmetička sredina, medijan, mod, ali i suma, udio, itd.

a) Udio (eng. Proportions) - p

Rezultat istraživanja koji predstavlja udio uzorka sa nekom karakteristikom može direktno biti izračun veličine učinka sa vrijednostima između 0.00 i 1.00, a vrši se na sljedeći način:

$$\begin{aligned}
 ES_p &= p = \frac{k}{n}, \\
 SE_p &= \sqrt{\frac{p(1-p)}{n}}, \\
 w_p &= \frac{1}{SE_p^2} = \frac{n}{p(1-p)},
 \end{aligned}
 \tag{2.1}$$

gdje su ES , SE i w kratice za veličinu učinka (eng. Effect Size), standardnu pogrešku (eng. Standard Error) i težinu (eng. weight) redom, i takve će se koristiti u daljnjem tekstu, k je broj ispitanika s nekom karakteristikom, a n je ukupan broj ispitanika.

b) Aritmetička sredina (eng. Arithmetic Means) - m

Uobičajena statistička metoda za obradu rezultata istraživanja jedne varijable je izračun aritmetičke sredine kao veličine učinka, kako slijedi:

$$\begin{aligned}
 ES_m &= \bar{X} = \frac{\sum x_i}{n}, \\
 SE_m &= \frac{s}{\sqrt{n}}, \\
 w_m &= \frac{1}{SE_m^2} = \frac{n}{s^2},
 \end{aligned}
 \tag{2.2}$$

gdje je x_i individualni iznos i -tog ispitanika ($i = 1$ do n), s je standardna devijacija od x .

VEZE KOJE UKLJUČUJU DVIJE VARIJABLE (ENG. TWO-VARIABLE RELATIONSHIPS)

Veze dvije varijable su najprisutnije u istraživanjima i najčešće su meta-analizirane. One uključuju metode kao što su kontrasti inicijalnog i završnog mjerenja, kontraste u grupi te povezanost dvije varijable.

a) *Kontrasti inicijalnog i završnog mjerenja*

Spomenuto je, da se tu podrazumjeva uspoređivanje centralne tendencije (aritmetičke sredine) varijable mjerene u jednom trenutku, sa centralnom tendencijom te iste varijable mjerene na isti način i na istom uzorku poslije određenog vremena. Obično se provodi da bi se usporedila razlika npr. djelovanje nekog tretmana. Za obradu rezultata ovog tipa, koriste se dva izračuna veličine učinka, nestandardno i standardno srednje povećanje.

- **Nestandardno srednje povećanje (eng. The Unstandardized Mean Gain) - ug**

Ovaj tip izračuna se primjenjuje na situacije kad su svi rezultati inicijalno-završnih mjerenja dobiveni istim mjernim postupkom na nekoj varijabli. Dakle, kad se isti mjerni instrument (npr. neki standardizirani test za provjeru znanja) koristi na inicijalnom kao i završnom mjerenju za sve uzorke ispitanika, tada su dobivene vrijednosti usporedive, pa se može direktno izračunati veličina učinka na sljedeći način:

$$ES_{ug} = \bar{X}_{T2} - \bar{X}_{T1},$$

$$SE_{ug} = \sqrt{\frac{2s_p^2(1-r)}{n}} = \sqrt{\frac{s_g^2}{n}}, \quad (2.3)$$

$$w_{ug} = \frac{1}{SE_{ug}^2} = \frac{n}{2s_p^2(1-r)},$$

gdje su \bar{X}_{T1} i \bar{X}_{T2} aritmetičke sredine u vremenu 1 i 2, s_p^2 je zajednička varijanca iznosa u vremenu 1 i 2, tj. $s_p^2 = (s_{T1}^2 + s_{T2}^2) / 2$, s_g^2 je varijanca povećanja, r je korelacija između iznosa u vremenu 1 i 2.

- **Standardno srednje povećanje (eng. The Standardized Mean Gain) - sg**

Kontrasti inicialnog i završnog mjerenja, po definiciji, koriste iste mjerne procedure za varijable u oba vremena mjerenja. Međutim, različita istraživanja mogu u tom postupku koristiti različite metode mjerenja, ovisno o tome koje istraživač odabere kao najprikladnije. Tada se dobiveni numerički rezultati ne smiju uspoređivati, niti ih se tako smije uvrstiti u istu meta-analizu. Potrebna je standardizacija inicijalno-završnih mjerenja na taj način da se dobivene vrijednosti mogu smisljeno uspoređivati kroz sve uzorke i istraživanja. Izračun veličine učinka za takvu vrstu mjerenja razvio je Becker (prema [LIPS2001]), a definiira se:

$$ES_{sg} = \frac{\bar{X}_{T2} - \bar{X}_{T1}}{s_p},$$

$$SE_{sg} = \sqrt{\frac{2(1-r)}{n} + \frac{ES_{sg}^2}{2n}}, \quad (2.4)$$

$$w_{sg} = \frac{1}{SE_{sg}^2} = \frac{2n}{4(1-r) + ES_{sg}^2},$$

gdje je s_p zajednička standardna devijacija iznosa u vremenu 1 i 2 tj.

$$s_p = \sqrt{(s_{T1}^2 + s_{T2}^2)/2} . \quad (2.5)$$

b) Kontrasti u grupi

Rezultati istraživanja u ovom obliku uključuju varijablu, koju se mjeri u dvije ili više grupa ispitanika i uspoređuje kroz grupe. Deskriptivna statistika koja se koristi za obradu ovakvih rezultata je aritmetička sredina, standardna devijacija i veličina uzorka svake grupe i za svaku varijablu. Grupe se mogu uspoređivati i pomoću udjela, tako da se pokaže u kojem udjelu neka grupa posjeduje ili ne posjeduje određenu karakteristiku. Razlike između grupa isto tako se mogu obrađivati pomoću t-testa, analize varijance i sl.

Gotovo sva eksperimentalna i kvazi eksperimentalna istraživanja daju rezultate u ovom obliku, npr. usporedba eksperimentalne i kontrolne grupe u jednoj ili više zavisnih varijabli. Isto je tako česta usporedba grupa koje nisu dio nekog eksperimenta, npr. različite demografske grupe kao rezultate daju razlike vezano za stajalište, mišljenje, ponašanje itd. Prema tome, postoje različiti izračuni veličine učinka za rezultate ovog tipa istraživanja, od kojih su najvažniji sljedeći: nestandardna i standardna razlika aritmetičkih sredina, razlika udjela i omjer šansi.

- **Nestandardna razlika aritmetičkih sredina (eng. The Unstandardized Mean Difference) - um**

Kad se u svim istraživanjima koristi ista mjerna procedura i ista numerička skala za varijablu koju ispitujeemo, i varijabla je kontinuirana, veličina učinka se može dobiti direktno razlikom aritmetičkih sredina grupa:

$$ES_{um} = \bar{X}_{G1} - \bar{X}_{G2},$$

$$SE_{um} = s_p \sqrt{\frac{1}{n_{G1}} + \frac{1}{n_{G2}}}, \quad (2.6)$$

$$w_{um} = \frac{1}{SE_{um}^2} = \frac{n_{G1}n_{G2}}{s_p^2(n_{G1} + n_{G2})},$$

gdje su \bar{X}_{G1} i \bar{X}_{G2} aritmetičke sredine grupa 1 i 2, n_{G1} i n_{G2} su brojevi ispitanika u grupama 1 i 2, s_p je zajednička standardna devijacija tj.

$$s_p = \sqrt{\frac{(n_{G1} - 1)s_{G1}^2 + (n_{G2} - 1)s_{G2}^2}{(n_{G1} - 1) + (n_{G2} - 1)}}, \quad (2.7)$$

gdje su s_{G1} i s_{G2} standardne devijacije za grupe 1 i 2.

- **Standardna razlika aritmetičkih sredina (eng. The Standardized Mean Difference) - sm**

Ovaj tip izračuna veličine učinka se koristi kad se uspoređuju dvije grupe na nekoj zavisnoj varijabli koja nije mjerena istom mjernom procedurom u svim istraživanjima. Najčešće se koristi za usporedbu rezultata eksperimentalne i

kontrolne grupe, kod proučavanja učinka nekog tretmana. Primjenjuje se i za usporedbu razlika grupa koje nisu dio nekog eksperimenta, npr. muškarci i žene i razlike između njih vezano uz spol. Izračunava se na sljedeći način:

$$ES_{sm} = \frac{\bar{X}_{G1} - \bar{X}_{G2}}{s_p}, \quad (2.8)$$

gdje je s_p zajednička standardna devijacija kao u prijašnjem slučaju.

Međutim, ovaj izračun pokazuje lažno veći učinak kad se temelji na malom uzorku ispitanika, pogotovo ispod 20. To je dokazao Hedges (prema [LIPS2001]), koji je pronašao jednostavan ispravak ove greške, tako da se veličina učinka računa prema sljedećoj formuli:

$$ES'_{sm} = \left[1 - \frac{3}{4N - 9} \right] ES_{sm},$$

$$SE_{sm} = \sqrt{\frac{n_{G1} + n_{G2}}{n_{G1}n_{G2}} + \frac{(ES'_{sm})^2}{2(n_{G1} + n_{G2})}}, \quad (2.9)$$

$$w_{sm} = \frac{1}{SE_{sm}^2} = \frac{2n_{G1}n_{G2}(n_{G1} + n_{G2})}{2(n_{G1} + n_{G2})^2 + n_{G1}n_{G2}(ES'_{sm})^2},$$

gdje je N ukupan broj ispitanika u obje grupe zajedno tj. $(n_{G1} + n_{G2})$.

Kod usporedbe eksperimentalne i kontrolne grupe, po dogovoru se uzima pozitivna vrijednost veličine učinka kada je eksperimentalna grupa "bolja" od kontrolne, a negativna vrijednost znači da je eksperimentalna grupa "lošija" od kontrolne. Međutim, taj se predznak ne poklapa uvijek sa aritmetičkim znakom kojeg dobijemo oduzimanjem. Kad velika vrijednost zavisne varijable pokazuje bolju učinkovitost, tada se poklapaju, ali kad su niske vrijednosti znak bolje učinkovitosti, onda se dobiveni predznak mora promijeniti.

- **Razlika udjela (eng. The Proportion Difference) - pd**

Rezultati usporedbe grupa mogu biti izraženi u obliku razlike između udjela, pa se taj jednostavni izraz koristi kao izračun veličine učinka:

$$ES_{pd} = p_{G1} - p_{G2},$$

$$SE_{pd} = \sqrt{p(1-p) \left(\frac{1}{n_{G1}} + \frac{1}{n_{G2}} \right)}, \quad (2.10)$$

$$w_{pd} = \frac{1}{SE_{pd}^2} = \frac{n_{G1}n_{G2}}{p(1-p)(n_{G1} + n_{G2})},$$

gdje je p_{G1} udio Grupe 1 a p_{G2} udio Grupe 2 u nekom događaju ili karakteristici, p je mjereni prosjek od p_{G1} i p_{G2} , tj. $p = (n_{G1}p_{G1} + n_{G2}p_{G2}) / (n_{G1} + n_{G2})$.

Ovaj izračun je jednostavan, ali to je jedina njegova prednost. Naime, nedostatak je taj što veličina učinka ovisi o udjelu p_{G1} ili p_{G2} , jer ako je $p_{G2} = 0.5$ tada je maksimalna veličina učinka također 0.5, dok je za $p_{G2} = 0.1$ maksimalna veličina

učinka = 0.9, što je teško interpretirati uspoređujući ovakve rezultate kroz više istraživanja.

- **Omjer šansi (eng. The Odds-Ratio) - OR**

Kod ovog izračuna uspoređuju se dvije grupe u terminima relativnih šansi nekog događaja, npr. uspješnog ishoda, smrti, bolesti itd. Često ga se zabunom interpretira kao omjer dvije vjerojatnosti. Da bi se shvatio pojam omjera šansi, prvo definiramo značenje šansi nekog događaja kao

$$\text{šansa nekog događaja} = \frac{p}{1-p}, \quad (2.11)$$

gdje je p vjerojatnost tog događaja.

Npr. ako je vjerojatnost uspješnog ishoda nekog tretmana 0.25, tada su šanse tog ishoda 0.33 tj. $0.25/(1-0.25)$. Znači da su šanse 1 od 3, gdje je vjerojatnost uspješnog ishoda 1 u 4, (jedan uspjeh u četiri slučaja). Neka je sad vjerojatnost uspješnog ishoda za kontrolnu grupu 0.2, znači da su šanse za uspješan ishod bez tretmana 0.25. Iz šansi uspješnog ishoda tretmana i kontrolne grupe izračunamo omjer šansi, a to je omjer te dvije šanse tj. $0.33/0.25 = 1.33$. Znači, za ovaj slučaj, šanse za uspješan ishod su 1.33 puta veće za eksperimentalnu nego za kontrolnu grupu.

Ovakav tip izračuna se primjenjuje na rezultate koji se temelje na dihotomnim varijablama, tj. takvim koje se dijele u dvije jasno odvojene kategorije (npr. prošao ispit-pao ispit) predstavljenim u formi relativnih frekvencija (Tablica 2.1.) ili udjela (Tablica 2.2.).

Tablica 2.1. Frekvencija tj. broj pripadnika neke grupe u određenom događaju

	Frekvencije	
	Događaj A	Nije događaj A
Grupa 1	a	b
Grupa 2	c	d

Tablica 2.2. Udio grupe u određenom događaju

	Udjeli	
	Događaj A	Nije događaj A
Grupa 1	$p_a = a/(a+b)$	$p_b = b/(a+b)$
Grupa 2	$p_c = c/(c+d)$	$p_d = d/(c+d)$

Događaji su tipa prošao ispit - pao ispit, izrazi a , b , c i d su frekvencije, odnosno broj ispitanika u određenom događaju, a p_a , p_b , p_c i p_d su udjeli svake grupe u određenom događaju.

Sada se omjer šansi kao izračun veličine učinka računa koristeći prethodne tablice na sljedeći način:

$$ES_{OR} = \frac{ad}{bc} = \frac{p_a p_d}{p_b p_c}. \quad (2.12)$$

Dobivena vrijednost se interpretira kao broj koliko puta Grupa 1 ima veću šansu u događaju A, (ili manju, ako je dobivena vrijednost ES_{OR} manja od 1). Iz toga je vidljivo da ovakav tip veličine učinka ima nepraktičnu tendenciju centriranja oko 1 umjesto oko 0. Tj. vrijednosti između 0 i 1 predstavljaju negativnu vezu (odnosno uspješniju kontrolnu grupu, ako se radi o eksperimentalnom istraživanju), u 1 nema nikakve veze (odnosno grupe su jednako uspješne), a vrijednosti iznad 1 pozitivnu vezu (uspješniju eksperimentalnu grupu). Tako, veličina učinka od 0.5, odgovara jačini veze veličine učinka 2 (obrnuto proporcionalno od 0.5) ali u suprotnom smjeru.

Da bi se to izbjeglo, analize se provode koristeći prirodni logaritam te veličine učinka. Tako dobijemo relativno normalnu distribuciju veličine učinka sa aritmetičkom sredinom u 0 i standardnom devijacijom 1.83. Ako tako dobivenu veličinu učinka (ES_{LOR}), podijelimo sa 1.83, dobijemo vrijednost koja je direktno usporediva sa veličinom učinka kao standardnom razlikom aritmetičkih sredina (ES_{sm}). Također, sada vrijedi da negativne vrijednosti odražavaju negativnu, a pozitivne vrijednosti pozitivnu vezu.

$$ES_{LOR} = \log_e(ES_{OR}),$$

$$SE_{LOR} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}, \quad (2.13)$$

$$w_{LOR} = \frac{1}{SE_{LOR}^2} = \frac{abcd}{ab(c+d) + cd(a+b)},$$

Problem se javlja ako je neka od vrijednosti u tablici jednaka 0. Tada se ne može izvršiti izračun većine danih formula, pa se to rješava, uz minimalni utjecaj na točnost veličine učinka, tako da se u tablici svakoj vrijednosti doda 0.5.

c) Povezanost između varijabli

Ponekad istraživači žele pronaći odgovor na pitanje koliko su neke dvije varijable međusobno povezane, npr povezanost socijalno-ekonomskog statusa obitelji pojedinca i njegovog uspjeha u srednjoj školi. Za izračun veličine učinka ovog tipa istraživanja koriste se: *Pearsonov koeficijent korelacije* (eng. Pearson product-moment correlation coefficient) i njegove varijacije, ali i standardna razlika aritmetičkih sredina i omjer šansi.

Standardna korelacija se koristi kad se uspoređuju dvije kontinuirane varijable (takve koje mogu poprimiti različite vrijednosti, od niskih do visokih). Kad se mjere varijable koje se razlikuju od navedenih, tada se koriste varijacije koeficijenta korelacije, tako se npr. *point-biserijalni* (eng. point-biserial correlation coefficient) koristi za povezanost dihotomne i kontinuirane varijable, a *fi-koeficijent* (eng. phi coefficient), za povezanost dvije dihotomne varijable.

- **Povezanost dvije dihotomne varijable**

Kako je već spomenuto, rezultati koji daju podatke o povezanosti dvije dihotomne varijable mogu se obrađivati omjerom šansi za izračun veličine učinka, ili koeficijentom korelacije (fi- koeficijentom).

Dok se omjer šansi općenito interpretira kao usporedba dvije podgrupe na nekoj zavisnoj dihotomnoj varijabli, fi-koeficijent se prirodno definira kao povezanost između dvije varijable mjerene na zadanom uzorku. Stoga, kad se rezultati

istraživanja fokusiraju na razlike između dvije prirodno različite grupe vezano uz dihotomnu zavisnu varijablu, npr. razlike spolova vezano uz stupanj uspješnosti u nekom određenom području znanosti, preporuča se upotrijebiti omjer šansi za izračun veličine učinka. Kad se rezultati fokusiraju na povezanost dvije varijable npr. uspjeh-neuspjeh na testovima iz dva školska predmeta, tada je bolje koristiti koeficijent korelacije.

- **Povezanost dihotomne i kontinuirane varijable**

Rezultati istraživanja u formi povezanosti između dihotomne (npr. muški - ženske, eksperimentalna-kontrolna grupa) i kontinuirane varijable (npr. rezultati testa) mogu se obrađivati standardnom razlikom aritmetičkih sredina ili koeficijentom korelacije tj. point-biserijalnim koeficijentom. Taj koeficijent ima restrikcije ako je dihotomna varijabla ekstremno raspoređena, npr. 80%-20% ili više, pa se ne preporuča korištenje u tim slučajevima. Tada se koristi standardna razlika aritmetičkih sredina koja se zbog lakše interpretacije može pretvoriti u formulu za izračun point-biserijalnog koeficijenta:

$$r_{pb} = \frac{ES_{sm}}{\sqrt{\frac{1}{p(1-p)} + ES_{sm}^2}} \quad (2.14)$$

gdje je ES_{sm} standardna razlika aritmetičkih sredina, p je udio ispitanika u grupi 1, a $(1-p)$ udio ispitanika u grupi 2.

Ako su grupe jednako brojne, odnosno $p = (1-p) = 0.5$, tada se prethodna formula reducira na:

$$r_{pb} = \frac{ES_{sm}}{\sqrt{4 + ES_{sm}^2}} \quad (2.15)$$

Isto se tako može izračunati standardna razlika aritmetičkih sredina, ako je poznat point-biserijalni koeficijent:

$$ES_{sm} = \frac{r_{pb}}{\sqrt{p(1-p)(1-r^2)}} \quad (2.16)$$

Ako je $p = 0.5$, ova formula se reducira na:

$$ES_{sm} = \frac{2r_{pb}}{\sqrt{1+r^2}} \quad (2.17)$$

- **Povezanost dvije kontinuirane varijable**

Kod rezultata istraživanja koji uspoređuju vezu dvije kontinuirane varijable, za izračun veličine učinka koristi se koeficijent korelacije. Ostale metode se vrlo

rijetko koriste u ovakvom slučaju. Koeficijent korelacije koji opisuje povezanost varijabli x i y se definira na sljedeći način:

$$r_{xy} = \frac{\sigma_{xy}^2}{\sigma_x \sigma_y}, \quad (2.18)$$

gdje je σ_{xy}^2 kovarijanca između x i y , a σ_x i σ_y su redom standardne devijacije od x i y . Ovaj izraz za koeficijent korelacije je već u standardiziranom obliku pa se kao takav koristi za usporedbu dvije varijable koje mogu biti obrađene i mjerene međusobno različitim mjernim metodama, i daje izlazne rezultate u intervalu od -1 do 1. Međutim, problem se javlja kad je potrebno pronaći standardnu pogrešku. Zato, zbog jednostavnijeg izračuna veličine učinka, koeficijent korelacije se Fisherovom Z_r transformacijom (Hedges i Olkin, prema [LIPS2001]) definira kao:

$$ES_{Z_r} = 0.5 \log_e \left[\frac{1+r}{1-r} \right], \quad (2.19)$$

gdje je r koeficijent korelacije. Zbog lakše interpretacije, inverznim se postupkom može iz ove formule dobiti koeficijent korelacije, na sljedeći način:

$$r = \frac{e^{2ES_{Z_r}} - 1}{e^{2ES_{Z_r}} + 1}, \quad (2.20)$$

gdje je r individualni ili srednji koeficijent korelacije.

Sad možemo napisati sve formule potrebne za izračun veličine učinka:

$$\begin{aligned} ES_r &= r, \\ ES_{Z_r} &= 0.5 \log_e \left[\frac{1+ES_r}{1-ES_r} \right], \\ SE_{Z_r} &= \frac{1}{\sqrt{n-3}}, \\ w_{Z_r} &= \frac{1}{SE_{Z_r}^2} = n-3, \end{aligned} \quad (2.21)$$

gdje je r koeficijent korelacije a n ukupan broj ispitanika.

VEZE KOJE UKLJUČUJU VIŠE VARIJABLI (ENG. MULTIVARIATE RELATIONSHIPS)

Postoji mnogo zanimljivih istraživanja koja uključuju više od dvije varijable. Tako se npr. neka pitanja istražuju primjenom multiple regresije, analizom faktora, analizom diskriminante, i sl.

Ovakva istraživanja predstavljaju veliki izazov za meta-analizu, pa su tako Greenwald, Hedges i Laine (Greenwald prema [LIPS2001]) sintetizirali kvantitativni više-varijableni model veze između ulaznih resursa i školskog uspjeha. Rezultati, koji su bili važni za ovo istraživanje, predstavljeni su u obliku koeficijenta regresije za ulazne resurse (učenički izdatak, učiteljeva plaća i sl.) u procjeni uspjeha u školi. Međutim, različite nezavisne

varijable kompliciraju sintezu, pa dobiveni standardni koeficijent regresije iz svake analize pokazuje različiti parametar populacije. Isto tako standardna pogreška svakog koeficijenta, općenito se ne može izračunati iz podataka dobivenih iz istraživanja. Zato se ne može formalno, kao do sad, definirati formula za izračun veličine učinka. Greenwald i ostali, ovakve su situacije rješavali u dva koraka. Prvo su provodili kombinirani test značajnosti, koji pokazuje statističku značajnost, tj. odbijanje ili prihvatanje nul-hipoteze koja glasi: "Nema nikakve povezanosti između ulaznih resursa i uspjeha u školi". U drugom dijelu su proveli "ograničenu" analizu standardnog koeficijenta regresije, računajući medijan koeficijenta regresije za svaku nezavisnu varijablu (tj. učenički izdatak, plaća učitelja, i drugi resursi), pa su tako dobili određeni skup zaključaka.

Općenito, nije moguće napraviti kompletan izračun veličine učinka za meta-analizu više-varijablenih istraživanja, ali se za određenu svrhu može napraviti neka vrsta opisne usporedbe kao što je to napravio Greenwald.

Drugačiji pristup ovom problemu, koji se sve više razvija u meta-analizi, je uzimanje korelacija iz originalnih istraživanja, pa se tad analiza više varijabli može obavljati na sintetiziranoj matrici korelacije. Primjeri ovakvih analiza se nalaze u radovima Becker-a, te Premack-a i Hunter-a (prema [LIPS2001]). Becker je tako sintetizirao korelaciju između prediktora uspjeha u znanosti za muškarce i žene posebno, i izveo više-varijablenu analizu za utvrđivanje da li je isti skup varijabli bio važan za oba spola.

Ovakav pristup je obećavajući, međutim, izostavljanje cijele matrice korelacije iz većine izvještaja više-varijablenih istraživanja, ograničava njegovu upotrebu. Meta-analiza više-varijablenih rezultata će biti uvelike olakšana, ako se promjene publikacijske norme istraživačkih rezultata na način da objavljivanje matrica korelacija postane standard.

2.3. PRIKUPLJANJE VAŽNIH INFORMACIJA IZ POJEDINOG ISTRAŽIVANJA

Kod provedbe detaljne meta-analize, preporučuje se primijeniti postupak kojim će se specificirati koje informacije treba izdvojiti iz odabranih istraživanja. Tako analitičar najprije sastavi *upitnik* (eng. questionnaire), tj. niz pitanja o istraživanjima i njihovim rezultatima, unaprijed ih formulirajući tako da za odgovore dobije sve važne informacije iz pojedinog istraživanja. Zatim pregledava istraživanje po istraživanje i popunjava taj upitnik. Pri tom procesu je vrlo važno voditi brigu o točnosti odnosno pouzdanosti podataka pri upisivanju.

2.3.1. IZRADA UPITNIKA

Upitnik je sastavljen tako da prvi dio sadrži pitanja o samom istraživanju, o njegovim karakteristikama, a u drugom dijelu se pitanja odnose na dobivene rezultate (npr. veličina učinka). Razlika između ta dva dijela je slična onoj između nezavisne i zavisne varijable, tako su rezultati istraživanja u obliku veličine učinka zavisne varijable meta-analize, tj. one su

izlazne vrijednosti empirijskih istraživanja. Karakteristike nekog istraživanja, npr. metode, mjerne procedure, uzorci, veze, tretmani i sl. predstavljaju nezavisnu varijablu meta-analize, tj. one predstavljaju faktore koji mogu utjecati na prirodu i veličinu dobivenih rezultata.

Cilj ovakvog upitnika je stvaranje baze podataka za statističku analizu, zato bi analitičar trebao unaprijed odrediti koji tip odgovora se upisuje za pojedina pitanja. To bi značilo definiranje unaprijed formuliranih odgovora od kojih se odabire odgovarajući, ili se unosi specifični iznos (npr. broj ispitanika). Ponekad je nemoguće odrediti kategoriju odgovora, pa se ostavlja otvoreno mjesto za odgovor, koji kasnije treba detaljno proučiti i obraditi.

2.3.2. POPUNJAVANJE UPITNIKA

Prvi dio upitnika sadrži opis samog istraživanja. Na početku se svako istraživanje označi identifikacijskim brojem (study ID). Nakon toga slijede pitanja kojima se opisuje uzorak ispitanika: prosječna starost ispitanika, rasa koja dominira (npr. više od 70% bijelci), spol koji dominira (više muškaraca), itd. Zatim slijedi opis strukture istraživanja, odnosno pitanja o tipu istraživanja: da li se istraživanje provodilo uz pomoć eksperimenta, način formiranja kontrolne i eksperimentalne grupe, provođenje testa ekvivalencije grupa prije početka eksperimenta, veličina uzorka svake grupe, itd. Zatim opis prirode eksperimenta (tretmana): trajanje tretmana, intenzitet tretmana, status kontrolne grupe (npr. ne dobiva nikakav tretman, ili dobiva placebo...).

Drugi dio sadrži pitanja o podacima vezanim uz statističke izračune: opis mjernih procedura za zavisne varijable, zatim koja grupa pokazuje veću uspješnost, aritmetičke sredine kontrolne i eksperimentalne grupe, standardna devijacija, udio itd. I na kraju veličina učinka i njezina preciznost (težina).

2.3.3. STVARANJE PODATKOVNIH DATOTEKA

Nakon prethodnog procesa, sastavljena je neka vrsta zbirke u kojoj su pohranjeni najvažniji podaci iz svih istraživanja, koji su potrebni za meta-analizu. Međutim, nisu svi tako dobiveni podaci potrebni za izračun veličine učinka i provedbu ostalih statističkih analiza. Zato je potrebno izdvojiti one rezultate, koji su potrebni za vršenje potrebnih izračuna i te podatke na neki način složiti, grupirati, da bi se mogla lakše provesti međusobna usporedba i analiza. To se naziva stvaranje *meta-analitičkih podatkovnih datoteka* (eng. meta-analytic data files). Obično se provodi na način da se kreiraju tablični zapisi, koji sadrže imena najvažnijih varijabli i u polja tablice se unose vrijednosti tih varijabli iz svih istraživanja. Iako bi se to moglo napraviti ručno za neku manju meta-analizu, ovaj proces ipak najčešće zahtjeva upotrebu računala. Izbor programske podrške koja će se koristiti u ovu svrhu ovisi o meta-analizi, ali i o dostupnim programima i samom autoru. Tako se podaci mogu unositi u računalo direktno kao dokument u neki program za obradu teksta kao što je Microsoft Word, u obliku tablice u Microsoft Excel, Lotus 123 ili neki slični program, gdje stupci predstavljaju varijable a redovi zapise. Isto se tako mogu upisivati u obliku baze podataka u Microsoft Access ili u neki drugi slični program. Složeni u podatkovne datoteke, podaci su spremni za analizu.

2.4. ANALIZA PODATAKA

U meta-analizi jedinica analize je individualno istraživanje. Prvi problem predstavlja činjenica da istraživanja u meta-analizi često generiraju više od jedne veličine učinka. Iako je za analizu važan cijeli taj skup dobivenih veličina učinka, svake dvije ili više koje dolaze iz istog istraživanja (od istog uzorka ispitanika) statistički su zavisne. Ako ih uključimo u istu meta-analizu, kršimo pretpostavku nezavisnih podataka koja je osnovna za statističku analizu.

Drugi problem je baš u tome što jedinicu analize čini individualno istraživanje, a kolekcija istraživanja je obično šarolika, u smislu da različita istraživanja obično imaju različit broj ispitanika (nejednaki uzorci). Zbog činjenice da su veličine učinka izvedene iz podataka vezanih uz uzorak (aritmetička sredina, standardna devijacija), njihova statistička svojstva dijelom ovise o toj veličini uzorka. Kao što je već bilo spomenuto, veličina učinka koja se temelji na velikom uzorku sadrži manju grešku, odnosno preciznija je od one veličine učinka temeljene na malom uzorku. Ako provodimo statističku analizu veličine učinka bez vođenja računa o veličini uzorka, mi ih tretiramo kao da svaka veličina učinka daje isti doprinos rezultatu. Međutim, intuitivno je jasno da bi veličine učinka temeljene na većim uzorcima, trebale imati veću ulogu u analizi upravo zbog manje greške, tj. trebale bi više "težiti" u statističkom izračunu od onih temeljenih na manjim uzorcima. Zato je u meta-analizi veličina učinka "mjerena" određenom vrijednošću, njenom težinom, tako da je njezin doprinos statističkoj analizi proporcionalan s njenom preciznošću.

Za analizu prikupljenih podataka, koji su pohranjeni u obliku podatkovnih datoteka, potrebno je slijediti nekoliko koraka. Najvažniji dio tih datoteka su, naravno, dobivene veličine učinka svih korištenih istraživanja kao i njihove težine. Slijede faze analize tih podataka.

2.4.1. PRILAGODBE VELIČINE UČINKA

Ponekad je potrebno dodatno prilagoditi odnosno korigirati dobivene rezultate, pa tako i veličinu učinka. Ako nije obavljena korekcija tijekom samog izračuna, onda se u ovoj fazi izrade meta-analize dobiveni rezultati ponovno provjeravaju, i vrše se određeni ispravci ako su potrebni.

TRANSFORMACIJE I ISPRAVCI ZABLUDA

Kod izračuna veličine učinka kao standardne razlike aritmetičkih sredina, bilo je pokazano korištenje korekcije izračuna jer originalna formula daje lažno veću vrijednost ako se temelji na malom uzorku. Slično vrijedi i za koeficijent korelacije, koji se obično transformira u prikladan oblik za analizu a onda opet vrati u originalan oblik za bolju prezentaciju rezultata, npr. Fisher-ova Z_r - transformacija.

IZOLIRANE (EKSTREMNE) VRIJEDNOSTI

Ponekad se u analizi naiđe na ekstremne vrijednosti veličine učinka, takve koji su vidljivo proturječni većini drugih podataka u istraživanju i tako nisu reprezentativni za to istraživanje, a mogu biti i lažni. Takve vrijednosti imaju neproporcionalan utjecaj na aritmetičku sredinu,

varijancu i ostale izračune koji se koriste u meta-analizi i mogu odvratiti analizu u pogrešan smjer. Zato je preporučljivo pregledati distribuciju veličina učinka da se ustanovi prisutnost ekstremnih vrijednosti. Ako takve postoje, preporuča ih se izostaviti iz analize, ili ih podesiti na niže vrijednosti – to se ponekad u analizi zove „*windsorizing*“.

2.4.2. ANALIZA SKUPA VELIČINA UČINKA I NJIHOVE DISTRIBUCIJE

Najvažniji dio svake meta-analize svakako je izračun ukupne veličine učinka svih korištenih istraživanja, jer ona predstavlja bit same meta-analize. Odgovara na njezina najvažnija pitanja: da li je metoda ili tretman, čije se djelovanje istražuje, učinkovit? Kolika je ta učinkovitost i da li je statistički značajna? Slijedi opis analize skupa dobivenih veličina učinka.

2.4.2.1 STVARANJE NEZAVISNOG SKUPA VELIČINA UČINKA

Veličina učinka se smatra statistički nezavisnom, ako za danu distribuciju ne dolazi više od jedne veličine učinka od istog uzorka ispitanika. Ako istraživanje donosi više od jedne veličine učinka za danu vezu, npr. korištenjem različitih mjernih instrumenata, oni ne bi smjeli biti uključeni u istu analizu kao da su nezavisni podaci. Neka istraživanja ispituju učinke eksperimentalnog faktora različitim testovima, pa to rezultira sa više veličina učinka dobivenih istim istraživanjem. Tada se savjetuje odabir jedne, slučajnim odabirom ili na temelju nekog kriterija, npr. ako analitičar utvrdi da je određeni test kvalitetniji od ostalih ili se rezultati tog testa temelje na najvećem uzorku, uzima u obzir samo njegove rezultate. Ako analitičar procijeni da su testovi podjednako kvalitetni, a isti broj ispitanika je sudjelovao u njihovom rješavanju, može odrediti aritmetičku sredinu veličina učinka dobivenih tim testovima. Ta prosječna veličina učinka predstavlja stvarni učinak eksperimenta kojeg istraživanje proučava, i kao takva se koristi za daljnju analizu.

Postoji alternativa stvaranju nezavisnog skupa veličina učinka. To je statističko modeliranje zavisnosti između veličina učinka i tako dopuštanje korištenja više veličina učinka iz istog uzorka u analizi. Za takvu analizu treba biti poznata kovarijanca između zavisnih veličina učinka (Gleser i Olkin prema [LIPS2001]). Nažalost, malo situacija u meta-analizi daje dovoljno informacija da bi se dopustilo implementiranje ove metode, koja od provoditelja analize zahtjeva i veću statističku sofisticiranost.

Pretpostavimo, dakle, da je skup veličina učinka, koje su spremne za daljnju analizu, statistički nezavisan.

2.4.2.2 IZRAČUN UKUPNE VELIČINE UČINKA

Ukupna veličina učinka svih dobivenih veličina učinka u nekoj meta-analizi, računa se „mjeranjem“ svake veličine učinka (ES_i) njenom „težinom“ (w_i). To se naziva *prosječna mjerena veličina učinka* (eng. weighted mean effect size) i računa se prema sljedećoj formuli:

$$\overline{ES} = \frac{\sum (w_i ES_i)}{\sum w_i}, \quad (2.22)$$

gdje su ES_i vrijednosti dobivenih veličina učinka, w_i njihove težine, gdje je i od 1 do k , a k je ukupan broj veličina učinka.

2.4.2.3 INTERVAL POUZDANOSTI OKO PROSJEČNE VELIČINE UČINKA

Interval pouzdanosti pokazuje raspon vrijednosti u kojem će se nalaziti prosječna veličina učinka neke populacije, prema zadanim podacima.

Interval pouzdanosti za prosječnu veličinu učinka se temelji na standardnoj pogrešci tog prosjeka i kritičnoj vrijednosti iz z-distribucije (npr. 1.96 za $\alpha=0.05$). Formula za standardnu pogrešku prosječne veličine učinka je

$$SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_i}}, \quad (2.23)$$

gdje je $SE_{\overline{ES}}$ standardna pogreška prosječne veličine učinka a w_i kao prethodno.

Da bi se sad odredio interval pouzdanosti koriste se formule za donju i gornju granicu na sljedeći način:

$$\begin{aligned} \overline{ES}_L &= \overline{ES} - z(SE_{\overline{ES}}), \\ \overline{ES}_U &= \overline{ES} + z(SE_{\overline{ES}}), \end{aligned} \quad (2.24)$$

gdje su \overline{ES}_L i \overline{ES}_U , redom donja i gornja granica tog intervala, z je kritična vrijednost z-distribucije i za interval pouzdanosti od 95% vrijedi $z = 1.96$ ($\alpha=0.05$), a za interval pouzdanosti od 99%, $z = 2.58$ ($\alpha=0.01$). Ako interval pouzdanosti ne uključuje 0, prosječna veličina učinka je statistički značajna za $p \leq \alpha$.

Direktan se test značajnosti prosječne veličine učinka može dobiti računanjem z-testa kao:

$$z = \frac{|\overline{ES}|}{SE_{\overline{ES}}}, \quad (2.25)$$

gdje je $|\overline{ES}|$ apsolutna vrijednost prosječne veličine učinka. Ako ova vrijednost prelazi 1.96 tada je statistički značajna za $p \leq 0.05$, a ako prelazi 2.58 tada je statistički značajna za $p \leq 0.01$.

2.4.2.4 ANALIZA HOMOGENOSTI

Postavlja se pitanje, da li se sve dobivene različite veličine učinka koje su skupljene u prosječnu vrijednost, odnose na istu populaciju. To je pitanje homogenosti distribucije veličine učinka.

U homogenoj distribuciji, disperzija veličina učinka oko njihovog prosjeka nije veća od očekivane greške vezane uz uzorak. Drugim riječima, u homogenoj distribuciji individualna se veličina učinka razlikuje od prosječne samo za tu grešku.

Ako distribucija nije homogena, to može značiti da razlike između veličina učinaka variraju zbog nekog drugog razloga, možda zbog različitih karakteristika korištenih istraživanja.

Test homogenosti se temelji na izračunu Q vrijednosti, koji predstavlja distribuciju χ^2 (hi-kvadrat) sa $k-1$ stupnjeva slobode gdje je k broj veličina učinka (Hedges i Olkin prema [LIPS2001]). Formula glasi:

$$Q = \sum w_i (ES_i - \overline{ES})^2, \quad (2.26)$$

Ako dobivena vrijednost Q prijeđe kritičnu vrijednost za χ^2 , za $k-1$ stupnjeva slobode, tada se nul hipoteza homogenosti odbacuje. Analogna formula za homogenost, lakša za izračun glasi:

$$Q = \left(\sum w_i ES_i^2 \right) - \frac{\left(\sum w_i ES_i \right)^2}{\sum w_i}. \quad (2.27)$$

Ako distribucija nije homogena, postoje analitički pristupi obrade takvih situacija. Tada su potrebna dodatna ispitivanja mogućih faktora koji su utjecali na heterogenost distribucije, i utvrđivanje da li su i u kolikoj mjeri utjecali na tako dobivene veličine učinka.

2.5. INTERPRETACIJA I KORIŠTENJE REZULTATA META-ANALIZE

Okosnica svake meta-analize je prikaz rezultata u obliku ukupne veličine učinka. Zato se javlja poteškoća u razumijevanju tih rezultata, jer je upravo veličina učinka ta, koja nije intuitivno jasna kao brojčani izraz.

Da bi se pravilno interpretirala veličina učinka, potrebno je dobivene rezultate uvrstiti u neki kontekst koji je moguće jednostavnije interpretirati. Postoje različiti načini da bi se to postiglo i nijedan nije potpuno zadovoljavajući u svakoj situaciji.

Tako je npr. Cohen (prema [LIPS2001]), objavio svoje generalne zaključke, tj. da veličina učinka izračunata kao standardna razlika aritmetičkih sredina pada u jednu od tri kategorije (Tablica 2.3.).

Tablica 2.3. Cohen-ova kategorizacija veličine učinka kao standardne razlike aritmetičkih sredina

mala	srednja	velika
$ES \leq 0.2$	$ES = 0.5$	$ES \geq 0.8$

Analogne vrijednosti za veličinu učinka u obliku koeficijenta korelacije su prikazane u (Tablica 2.4.).

Tablica 2.4. Cohen-ova kategorizacija veličine učinka kao koeficijenta korelacije

mala	srednja	velika
$r \leq 0.1$	$r = 0.25$	$r \geq 0.4$

Cohen, međutim, nije sistematski objasnio svoju generalizaciju o tome što je tipično "mali", "srednji" i "veliki" učinak u istraživanju.

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

2.5.1. PREVOĐENJE UKUPNE VELIČINE UČINKA U DRUGE METRIKE

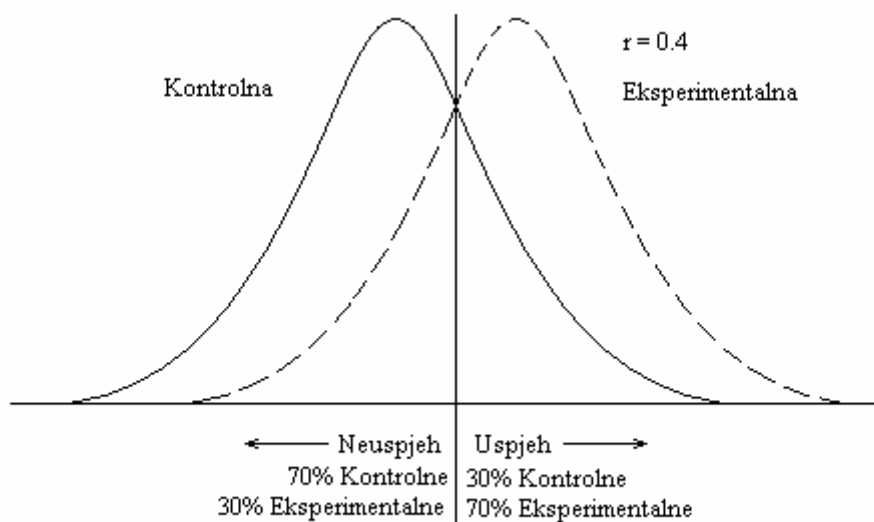
Sljedeći način interpretacije vrijednosti veličine učinka je njezino translatiranje u neku drugu, intuitivno jasniju metriku. U tom slučaju analitičar odredi aritmetičku sredinu i standardnu devijaciju kontrolne grupe eksperimenta koji ga zanima. Sada se veličina učinka tog eksperimenta prikaže tako da se aritmetička sredina eksperimentalne grupe odredi na toj istoj skali. To jednostavno znači da je aritmetička sredina eksperimentalne grupe jednaka zbroju aritmetičke sredine kontrolne grupe i umnoška veličine učinka i standardne devijacije.

Npr. ako neki eksperiment pokazuje veličinu učinka 0.3, a aritmetička sredina je 3.5 i standardna devijacija 1.5 za kontrolnu grupu na skali koja poprima vrijednosti od minimalno 1 do maksimalno 5. Sada znamo da je aritmetička sredina eksperimentalne grupe 0.3 standardne devijacije veća od aritmetičke sredine kontrolne grupe. Umnožak 0.3 i 1.5 nam govori da je to ekvivalentno 0.45 jedinica na danoj skali, tj. aritmetička sredina eksperimentalne grupe je 3.95 (oko 4). Tako smo dobili da je prosječna veličina učinka od 0.3 ekvivalentna porastu od 3.5 do 4 na danoj skali (od 1 do 5), što je većini istraživača i analitičara značajniji i jasniji podatak, nego samo saznanje da postoji razlika od 0.3 standardne devijacije između eksperimentalne i kontrolne grupe.

2.5.2. INTERPRETACIJA UKUPNE VELIČINE UČINKA U TERMINIMA PREKLAPANJA DISTRIBUCIJA

Rezultat meta-analize, odnosno ukupna veličina učinka se može interpretirati na nekoliko načina u terminima preklapanja distribucija uspjeha dvije grupe (npr. kontrolne i eksperimentalne), tako da je razlika između njih jasnije vidljiva. Najčešće korišten pristup su

razvili Rosenthal i Rubin (prema [LIPS2001]), a naziva se *binomni prikaz veličine učinka* (eng. Binomial Effect Size Display - BESD). Ovaj prikaz prevodi koeficijent korelacije u ekvivalentnu razliku u stopi uspjeha za dvije populacije, na sljedeći način. Prvo se "prag uspjeha" odredi na medijanu distribucija zavisne varijable (u ovom slučaju veličine učinka) obje grupe. Sada razlikujemo populaciju na temelju nezavisne varijable odnosno pripadnosti kontrolnoj ili eksperimentalnoj grupi. Na kraju se postavlja pitanje koliki je udio određene grupe iznad "praga uspjeha" i uspoređujemo grupe da bi dobili slikovit prikaz razlike (učinka) u terminima jednostavne procjene uspjeha (Slika 2.1.).



Slika 2.1. Distribucija BESD prikaza za koeficijent korelacije 0.4

Slika pokazuje unaprijed zadani „prag uspjeha“ postavljen na medijanu obje distribucije, i po definiciji je uvijek na 50%. Korisno svojstvo BESD-a je da je razlika uspjeha između dvije dane distribucije uvijek jednaka koeficijentu korelacije. Ako je ukupna veličina učinka izražena kao koeficijent korelacije $r = 0.4$, tada razlika između postotka niže distribucije iznad praga uspjeha i postotka više distribucije iznad praga uspjeha iznosi 40%. Dakle, ako je prag postavljen na 50%, to bi značilo 20% ispod i 20% iznad tog praga, tj. postotak uspješnosti od 30% za nižu i 70% za višu distribuciju. Za većinu je tako jednostavnije razumjeti ukupnu veličinu učinka ako je prikazana kao razlika između 30% i 70% uspješnosti nego ako je izražena kao koeficijent korelacije u vrijednosti $r = 0.4$.

Vidljivo je da je ovakav diferencijal jednostavnije izračunati ako je veličina učinka izražena kao koeficijent korelacije. Međutim, ako je poznata veličina učinka kao standardna razlika aritmetičkih sredina (ES_{sm}), ona se može lako pretvoriti u odgovarajući koeficijent korelacije formulom:

$$r = \frac{ES_{sm}}{\sqrt{4 + ES_{sm}^2}} \quad (2.28)$$

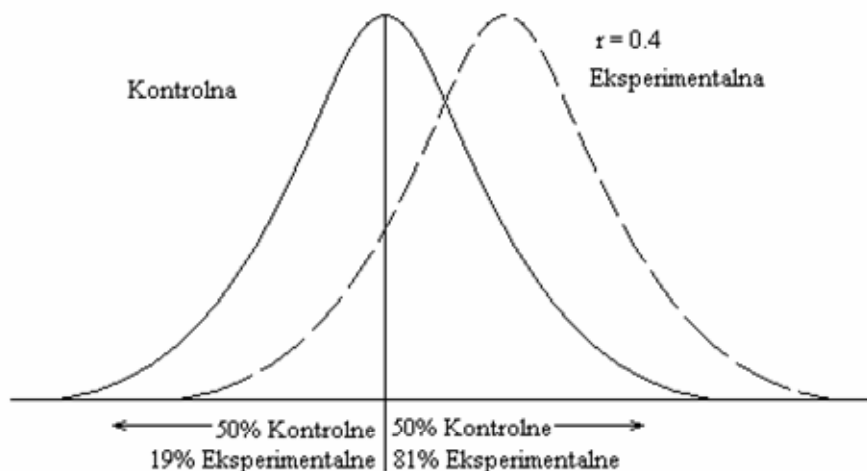
Još je jednostavnije pronaći odgovarajući postotak uspješnosti za veličinu učinka izraženu kao standardnu razliku aritmetičkih sredina ili kao koeficijent korelacije u već gotovoj tablici (Tablica 2.5.).

Tablica 2.5. Veličina učinka u obliku standardne razlike aritmetičkih sredina i njezini ekvivalenti u obliku koeficijenta korelacije r , U3 i BESD distribucije

ES_{sm}	r	U3:% E.G. iznad prosjeka K.G	BESD: % uspješnosti K.G. nasuprot E.G.		BESD razlika između E.G i K.G.
0.1	0.05	54	0.47	0.52	0.05
0.2	0.10	58	0.45	0.55	0.10
0.3	0.15	62	0.42	0.57	0.15
0.4	0.20	66	0.40	0.60	0.20
0.5	0.24	69	0.38	0.62	0.24
0.6	0.29	73	0.35	0.64	0.29
0.7	0.33	76	0.33	0.66	0.33
0.8	0.37	79	0.31	0.68	0.37
0.9	0.41	82	0.29	0.70	0.41
1.0	0.45	84	0.27	0.72	0.45
1.1	0.48	86	0.26	0.74	0.48
1.2	0.51	88	0.24	0.75	0.51
1.3	0.54	90	0.23	0.77	0.54
1.4	0.57	92	0.21	0.78	0.57
1.5	0.60	93	0.20	0.80	0.60
1.6	0.62	95	0.19	0.81	0.62
1.7	0.65	96	0.17	0.82	0.65
1.8	0.67	96	0.16	0.83	0.67
1.9	0.69	97	0.15	0.84	0.69
2.0	0.71	98	0.14	0.85	0.71
2.1	0.72	98	0.14	0.86	0.72
2.2	0.74	99	0.13	0.87	0.74
2.3	0.75	99	0.12	0.87	0.75
2.4	0.77	99	0.11	0.88	0.77
2.5	0.78	99	0.11	0.89	0.78
2.6	0.79	99	0.10	0.89	0.79
2.7	0.80	99	0.10	0.90	0.80
2.8	0.81	99	0.09	0.90	0.81
2.9	0.82	99	0.09	0.91	0.82
3.0	0.83	99	0.08	0.91	0.83

Postoji i varijacija na BESD prikaz. To je još jedna slična forma prikaza, u kojoj se prag uspješnosti postavi na medijan niže distribucije, umjesto na zajednički medijan obiju distribucija. Iako je odabir tog praga uspješnosti dogovoran, ipak postoji prednost ako ga stavimo na "sredinu" kontrolne grupe, kod eksperimentalnih usporedbi. U ovom slučaju uspješnost je 50% za donju distribuciju (po definiciji), a veličina učinka se interpretira tako da se utvrdi koliko je bolja uspješnost gornje distribucije.

Slika 2.2. pokazuje udjele u uspješnosti pojedinih grupa, i odgovara mjerenju koje je Cohen (prema [LIPS2001]) nazvao U3. Isto tako Tablica 2.5. pokazuje udjele uspješnosti za ove U3 usporedbe. Npr. ako imamo ukupnu veličinu učinka nekog eksperimenta kao standardnu razliku aritmetičkih sredina koja iznosi 0.9, to je ekvivalentno koeficijentu korelacije $r = 0.41$ (po tablici). Tada se taj učinak može prikazati kao postotak uspješnosti od 82% za eksperimentalnu grupu, nasuprot uspješnosti od 50% za kontrolnu grupu. Odnosno, 82% ispitanika koji su bili podvrgnuti nekom eksperimentu se nalaze iznad medijana kontrolne grupe, dok je samo 50% ispitanika kontrolne grupe iznad tog nivoa.



Slika 2.2. Distribucija U3 usporedbi za koeficijent korelacije 0.4

2.5.3. KVALITETA REZULTATA META-ANALIZE

Meta-analiza predstavlja sintezu istraživanja. Stoga su njeni rezultati utoliko kvalitetni koliko su kvalitetna istraživanja koja uključuje. Ako ne postoje istraživanja sa visokom metodološkom kvalitetom u bazi istraživanja, teško se može očekivati da će udruživanje njihovih otkrića dati valjane i korisne rezultate. Meta-analitičar zato mora oprezno birati i koristiti istraživanja koja daju valjane i pouzdane rezultate, premda to nije nimalo jednostavna zadaća.

Istraživanja se često razlikuju u mnogo metodoloških dimenzija, tako da je nemoguće generalno odrediti koja su "dobra" a koja "loša". Analitičari se zato, u provođenju analize "snalaze" na različite načine, da bi dobili što kvalitetnije i pouzdanije rezultate. U kolekciji istraživanja, ponekad odrede istraživanja koja su visoke kvalitete, te njihove rezultate uspoređuju sa ostalim rezultatima. Ako nema značajnog odstupanja, koriste sva dobivena otkrića. Ako je slučaj obrnut, dobar analitičar će radije odbaciti rezultate koji mnogo odstupaju, ako su dobiveni istraživanjima "sumnjive" kvalitete. Ponekad pomoću kvalitetnih istraživanja unaprijed odrede raspon vrijednosti u kojem bi trebala biti krajnje dobivena ukupna veličina učinka.

U tom postupku provjere da li su korištena odgovarajuća istraživanja, pomaže i analiza homogenosti, koja također predstavlja vrstu "testa", pa ako je dobivena distribucija homogena, rezultat meta-analize, odnosno ukupna veličina učinka jasno i smisleno reprezentira korištena istraživanja.

U meta-analizi koja ispituje učinke nekoliko različitih metoda, krajnji rezultati se moraju oprezno obrazložiti i ne smiju se donositi površni zaključci na način da se neka metoda proglasi učinkovitijom od neke druge, i isključivo preporučiti za daljnje korištenje. Naime, često je slučaj da je pojedina metoda učinkovita u određenim uvjetima, a neka druga prikladnija pa zato i učinkovitija u nekim drugim uvjetima. Zato treba voditi računa o generalizaciji i donošenju zaključaka iz rezultata meta-analize. Ako nisu provedena dodatna ispitivanja i analize, koje podupiru te zaključke, bolje se suzdržavati od donošenja istih.

2.5.4. GRAFIČKO PRIKAZIVANJE REZULTATA

Grafičko prikazivanje rezultata je sastavni dio statističke analize, odnosno predstavlja svojevrsnu pomoć u analizi. Takvo vizualno predočavanje rezultata korisno je zbog mnogo razloga. Prikažemo li dobivene veličine učinka grafički, odmah možemo uočiti da li je njihova distribucija asimetrična, postoje li rezultati koji odudaraju od ostalih (ekstremi), imaju li podaci neku drugu karakteristiku ili nepravilnost. To su sve pojedinosti koje nam mogu promaknuti ako su rezultati izraženi samo numerički.

Većina grafičkih prikaza koji se primjenjuju u meta-analizi, samo su prilagodbe vizualnih prikaza koji su uobičajeni u društvenim i prirodnim znanostima. Nadalje će biti prikazani samo oni koji se zbog određenih svojstava koriste u meta-analizi. To su primjerice: steam-and-leaf prikaz, error-bar chart, box-and-whisker plot i drugi.

STEAM-AND-LEAF DISPLAY

Ovaj tip grafičkog prikaza daje uvid u aritmetičku sredinu, varijabilnost i normalnost distribucije veličina učinka. Često se koristi u meta-analizi, i njegova prednost pred ostalim grafičkim prikazima je u tome što se originalni podaci mogu ponovno izvući iz slike, znači svaki pojedinačni dobiveni rezultat se može očitati iz tog prikaza. Međutim, ovaj graf ima ograničenja, u smislu da nije pogodan za prikaz distribucija velikog broja rezultata istraživanja (Slika 2.3.).

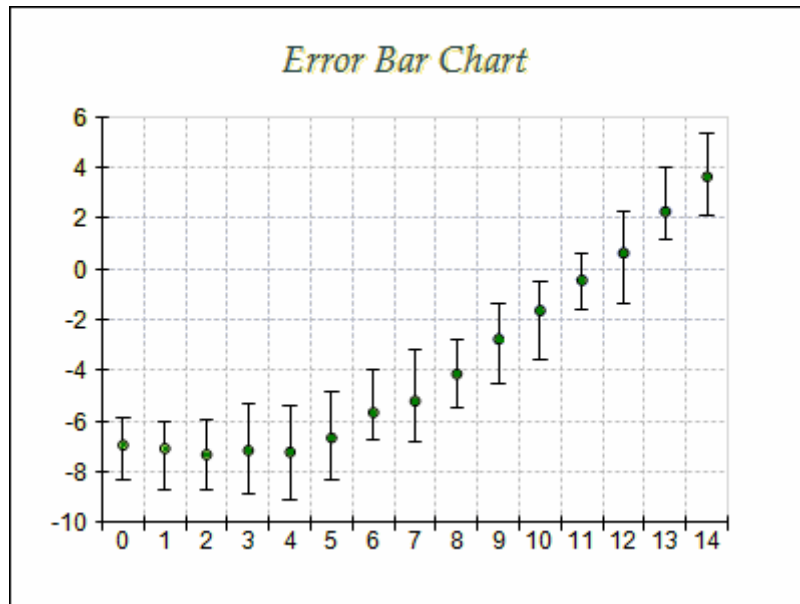
0.6	4
0.5	
0.4	
0.3	129
0.2	
0.1	57
0.0	0
-0.0	2
-0.1	
-0.2	
-0.3	33

Slika 2.3. Steam-and-leaf display s prikazom 10 veličina učinka

Slika pokazuje rezultate meta-analize koja obrađuje 10 veličina učinka, i te se vrijednosti mogu pročitati sa samog prikaza; npr 0.64, 0.31, 0.32, 0.39, 0.15, 0.17 itd.

ERROR-BAR CHART

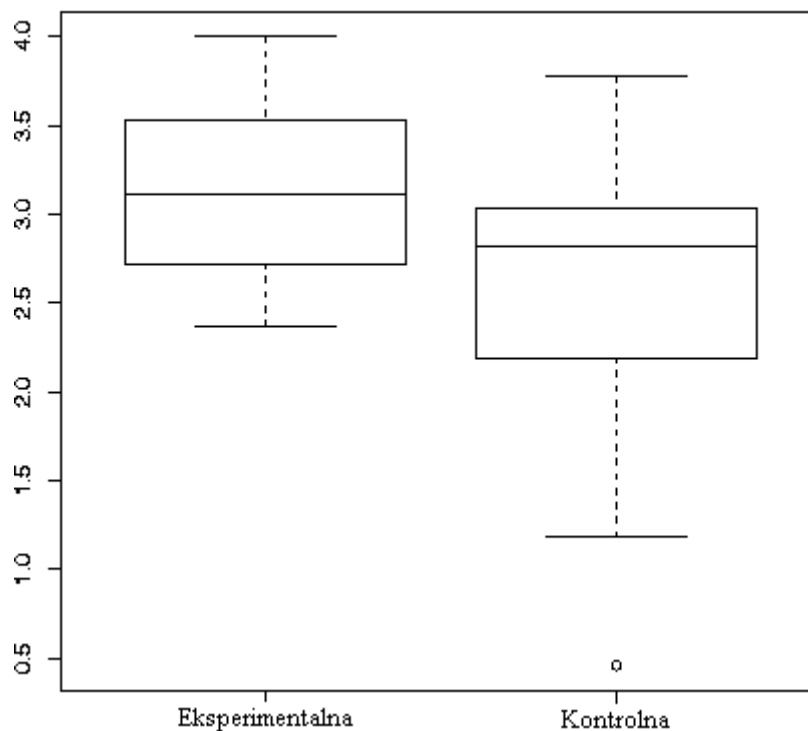
Error-bar chart pokazuje veličinu parametra koji nas zanima kao točku i njezin interval pouzdanosti kao vodoravnu ili okomitu crtu. Slika 2.4. je primjer ovakvog grafa s okomitim prikazom intervala pouzdanosti. Često se koristi u meta-analizi za prikazivanje individualnih veličina učinka i njihovih intervala pouzdanosti od 95%.



Slika 2.4. Error-bar prikaz parametra i njegovog intervala pouzdanosti

BOX-AND-WHISKER PLOT

Uobičajen grafički prikaz koji omogućuje uvid u centralnu tendenciju kao i disperziju distribucije određene varijable je box-and-whisker plot. Ovaj graf prikazuje medijan kao vodoravnu središnju liniju, 25. i 75. percentil čine vodoravne stranice pravokutnika, opseg vrijednosti su krajnje linije a ekstremi kružići (Slika 2.5.).



Slika 2.5. Box-and-whisker plot prikaz dvije distribucije neke varijable (eksperimentalna i kontrolna grupa)

2.6. PREDNOSTI I NEDOSTACI META-ANALIZE

Uzimajući u obzir sve navedeno, meta-analiza je zapravo istraživački proces koji se provodi u nekoliko faza: definiranje teme, pronalazak prikladnih istraživanja, odabir metode i izračun veličine učinka, obrada i analiza rezultata. Zato postoji nekoliko osnovnih razloga koji čine prednosti meta-analize pred konvencionalnim metodama.

Prvo, dobra meta-analiza se provodi kao strukturirana tehnika istraživanja i stoga zahtjeva da svaki korak bude dokumentiran i otvoren za pomno istraživanje, pa i za provjeru ispravnosti ako je potrebno. Uključuje definiranje kriterija, koji određuju vrstu rezultata istraživanja za tu meta-analizu, organizirane strategije traženja koje će identificirati i pronaći prikladna istraživanja. Zatim, sakupljanje i obradu podataka iz svakog istraživanja, te na kraju analizu rezultata koja će potvrditi zaključke koji su izvedeni. Takav proces istraživačkog sumiranja koje je eksplicitno i sistematsko, dopušta svakome pristup u autorove pretpostavke, procedure, dokaze i zaključke, je daleko bolje nego vjerovati na riječ autoru da su njegovi zaključci ispravni.

Drugo, meta-analiza prezentira rezultate istraživanja na način koji je različit i osjetljiviji od konvencionalnih postupaka, koji se oslanjaju na kvalitativne sažetke i statističku značajnost. Veličina učinka može poprimiti vrijednosti različite veličine i predznaka, u predstavljanju rezultata istraživanja. Na taj način, dobiveni rezultati imaju veću "vrijednost" nego samo korištenje statističke značajnosti, koja razlikuje istraživanja po tome koja pronalaze učinak, a koja ne.

Treće, meta-analiza pronalazi učinke i veze koji su prikriveni u drugim pristupima analize rezultata istraživanja. Kvalitativni sažeci istraživačkih rezultata ne dopuštaju tako pomno ispitivanje ispravnosti svojih zaključaka, i usporedbu razlika među istraživanjima. S druge strane sistematska obrada, karakteristična za meta-analizu, dopušta analitički preciznu provjeru veza između rezultata istraživanja, zatim karakteristike ispitanika, prirodu tretmana koji se provodi, strukturu istraživanja i mjerne postupke koji će se koristiti.

Četvrto, meta-analiza omogućava organizirani način postupanja sa informacijama velikog broja rezultata istraživanja. Kad broj istraživanja ili količina informacija koja se uzima iz tih istraživanja prijeđe određenu svotu, bilježenje ili obrađivanje na konvencionalni način ne može učinkovito voditi računa o svim detaljima. Sistematska obrada u meta-analizi i računalne baze podataka, suprotno prethodnome, imaju skoro neograničenu mogućnost bilježenja informacija iz svakog istraživanja i pokrivanja velikog broja takvih istraživanja. Npr. meta-analiza koju je proveo Lipsey 1992. (prema [LIPS2001]) rezultirala je bazom podataka većom od 150 vrsta podataka za svako od 500 istraživanja. Međutim, meta-analiza ne zahtjeva nužno veliki broj istraživanja i u nekim okolnostima može biti korisno primjenjena i na mali broj rezultata istraživanja, kao što su 2-3 istraživanja.

Meta-analiza ima i svojih nedostataka, pa je predmetom oštrih kritika.

Jedan od nedostataka meta-analize je jednostavno količina uloženog truda i stručnosti koje zahtjeva. Dobro izvedena, meta-analiza koja uključuje više od neke određene količine rezultata istraživanja je intenzivan rad koji uzima mnogo više vremena nego konvencionalni kvalitativni pregledi istraživanja.

Povrh toga, mnogi aspekti meta-analize zahtijevaju specijalizirano znanje, posebno odabir i izračun prikladnih veličina učinka i primjena statističke analize na njih.

Možda najveća kritika meta-analize, odnosi se na integriranje različitih istraživanja, tako i onih s neusporedivim rezultatima, kako to analitičari nazivaju, problem jabuka i naranči. Kritike govore da aritmetička sredina veličine učinka i slične metode meta-analize nisu smislene, ako su spojeni rezultati istraživanja bez zajedničke mjerne jedinice. Taj problem nastaje ako se različiti tipovi istraživanja spoje zajedno u jednu veliku prosječnu veličinu učinka. Međutim, meta-analitičari koji koriste široki krug tema, češće koriste usporedbu nego spajanje. A kada koriste različite podkategorije rezultata u meta-analizi, analitičari ih mogu razbiti pojedinačno, pa se distribucije veličine učinka i time povezani izračuni, mogu prikazati posebno za svakog, dopuštajući usporedbe među njima.

Međutim ovaj će problem uvijek ostati prisutan jer je definicija rezultata istraživanja koji su međusobno usporedivi u meta-analizi ovisna o analitičaru. Naime, rezultati koji nekom analitičaru izgledaju kategorički različiti, nekom se drugom mogu činiti srodni. Tako je meta-analiza koju je proveo Glass 1977.g. (Smith i Glass prema [LIPS2001]), a koja je proučavala učinke psihoterapije, bila predmetom kritika zbog kombiniranja rezultata iz potpuno različitih terapija. To su primjerice kognitivno bihevioristička, psihodinamična, gestalt, itd. Glass je tvrdio da je njegov cilj bio ispitivanje ukupne učinkovitosti različitih metoda psihoterapije, zato ih je uključio u istu meta-analizu. Drugi analitičari možda imaju uži pogled na temu meta-analize, pa uključuju rezultate samo jedne te iste terapije. U svakom slučaju, bitno je da analitičar ima definiciju domene koja ga zanima i logičku podlogu na temelju koje uključuje ili isključuje istraživanja za svoju meta-analizu. Drugi mogu kritizirati tu domenu i odabir istraživanja, ali dok su oni eksplicitno obrazloženi, svaki kritičar posebice može suditi je su li smisleni.

3. PRIMJERI VREDNOVANIH SUSTAVA E-UČENJA

U svrhu postizanja što boljeg rezultata u učenju i poučavanju, iz dana u dan se teži boljim i naprednijim metodama koje bi olakšale taj proces i učinile ga učinkovitijim. Tradicionalna nastava, koja se u praksi provodi pomoću jednog učitelja koji poučava 20-30 učenika, ne daje zadovoljavajuće rezultate. Takav proces poučavanja ne može učinkovito voditi računa o svakom učeniku pojedinačno. S druge strane, individualizirana nastava koja osigurava svakom učeniku individualnog učitelja, zasigurno bi davala dobre rezultate ali je u praksi neostvariva. Razlozi su očiti; visoki troškovi takvog sustava obrazovanja, manjak učitelja-tutora itd.

Zbog visoke učinkovitosti individualnog načina poučavanja, koje je pokazalo učinak od 2 sigma [BLOO1984], nastoji ga se uvesti u nastavu na takav način koji ne zahtijeva nužno prisutnost učitelja za svakog učenika posebice. To se postiže primjenom informacijske i komunikacijske tehnologije. Svakodnevno se dizajniraju i kreiraju sustavi koji integriraju različite pogodnosti koje navedena tehnologija pruža, od *web-temeljenog poučavanja* (eng. web-based learning), *računalno-temeljenog učenja* (eng. computer based learning), *digitalnih medija* (eng. digital collaboration) do *virtualnih učionica* (eng. virtual classrooms). Svi navedeni oblici poučavanja čine *e-učenje* (eng. e-learning) [LEAR2006], koje u novije vrijeme uključuje i primjenu mobilne tehnologije (mobilnih uređaja, PDA uređaja te MP3 player-a) [WIKI2006].

Učinkovitost individualiziranog poučavanja, se objašnjava utjecajem dva faktora. Prvo, ako se mjeri broj pitanja koja su postavljena i odgovorena, učenici i tutori naprave mnogo više *instrukcijskih interakcija* (eng. instructional interaction) po jedinici vremena nego što je to moguće u učionici. Naime, provedeno istraživanje o količini tih instrukcijskih interakcija (Graesser i Person prema [FLET2005]) pokazalo je da cijeli razred u učionici postavi prosječno 3 pitanja u jednom satu, što iznosi oko 0.11 pitanja po jednom učeniku. Suprotno tome, učenik u individualnoj sesiji s tutorom postavi oko 20-30 pitanja po satu, a tutor zahtjeva odgovor na njih 117-146. Ispitivanje intenziteta interakcije, koja se dogodi primjenom nekog sustava e-poučavanja pokazuje još aktivniji stupanj interakcije od navedenog. Drugo, individualizirano poučavanje smanjuje vrijeme potrebno za učenje tako što prilagođava poučavanje sposobnostima učenika. Manje vremena se utroši na materiju koja je učeniku već poznata ili koju je već savladao, i poučavanje se usmjeri na novo, nepoznato gradivo. Istraživanja su pokazala da korištenje nekog od sustava e-učenja prosječno uštedi 30% vremena (Orlansky i String prema [FLET2005]).

Osim navedenog, uspješnost sustava e-učenja je dijelom posljedica pozitivnog mišljenja učenika i studenata o takvim sustavima. Naime, u provedenim istraživanjima [COTT1991] oni su svojim riječima trebali opisati zbog čega osobno preferiraju učenje pomoću nekog od sustava e-učenja pred tradicionalnim načinom. Učenici i studenti su u korist sustava naveli sljedeće: da su beskrajno strpljivi, nisu nikad umorni, da se nikad ne uznemire niti se naljute, dopuštaju učenicima privatnost u učenju, nikad ne zaboravljaju ispravak i pohvalu, zabavni su i zanimljivi, individualiziraju učenje, omogućavaju učeniku učenje vlastitim tempom, ne omalovažavaju učenika kad pogriješi, u trenutku pružaju povratnu informaciju, objektivniji su

od nastavnika i profesora, a njima ostavljaju vremena za kvalitetniji kontakt sa učenicima (za razgovor, davanje savjeta, pomoć u rješavanju različitih problema vezanih i nevezanih uz nastavu...), nemaju predrasude o rasnoj i etničkoj pripadnosti, dobri su motivatori, poučavaju u malim pomacima, a rade izuzetnom brzinom - sličnom brzini ljudskih misli, itd...

E-učenje je danas od iznimne važnosti za proces poučavanja zbog potrebe permanentnog obrazovanja. Moderno doba ne dopušta završetak obrazovanja izlaskom iz škole ili fakulteta. Napredak u svim granama znanosti i u tehnologiji, zahtjeva od pojedinca praćenje napretka, novih spoznaja i svakodnevno usavršavanje. Zbog činjenice da e-učenje omogućava pristup izvorima za učenje u bilo koje vrijeme i bilo gdje, učenje je omogućeno svima, pa tako i zaposlenima, majkama ili učenicima koji žive na udaljenim mjestima, ili im je zbog nekog drugog razloga otežan svakodnevni odlazak u školu i fakultet. Takav oblik e-učenja još se naziva *učenje na daljinu* (eng. distance learning), a izvori učenja, između ostalog, su dostupni i putem blog-ova, foruma, chat-ova, meeting-a itd [WIKI2006].

3.1. PRIMJERI SUSTAVA E-UČENJA

Sustavi e-učenja zamišljeni su kao osobni učitelji za svakog učenika, a poučavanje je poduprto spomenutom tehnologijom. Njihova je zadaća uzimanje u obzir sposobnost i znanje svakog učenika, i ovisno o tome, provođenje poučavanja koje će biti prilagođeno svakom individualnom učeniku. Iako još nisu postigli učinak od 2-sigma, stručnjaci se slažu da njihova učinkovitost raste, te će isto tako rasti i njihova uloga u procesu učenja i poučavanja.

U nastavku slijede primjeri takvih sustava. Koliko su ti sustavi zaista uspjeli u svojoj osnovnoj namjeni, ovisilo je o tome koliko su dobro osmišljeni i dizajnirani. Primjeri sustava koji će ovdje biti opisani su vrednovani, odnosno, provedena su istraživanja o tome koliko je svaki od tih sustava doista pomogao učenicima u savladavanju gradiva. A odabir potencijalnih sustava, koji će tu biti opisani, je bio usmjeren na ona koja daju dovoljno podataka za izračun veličine učinka.

3.1.1. LEARNING BY ASKING - LBA

Learning by asking (LBA) [ZHAN2006] je sustav e-učenja s integriranim multimedijalnim sadržajima. Nastavni materijal uključuje interaktivni video, prezentacije i bilješke. LBA sustav pruža visok nivo interakcije omogućavajući studentima direktan pristup dijelovima videa i tako ponavljanje težeg ili nejasnog nastavnog sadržaja, preskakanje lakšeg ili već naučenog itd.

Provedeno je istraživanje kojim se htjelo ispitati da li interaktivni video povećava studentovo razumijevanje i povećava učinak učenja. U eksperimentu je sudjelovalo 138 studenata, koji su slučajnim odabirom svrstani u jednu od četiri grupe, kojima su pridružena različita okruženja učenja, kao što je prikazano (Tablica 3.1.).

Tablica 3.1. Grupe u eksperimentu

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

Studenti u grupama 1, 2 i 3 imali su isti scenarij. Najprije su objašnjeni ciljevi i procedure eksperimenta, te se pristupilo pisanju inicijalnog testa koji se sastojao od pitanja višestrukog izbora i da-ne pitanja. Studenti koji su trebali koristiti LBA sustav, prisustvovali su 5-minutnoj demonstraciji sustava, te su 5 minuta sami radili na sustavu. Nakon što su svi sudionici upoznali rad sustava, pristupili su učenju na sustavu u trajanju od 50 minuta, od kojih je video trajao 29 minuta. Bila je prikazana i prezentacija od 20 prikaza. Studenti tradicionalne nastave su slušali predavanja u trajanju od 50 minuta. Na kraju su svi pisali završni test. Testovi su bodovani na ljestvici 0-50 bodova.

U Tablica 3.2. su prikazani rezultati eksperimenta.

Tablica 3.2. Prikaz rezultata po grupama, aritmetičke sredine i standardne devijacije

Grupe	Aritmetičke sredine	Standardne devijacije
LBA sustav + interaktivni video (1)	34,1	8,87
LBA sustav + neinteraktivni video (2)	27,7	8,85
LBA sustav bez videa (3)	26,7	10,02
Tradicionalna nastava (4)	23,7	8,79

Grupa 2 i Grupa 3 (poučavanje s neinteraktivnim videom i bez videa) su ostvarile podjednak uspjeh na završnom testu, odnosno, između njihovih rezultata nije pronađena statistički značajna razlika, što govori u prilog tome da je interaktivni video taj koji pomaže studentima u boljem razumijevanju nastavnog sadržaja.

U daljnjoj analizi, tj. izračunu veličine učinka, Grupa 4 će se smatrati kontrolnom, jer je u njoj provedena nastava na tradicionalni način (slušanje predavanja), a Grupa 1 eksperimentalnom, odnosno, nastava se provodila pomoću spomenutog sustava i interaktivnog videa.

3.1.2. CONCEPTUAL HELPER

Conceptual Helper [ALBA2000] je inteligentni tutorski sustav za učenje i poučavanje fizike. Poučavanje se provodi pomoću domaćih zadaća orijentiranih na rješavanje kvalitativnih problema, takvih koje ne zahtijevaju upotrebu automatiziranih algebarskih izračuna, nego traže primjenu konceptualnog znanja. Koristi model traganja uz primjenu probabilističkih procjena za ispravljanje grešaka.

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

Iz rezultata inicijalnog testa utvrdila se ekvivalentnost kontrolne i eksperimentalne grupe, odnosno između tih grupa nije pronađena statistički značajna razlika. Naime, aritmetička sredina inicijalnog testa za kontrolnu grupu iznosila je 33.7 sa standardnom devijacijom 7.47, dok je aritmetička sredina inicijalnog testa za eksperimentalnu grupu iznosila 31.36 sa standardnom devijacijom 8.14. Zatim se pristupilo usporedbi razlika rezultata inicijalnog i završnog testa. Aritmetička sredina kontrolne grupe je 4.12 sa standardnom devijacijom 5.33. Aritmetička sredina eksperimentalne grupe je 7.47 sa standardnom devijacijom 5.03.

3.1.3. SQL- TUTOR

SQL-Tutor [MITR1999] je inteligentni tutorski sustav koji pomaže studentima u formuliranju upita u Structured Query Language. SQL-Tutor je dizajniran kao nadopuna tradicionalnoj nastavi i podrazumijeva da su se studenti već upoznali s osnovnim konceptom o bazama podataka, relacijskim modelom te osnovama SQL jezika.

Eksperiment je proveden sa studentima zadnje godine studija na Sveučilištu u Canterbury u Novom Zelandu. Oni su poslušali 6 predavanja o SQL-u i potrošili oko 8 sati na vježbanje definiranja upita. Eksperimentalnu grupu je činilo 20 studenata koji su radili sa sustavom SQL-Tutor dva sata. Sustav je provjeravao riješene probleme, a zatim bi kao povratnu informaciju naveo broj grešaka i specificirao samo jednu od njih, tako da student ispravlja jednu po jednu grešku. Koristio se takav oblik povratne informacije, jer se polazilo od činjenice da više navedenih grešaka odjednom, nepotrebno zbunjuje studente. Kontrolnu grupu je činilo 26 studenata.

Test za provjeru učinkovitosti je proveden dva tjedna nakon rada sa sustavom. Trebalo je riješiti 6 problema za 90 minuta. Test je bodovan na ljestvici 0-100. Aritmetička sredina rezultata eksperimentalne grupe je 82.7, a kontrolne grupe 71.2. Zajednička standardna devijacija je 15.4.

3.1.4. PSYCAL

PsyCAL [BUCH2000] (Psychology Computer Assisted Learning) je oblik web-temeljenog sustava ocjenjivanja, koje je primjenjeno kao pomoć u savladavanju gradiva iz kolegija Uvod u psihologiju.

Studenti su pristupali sustavu on-line, i riješavali PsyCAL vježbe. Nakon rješavanja, student bi dobio povratnu informaciju, koja je ukazivala na to koja područja je student slabo naučio. Važna karakteristika PsyCAL-a je ta, da nije otkrivao točne odgovore na netočno riješene zadatke. Umjesto toga, za svaki netočan odgovor, sustav je davao reference (na određeni dio udžbenika ili više njih) i tako uputio studenta gdje da pronađe točan odgovor. Tako je sustav na neki način "obvezao" studenta da koristi zadanu literaturu. Studentu se savjetovalo da ponovi test nakon što utvrdi gradivo koje mu je sustav pokazao kao nedovoljno naučeno, i tako ponavlja kružni postupak test-učenje-ponovno testiranje, sve dok ne savlada gradivo u potpunosti. Prilikom korištenja sustava, student je upisivao svoj identifikacijski broj, pa se

tako vodila evidencija o studentima koji su pristupali sustavu. Međutim, sustavu se moglo pristupiti i bez unošenja identifikacijskog broja.

Da bi se utvrdila učinkovitost ovog sustava, proveden je niz analiza, od kojih je za ovo razmatranje najvažnija usporedba uspjeha na završnom testu između studenata koji su koristili PsyCAL s onima koji nisu. Sa sigurnošću se utvrdio broj od 148 studenata koji su koristili sustav i 71 student za kojeg nema podataka o korištenju, (neki od studenata su možda anonimno koristili sustav). U svrhu analize, pretpostavilo se da tih 71 studenata nije koristilo sustav PsyCAL.

Prva grupa (korisnici sustava) na završnom testu su ostvarili uspjeh od 48.42, (standardna devijacija 12.32), a druga grupa je imala 39.75, (standardna devijacija 1.3).

3.1.5. COMPUTER TRAINING OF PHONOLOGICAL AWARENESS IN KINDERGARTEN

Computer training of phonological awareness in kindergarten [SEGE2005] ili računalno-temeljeno poučavanje fonološke svijesti, provedeno je u djece predškolske dobi u Nizozemskoj. Specijalno su dizajnirani programi u obliku različitih igara, kroz koje djeca uče određene vještine. Istraživanje, koje je provedeno, proučava dugotrajne učinke tog poučavanja na razvoj fonološke svijesti u djece predškolske dobi. Fonološka svijest se opisuje kao svijest o rimi, slogovima, početnim dijelovima riječi, završnim dijelovima riječi itd. Dokazano je da razvoj djetetove fonološke svijesti u predškolskoj dobi, kasnije olakšava proces učenja čitanja i pisanja.

Eksperimentalni program je činilo 5 edukacijskih CD-ROM-ova specijalno dizajniranih za predškolski uzrast. Svaki CD-ROM sadrži 9 različitih igara. Djeca mogu odabrati igru koju žele igrati, međutim nisu sve igre dostupne na samom početku poučavanja. Kasnije, kad dijete stekne određene sposobnosti, postoji mogućnost odabira i naprednijih igara. One su generalno podijeljene u *igre otkrivanja* (eng. discovery games) i *igre poučavanja* (eng. learning games). Kroz igre otkrivanja, dok traje pjesma na ekranu se pojavljuje tekst u karaoke stilu, pa dijete otkriva slova, riječi, rečenice, smjer pisanja (s lijeva na desno) itd. Zatim otkriva boje, tako da dobije zadatak obojiti predmet onom bojom koja je zapisana ispod tog predmeta. Zatim može izrezivati riječi i slagati, crtati i praviti razglednicu. Igre poučavanja započinju igrama s rimom. Dijete mora završiti danu rečenicu odabirući određenu sličicu predmeta čiji se naziv rimuje sa ostatkom rečenice. Težina igre se automatski povećava kad dijete pokaže određenu lakoću u rješavanju. Zatim slijede igre sa slovima, povezivanje sličice i slova, kojim počinje predmet na toj slici. Program stvara i "popis slova" koje dijete poznaje, a kad nauči dovoljno slova za slaganje barem 5-6 riječi, tada se stvara i "popis riječi" koje dijete zna "napisati". Kroz 40 tjedana, djeca su radila s računalnim igrama jednom tjedno po 15 minuta, a zadnjih 10 tjedana, djeca su mogla igrati računalne igre najviše 3 puta tjedno.

Djeca u kontrolnoj grupi su povremeno radila sa računalom i nekim programima, ali ne sa ovim programom po kojem je radila eksperimentalna grupa. Prvenstveno su koristili zabavne programe kao i programe koji su uključivali priče i igre sa oblicima i bojama.

Inicijalno testiranje je obavljeno na početku korištenja sustava i nije pokazalo značajnu razliku između grupa. Testiranje je vršeno još par puta: na pola poučavanja, na kraju poučavanja te u prvom razredu osnovne škole, nakon 4 mjeseca formalnog učenja čitanja. Da bi se provjerilo koliko je poučavanje utjecalo na djecu i na njihov uspjeh u savladavanju

gradiva, koristiti će se rezultati testiranja u prvom razredu, nakon što su djeca 4 mjeseca formalno učila čitati. Takav test, koji se provodi nekoliko mjeseci nakon što je završeno poučavanje, se još naziva testom *retencije* ili zadržavanja (eng. retention test), a posebno su bodovani njegovi segmenti kako je prikazano (Tablica 3.3.).

Tablica 3.3. Rezultati različitih segmenata testa: aritmetička sredina, standardna devijacija i broj djece

Retention test	Kontrolna grupa			Eksperimentalna grupa		
	Arit.sredina	St.devijacija	N	Arit.sredina	St.devijacija	N
Phonemic segmentation	14,11	4,76	44	13,26	4,45	34
Auditory blending	15,88	3,81	44	17,43	3,04	34
Grapheme knowledge	32,64	1,97	44	33,76	0,55	34
Decoding (no. of words/min)	17,93	7,39	44	20,38	10,47	34

U tablici je prikazan prosječan broj bodova za svaki dio testa, pa će ukupna veličina učinka biti prosjek četiri dobivene veličine učinka.

3.1.6. KNOWLEDGE MAPPING

Knowledge mapping [OSMU1999] ili mape znanja, su definirane kao dijagramski prikaz glavnih semantičkih veza između skupa pojmova. Čvorovi (eng. nodes) predstavljaju pojmove a linije između njih pokazuju vezu između tih pojmova. Uvođenje tehnologije u ovakav sustav mapi znanja, daje nove mogućnosti poučavanju i provjeravanju znanja na taj način što su podaci pohranjeni u obliku baze podataka, i u svakom trenutku im se može pristupiti. Zatim, omogućavaju odgovor u kratkom vremenu, brzu povratnu informaciju te na taj način povećavaju interakciju s učenikom. Cilj primjene ovog sustava je istraživanje kako računalno temeljene mape znanja mogu biti korištene istovremeno kao alat za poučavanje i kao alat za procjenu učenikova razumijevanja pojmova iz znanosti.

Sudionici u istraživanju su učenici 4. i 5. razreda osnovne škole, koji su nasumice raspoređeni u kontrolnu i eksperimentalnu grupu. Podaci su se skupljali tijekom 6 tjedana. Kao inicijalni test, svi učenici su na početku dobili zadatak da kreiraju individualnu mapu znanja. U nastavku eksperimenta, učenici su dobili uvid u te mape i to im je bila početna točka za međusobnu raspravu. Razlika u poučavanju između kontrolne i eksperimentalne grupe je bila kroz 2., 3. i 4. tjedan. Kod eksperimentalne, male grupe učenika su radile zajedno na stvaranju mape znanja ljudskog tijela, online, i to jednom na tjedan. Svaka od tih grupica je dobivala povratnu informaciju svaki tjedan, o tome koliko je kvalitetno do tad napravljena njihova mapa (smisleni čvorovi i veze). Kontrolna grupa je koristila neke druge resurse za istraživanje ljudskog tijela. Završni test se sastojao od dva dijela: kreiranje mape znanja i pisanje eseja o dišnom, probavnom i krvožilnom sustavu. Završni test je pisalo 22 učenika kontrolne i 20 učenika eksperimentalne grupe.

Zbog činjenice da je eksperimentalna grupa kroz postupak učenja i poučavanja radila na izradi i poboljšanju mapa znanja, uspjeh je, kako je i očekivano na završnom testu vezanom uz izradu mape znanja bio veći nego kod kontrolne grupe. Zato će se analiza rezultata temeljiti na rezultatima eseja. U esej je učenik trebao opširno izraziti sve što je naučio o navedenim sustavima u ljudskom tijelu, na koji način funkcioniraju i kakvo je njihovo međudjelovanje tj. na koji način su sustavi međusobno povezani. Uspješnija eksperimentalna grupa u eseju, je pokazatelj da navedeni sustav kroz mape znanja pomaže učenicima u boljem razumijevanju gradiva i povezivanju važnih pojmova. Zato će se u daljnjoj analizi koristiti samo rezultati tog eseja koji najbolje pokazuje kvalitetu učenikova znanja (Tablica 3.4.).

Tablica 3.4. Rezultati eseja o poznavanju dišnog, probavnog i krvožilnog sustava

	Kontrolna grupa		Eksperimentalna grupa	
	Arit.sredina	St.devijacija	Arit.sredina	St.devijacija
Ocjena cjelokupnog eseja	2,89	0,89	3,12	0,84
Ocjena međupovezivanja sustava	1,83	1,13	2,48	1,24

3.1.7. SHERLOCK

Znanstvenici na Air Force Armstrong Laboratory, Human Resources Directorate [GOTT1995], su razvili inteligentni tutorski sustav za otklanjanje kvarova (eng. troubleshooting) u zrakoplovstvu i nazvali ga Sherlock. Taj tutor predstavlja zapravo simulaciju okruženja u kakvom stručnjaci zrakoplovne tehnologije rade na ispitivanju ispravnosti te uklanjanju kvarova na djelovima zrakoplova. Njihova je zadaća, u ovoj domeni, popravak i održavanje elektroničkih podsustava i opreme zrakoplova F15.

Tutor omogućava da učenici *uče na način da sami rade* (eng. learning by doing) na elektroničkom otklanjanju kvarova. Najprije trebaju pronaći kvar, a to se provodi u računalnom okruženju koje je nastavak njihovog učenja u realnom svijetu. Ako student pogriješi u tom procesu rješavanja problema, sustav nudi objašnjenja, natuknice, podsjetnike na gradivo itd. Sherlock osigurava i napredniji sustav učenja koji se naziva *učenje kroz odraz* (eng. learning through reflection). To se provodi na način da se studentu nakon rješavanja problema prikažu postupci i odluke koje je proveo u tom procesu. Zatim se to uspoređi sa koracima koje bi proveo ekspert, i omogući studentu procjenu svog načina rješavanja. Ta ocjena se uspoređi sa ocjenom koju daje sustav. Navedeni način predstavlja izuzetan napredak u poučavanju jer rezultira sposobnošću *samo-ispravljanja* (eng. self-correction) i *samo-nadziranja* (eng. self-monitoring), što su izuzetne metakognitivne vještine nekog studenta.

Ovaj sustav, uz navedene, ima mnogo drugih prednosti pred tradicionalnom nastavom i praksom u stvarnoj radionici. Najveća prednost je u tome što treba prosječno 8-10 godina rada u stvarnoj radionici kako bi se došlo u kontakt sa velikim brojem kvarova. Nasuprot tome, Sherlock integrira sve moguće kvarove i njihovo otklanjanje u svoj sustav poučavanja kao probleme koje treba riješiti. Nadalje, ovaj sustav omogućuje studentu da naučeni teorijski dio primjeni na praktične zadatke, vršeći ispitivanja i otklanjajući kvarove samostalno, što nije slučaj kod tradicionalne nastave, koja nakon teorijske nastave predviđa praksu u radionicama gdje studenti mogu samo promatrati stručnjake u radu.

54 *naučnika* (eng. apprentice), *pomoćnika* (eng. journeyman) i *tehničkih stručnjaka* (eng. master) za zrakoplove, sudjelovalo je u eksperimentu za vrednovanje ovog sustava, koje se provodilo u 3 geografski odvojena mjesta: Langley, Eglin i Nellis. Naučnici i pomoćnici su imali prosječno 33 mjeseca iskustva u tom poslu, i u svrhu eksperimenta su raspodijeljeni u kontrolnu i eksperimentalnu grupu. Na inicijalnom testu, grupe su pokazale slične rezultate. Stručnjaci (eksperti) su imali prosječno 124 mjeseca iskustva, a njihovi rezultati su korišteni samo za usporedbu s navedene dvije grupe. Kreirana su dva instrumenta su za vrednovanje: VTT (eng. Verbal Troubleshooting Test) ili verbalni test otklanjanja kvarova i NIT (eng. Noninteractive Troubleshooting Test) ili neinteraktivni test otklanjanja kvarova. Rezultati su prikazani u Tablica 3.5.

Tablica 3.5. Prikaz rezultata na testovima i usporedba uspjeha sa iskusnim stručnjacima

Testovi	Naučnici i pomoćnici						Stručnjaci		
	Kontrolna			Eksperimentalna					
	Arit.sredina	St.devijacija	N	Arit.sredina	St.devijacija	N	Arit.sredina	St.devijacija	N
VTT3	59	37	23	95	5	18	85	12	13
VTT4	58	37	23	91	7	18	86	11	13
NIT	75	14	23	87	12	18	86	11	13

Rezultati pokazuju da sposobnosti početnika u zrakoplovnoj tehnologiji, primjenom navednog tutora, mogu konkurirati sposobnostima iskusnog stručnjaka u zadanom području. Naučnici u relativno kratkom vremenu postaju stručnjaci u otklanjanju kvarova, dok bi tradicionalnim načinom poučavanja trebala višegodišnja praksa da se stekne taj nivo uspješnosti.

3.1.8. ONLINE GENERAL COLLEGE CHEMISTRY COURSE

Online general college chemistry course [CASA2001] je sustav e-učenja i poučavanja kolegija Opća kemija koji kombinira male grupe studenata, tjedne kontakte i web-temeljene materijale. Kemija je grana znanosti čije je poučavanje vezano uz praktične vježbe u laboratoriju i mnogi temeljni koncepti su formulirani iz promatranja i eksperimenata napravljenih u laboratoriju. Zato je spomenuti sustav hibridni model koji ujedinjuje materijale za učenje dostupne preko interneta i 3 sata praktičnih vježbi u laboratoriju jednom tjedno.

Studenti koji su činili kontrolnu grupu, pohađali su tradicionalnu nastavu tj. imali su predavanja iz tog kolegija dva puta na tjedan, u trajanju od 75 minuta, i vježbe u laboratoriju 3 sata na tjedan. Kontrolnu grupu je činilo 117 studenata četverogodišnjeg studija na sveučilištu University of North Carolina (UNC). Na Cape Fear Community College (CFCC) korišten je model e-poučavanja. Eksperimentalnu grupu je činilo 25 studenata koji su dobivali materijale za učenje na daljinu, online, koji su uključivali animaciju i video sadržaje. Uz to su dobivali napomene za učenje za naredni tjedan, koje su trebale pomoći studentima da budu u tijeku, npr. preporučene lekcije koje treba proučiti, aktivnosti, tekst koji treba pročitati, probleme koje treba riješiti. Jednom tjedno studenti su imali vježbe u laboratoriju u trajanju od 3 sata. To se vrijeme koristilo i za ponavljanje zadataka iz prethodnog tjedna, i pisanje "blic" testa da bi se provjerilo da li studenti uistinu prate zadani program. Takvi testovi su bili dostupni i preko web-a, a studenti bi tada dobivali povratnu informaciju tj. točne odgovore i objašnjenja uz njih. Uz navedeno, mogli su komunicirati e-mailom sa svojim profesorom, slati riješene domaće zadaće itd. Korišten je isti udžbenik kao i kod tradicionalne nastave, a svi studenti su na kraju poučavanja pisali isti završni test.

Rezultati završnog testa pokazali su uspješniju eksperimentalnu grupu, sa aritmetičkom sredinom 75.84, varijanca 204.973. Aritmetička sredina kontrolne grupe na završnom testu je iznosila 61.98, varijanca 269.603. Zajednička varijanca 258.524. Nedostatak ovog eksperimenta je u tome što nije proveden inicijalni test za utvrđivanje ekvivalentnosti tih dviju grupa prije poučavanja.

3.1.9. DISTANCE LEARNING: BUSINESS COMMUNICATIONS CLASS

Na sveučilištu North Carolina je razvijen sustav e-učenja, i to učenja na daljinu za kolegij Business communications [TUCK2001], koji je dizajniran za razvijanje razumjevanja potrebe za učinkovitim načinima komunikacije u poslovanju.

Sudionici istraživanja koje je provedeno, su studenti koji su upisali kolegij Business communications. To su bili studenti različitih studija (društvenih, prirodnih, tehničkih), i različitih godina studija. Tradicionalnu nastavu je pohađalo 23 studenta, koji su imali predavanja jednom tjedno u trajanju od 2 sata. Za 24 studenta eksperimentalne grupe, materijali za učenje na daljinu su također bili poslani jednom na tjedan. Trebali su u tjedan dana proučiti navedene materijale, i u tom se roku "logirati" i ispuniti dodjeljene zadatke. Bilješke za učenje su bile i u formi audio i video linkova, pomoću kojih je studentu omogućeno slušanje i gledanje profesora dok objašnjava segmente tog kolegija. Uključivale su i tekstualne linkove, i prezentacije u PowerPoint-u. Kontrolna grupa je također na raspolaganju imala računala, folije i prezentacije u PowerPoint-u. Obje grupe su imale istog profesora, iste domaće uratke, isti inicijalni i završni test. Isti kriterij ocjenjivanja je vrijedio za obje grupe. Svaki student (kontrolne i eksperimentalne grupe) je mogao kontaktirati profesora e-mail-om, telefonom, faksom ili osobno u vrijeme konzultacija. Provedeno je više testova; inicijalni i završni test je proveden kao ocjena studentovog znanja osnovnih koncepata važnih za komunikaciju u poslovanju. Ocjenjeni su i domaći radovi, koji su pisani na kraju svake cjeline. Svaki student je trebao napraviti i napisati israživački projekt o međunarodnom poslovanju. U tu svrhu, student je trebao odabrati neku državu (osim SAD-a) i napisati izvještaj o kulturi te države i načinu na koji bi trebalo uspješno voditi poslovanje u toj državi. Studenti su pisali isti završni ispit znanja, koji je zahtijevao primjenu naučenog. Tablica 3.6. sadrži uspjehe u navedenim testovima, radovima i završnu ocjenu kolegija.

Tablica 3.6. Prikaz uspjeha dvije grupe različitim modelima ocjenjivanja

	Kontrolna grupa			Eksperimentalna grupa		
	Arit.sredina	St.devijacija	N	Arit.sredina	St.devijacija	N
Završni test	65,55	10,91	23	72,43	9,12	24
Cjelokupni ispit znanja	78,26	12,63	23	85,92	8,16	24
Istraživački izvještaj	87,45	28,6	23	91,39	12,32	24
Završna ocjena kolegija	80,57	16,16	23	85,42	13,11	24

Svi navedeni testovi i ocjene su podjednako važne za utvrđivanje učinkovitosti ovog načina poučavanja, a veličina uzorka je stalna, pa će ukupna veličina učinka biti izračunata kao prosjek veličina učinka dobivenih od sva četiri modela ocjenjivanja.

3.1.10. COMPUTER BASED INSTRUCTION IN AN ENGINEERING CLASS

U određivanju učinka računalno-temeljenog poučavanja (CBI) na inženjering, na sveučilištu North Carolina State University, dizajnirane su CBI jedinice (eng. units) za poučavanje kolegija "Civil Engineering: Traffic Engineering" [HUMM1996]. To je obvezni kolegij za sve studente civilnog inženjeringa. Studenti su radili na jedinicama sustava, koje su pisane u programskom jeziku CT, na Unix-temeljenoj radnoj stanici (eng. workstation). Jedinice su dizajnirane tako da rad na svakoj zahtjeva otprilike jedan sat, a uključivale su animacije, hipertekst (eng. hypertext) i nudile povratnu informaciju na uspjeh u "blic" testu.

Studenti kolegija su slučajnim odabirom raspoređeni u eksperimentalnu i kontrolnu grupu, i slušali kolegij pomoću CBI jedinica, odnosno, ekvivalentnih predavanja u učionici. Inicijalni test je pokazao statističku ekvivalentnost tih grupa, iako je kontrolna grupa pokazala bolje rezultate. Tijekom poučavanja, studentima je preporučeno da ne dijele bilješke sa studentima druge grupe, te da je pristup CBI jedinici ograničen na studente eksperimentalne grupe. Provedena su dva testiranja. Ocjenjivanje je vršio asistent na sveučilištu, koji nije pripadao timu koji je provodio eksperiment. Na rezultate prvog testiranja, kako je kasnije pokazano, utjecalo je ne razumijevanje pitanja od strane studenata, pa se nije moglo sa sigurnošću utvrditi da su rezultati posljedica učinka određene metode poučavanja. Ti rezultati se neće uzeti u obzir. Drugo testiranje je provedeno dva tjedna nakon završenog poučavanja. Rezultati drugog testiranja su dani u Tablica 3.7. i pokazuju nešto bolje postignuće kontrolne grupe.

Tablica 3.7. Rezultati testiranja: aritmetičke sredine, varijance i broj studenata

	Kontrolna grupa			Eksperimentalna grupa		
	Arit.sredina	Varijanca	N	Arit.sredina	Varijanca	N
Test question 1	10,33	36,1	18	9,6	42,6	20
Test question 2	3,94	3,58	18	3,85	3,50	20

Rezultati nisu postignuli statističku značajnost, što implicira da je računalno temeljeno poučavanje pokazalo podjednaku učinkovitost kao i tradicionalna predavanja.

Iako nije postignuta statistička značajnost, rezultati će se koristiti za izračun veličine učinka, koja će imati negativan predznak zbog uspješnije kontrolne grupe. Zato treba napomenuti da je ovaj sustav dizajniran i testiranje je provedeno u akademskoj godini 1993/94, dakle starije je od 12 godina, te da ovaj eksperiment spada među prve provjere učinkovitosti CBI-poučavanja u inženjeringu, što donekle objašnjava njegovu slabiju učinkovitost.

4. META-ANALIZA REZULTATA VREDNOVANIH SUSTAVA

U ovom poglavlju će biti provedena meta-analiza rezultata vrednovanih sustava opisanih u prethodnom poglavlju. Uzrok malom broju sustava, koji će biti uključeni u meta-analizu, je nedostupnost takvih sustava, a pogotovo onih koji su vrednovani. Naime, cijeli postupak vrednovanja sustava zahtjeva dosta uloženog vremena, truda i stručnosti. Dodatni kriterij, koji je vrednovani sustav trebao zadovoljiti kako bi bio uključen u ovu meta-analizu, je generiranje svih potrebnih podataka za izračun veličine učinka i njezine težine.

4.1. *TEMA META-ANALIZE*

Tema meta-analize je ujedno i zajednička tema svih sustava zajedno. Postavlja se pitanje koliko zapravo ti sustavi doprinose boljem i kvalitetnijem procesu učenja i poučavanja. Opisani sustavi su različito dizajnirani i uključuju različite oblike e-učenja, od inteligentnih tutorskih sustava, sustava na web-temeljenom poučavanju, sustava sa interaktivnim video sadržajima, sustava za poučavanje na daljinu itd. Svaki od sustava je napravljen za poučavanje specifičnog nastavnog sadržaja ili kolegija: fizike, algebre, psihologije, zrakoplovne tehnologije itd., a njihovi korisnici su djeca, učenici i studenti različitih dobnih skupina. Uz tolike međusobne razlike, postavlja se pitanje kako ih se može sve uvrstiti u istu meta-analizu, i međusobno uspoređivati i analizirati?

Odgovor leži u svojstvu koje im je zajedničko. Naime, prema provedenim istraživanjima i eksperimentima svaki od tih sustava (manje ili više) pomaže u savladavanju gradiva i tako povećava učinak i kvalitetu učenja. Tema ove meta-analize je, dakle, utvrđivanje učinkovitosti takvih sustava u procesu učenja i poučavanja u odnosu na tradicionalni način poučavanja. Postavlja se pitanje: Kako ovakvi sustavi utječu na proces učenja i poučavanja? Da li učenik ima koristi od takvog poučavanja? I ako ima, kolika je ta učinkovitost u odnosu na tradicionalni oblik poučavanja?

Rezultat meta-analize je veličina učinka u jedinicama standardne devijacije, pa u tom kontekstu pitanje glasi: Koliko standardnih devijacija je prosječan ovako poučavan učenik, uspješniji od prosječnog učenika koji je poučavan tradicionalnim putem?

4.2. ODABIR PRIKLADNIH ISTRAŽIVANJA

U kolekciji dostupnih istraživanja, njih otprilike 60, koja su se bavila tematikom meta-analize, odnosno primjenom informacijske i komunikacijske tehnologije u sustav obrazovanja, izdvojeno je 10 istraživanja. Glavni kriterij za odabir istraživanja je uključivao ona, koja generiraju sve potrebne podatke za izračun veličine učinka. Uz taj kriterij, postojao je i dodatan kriterij koji je trebalo zadovoljiti. Naime, neka od tih 60 istraživanja bavila su se usporedbom tzv. neintelligentnih sustava sa intelligentnim sustavima, i sličnim usporedbama postojećih sustava sa poboljšanim verzijama tih istih sustava. Takva istraživanja, nisu uključena u ovu meta-analizu, jer nisu odgovarala na njena glavna pitanja. Za ovu meta-analizu, u obzir su uzeta samo istraživanja koja se bave proučavanjem sustava, koji uključuje neke oblike e-učenja i njegovu usporedbu sa tradicionalnim načinom poučavanja.

4.3. ODABIR METODE ZA IZRAČUN VELIČINE UČINKA

Pošto su svi vrednovani sustavi, opisani u prethodnom poglavlju, eksperimentalne strukture, odnosno, dvije grupe (eksperimentalna i kontrolna) se uspoređuju na nekoj zavisnoj varijabli (rezultati testa), izračunavanje veličine učinka vršit će se prema formulama za kontraste u grupi i to izračunom standardne razlike aritmetičkih sredina (formule (2.7) i (2.8)). Koristiti će se i ispravak veličine učinka, zbog relativno malog uzorka ispitanika u pojedinim istraživanjima. Za svaku individualnu veličinu učinka potrebno je izračunati i njezinu težinu, kao što je opisano u drugom poglavlju (formule (2.9)).

Tek kad se pronađu sve pojedinačne veličine učinka i njihove težine, prijeći će se na izračun ukupne veličine učinka (formula (2.22)).

4.4. IZRAČUN POJEDINAČNIH VELIČINA UČINKA I NJIHOVIH TEŽINA

Iz svakog vrednovanog sustava, najprije će biti izdvojeni podaci koji su potrebni za izračun veličine učinka i njene težine. To su aritmetička sredina postignutog uspjeha na završnom testu (ili na više njih) za eksperimentalnu i kontrolnu grupu, te standardna devijacija i broj ispitanika za te dvije grupe. Zatim slijedi sam izračun, na način da se najprije izračuna zajednička standardna devijacija (s_p), veličina učinka kao standardna razlika aritmetičkih sredina (ES_{sm}), ispravak veličine učinka (ES'_{sm}) i njena težina (w_{sm}).

Kod pojedinih sustava se, prilikom vrednovanja, koristilo više različitih završnih testova, od kojih je svaki bio osmišljen za provjeru nekog užeg područja obrađene cjeline. Zbog pretpostavke nezavisnih podataka za ovu analizu, nije moguće koristiti rezultate svih provedenih testova, jer bi to rezultiralo sa više veličina učinka dobivenih od istog uzorka. Stoga će se za izračun veličine učinka svakog sustava, uzimati rezultati samo jednog testa, i to onog za kojeg se smatra da je "najosjetljiviji", u smislu da najbolje ocjenjuje kvalitetu učenikova znanja, razumjevanje naučenog i uzročno-posljedično povezivanje pojmova. Ako se procijeni da su svi provedeni testovi podjednako važni za vrednovanje određenog sustava, ili ako su provedeni isti testovi više puta, tada je dopušteno za konačnu veličinu učinka tog sustava uzeti prosječnu vrijednost dobivenih veličina učinka iz tih rezultata (naravno, ako je u testiranju sudjelovao uvijek isti broj ispitanika).

4.4.1. LEARNING BY ASKING - LBA

Learning by asking (LBA) je sustav e-učenja s integriranom multimedijom.

U eksperimentu za vrednovanje ovog sustava je sudjelovalo ukupno 138 studenata. Međutim, za ocjenu učinkovitosti ovog sustava pred tradicionalnim načinom učenja i poučavanja najvažnije su bile razlike u postignutim rezultatima između Grupe 1 i Grupe 4 (Tablica 3.2.). Grupa 1 će se smatrati eksperimentalnom jer se poučavanje provodilo pomoću LBA sustava s interaktivnim videom, dok je u Grupi 4 je primjenjeno tradicionalno poučavanje, pa će ona služiti kao kontrolna grupa. U Tablica 4.1. su izdvojeni svi važni podaci iz navedenog istraživanja, uz napomenu da je N korišten u tablici, ukupan broj ispitanika, u ovom slučaju 69 tj. $N = n_E + n_K$.

Tablica 4.1. Podaci vrednovanja sustava LBA, te izračun veličine učinka i njene težine

	EKSPERIMENTALNA	KONTROLNA
\bar{X}	34,1	23,7
s	8,87	8,79
n	35	34
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	8,83	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	1,18	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	1,16	
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (ES'_{sm})^2}$	14,75	

Dobivena je veličina učinka ovog sustava $ES'_{sm} = 1.16$ i njena težina $w_{sm} = 14.75$. Ti će podaci biti kasnije korišteni u izračunu ukupne veličine učinka spomenutih sustava.

4.4.2. CONCEPTUAL HELPER

Conceptual Helper je inteligentni tutorski sustav za učenje i poučavanje fizike.

Iz provedenog procesa vrednovanja, nisu poznati rezultati završnog testa za eksperimentalnu i kontrolnu grupu. Umjesto toga, navedeni su prosjeci povećanja uspješnosti (eng. mean gain score) između završnog i inicijalnog testa za obje grupe. Ti podaci su izdvojeni u Tablica 4.2. i biti će korišteni za analizu i potrebne izračune.

Tablica 4.2. Podaci vrednovanja sustava Conceptual Helper, te izračun veličine učinka i njene težine

	EKSPERIMENTALNA	KONTROLNA
\bar{X}	7,47	4,12
s	5,03	5,33
n	22	20
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	5,17	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	0,65	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	0,64	
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (ES'_{sm})^2}$	9,97	

Dobivena je veličina učinka ovog sustava $ES'_{sm} = 0.64$ i njena težina $w_{sm} = 9.97$. Ti će podaci biti kasnije korišteni u izračunu ukupne veličine učinka spomenutih sustava.

4.4.3. SQL-TUTOR

SQL-Tutor je inteligentni tutorski sustav za pomoć u formuliranju upita u Structured Query Language.

Podaci o vrednovanju ovog sustava ne sadrže informacije o standardnim devijacijama na završnom testu za eksperimentalnu i kontrolnu grupu. Umjesto toga, navedena je zajednička standardna devijacija objiju grupa, koja je korištena pri izračunu (Tablica 4.3.).

Tablica 4.3. Podaci vrednovanja sustava SQL-Tutor, te izračun veličine učinka i njene težine

	EKSPERIMENTALNA	KONTROLNA
\bar{X}	82,7	71,2
s	Nema podataka	Nema podataka
n	20	26
s_p	15,4	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	0,75	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	0,73	
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (ES'_{sm})^2}$	10,6	

Dobivena je veličina učinka ovog sustava $ES'_{sm} = 0.73$ i njena težina $w_{sm} = 10.6$. Ti će podaci biti kasnije korišteni u izračunu ukupne veličine učinka spomenutih sustava.

4.4.4. PSYCAL

Psychology Computer-Assisted Learning (PsyCAL) je oblik web-temeljenog sustava ocjenjivanja, koje je primjenjeno kao pomoć u savladavanju gradiva iz kolegija Uvod u psihologiju.

Prilikom pristupanja sustavu, student je upisivao svoj identifikacijski broj. Sustav je tako vodio evidenciju o učenicima koji koriste sustav; koliko učenika ga koristi i koliko puta. Problem je u tome što se sustavu moglo pristupiti i bez unošenja identifikacijskog broja. Zato se za navedenih 71 studenata, koji su svrstani u kontrolnu grupu, zapravo ne može sa sigurnošću znati jesu li koristili sustav. Pretpostavilo se da nisu koristili sustav. Tablica 4.4. sadrži rezultate završnog testa.

Zbog prethodne pretpostavke, postavlja se pitanje da li je možda stvarna veličina učinka ovog sustava veća od ovdje dobivene. Odgovor je vjerojatno potvrđan. Naime, ako su neki od 71 studenata u kontrolnoj grupi anonimno koristili sustav i zbog toga bolje napisali završni test, pridonijeli su povećanju uspjeha kontrolne grupe, odnosno smanjenju veličine učinka spomenutog sustava. Međutim, u nedostatku preciznijih mjerenja, za daljnju obradu će se koristiti dobiveni rezultati.

Tablica 4.4. Podaci vrednovanja sustava PsyCAL, te izračun veličine učinka i njene težine

	EKSPERIMENTALNA	KONTROLNA
\bar{X}	48,42	39,75
s	12,32	1,3
n	148	71
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	10,17	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	0,85	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	0,85	
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (ES'_{sm})^2}$	44,46	

Dobivena je veličina učinka ovog sustava $ES'_{sm} = 0.85$ i njena težina $w_{sm} = 44.46$.

4.4.5. COMPUTER TRAINING OF PHONOLOGICAL AWARENESS IN KINDERGARTEN

Računalno-temeljeno poučavanje fonološke svijesti je provedeno u djece predškolske dobi u Nizozemskoj. Specijalno su dizajnirani programi u obliku različitih igara, kroz koje djeca uče određene vještine. Dokazano je da razvoj djetetove fonološke svijesti u predškolskoj dobi, kasnije olakšava proces učenja čitanja i pisanja.

Provedeno je više testiranja, uz inicijalni i završni test, proveden je i jedan test između tih spomenutih. Međutim, za provjeru učinka ovakvog poučavanja najpogodnije je testiranje koje se provodilo nekoliko mjeseci nakon završetka spomenutog poučavanja, kad su djeca već krenula u prvi razred i počela učiti čitati. Taj test provjerava zadržavanje naučenog i primjenu na novo gradivo, odnosno na čitanje (Tablica 3.3.). Zato su za izračunavanje veličine učinka ovog sustava, u Tablica 4.5. izdvojeni dobiveni rezultati tog testa.

Tablica 4.5. Izračun veličine učinka i njene težine iz rezultata testa

	Phonemic		Auditory		Grapheme		Decoding	
	Eksp	Kont	Eksp	Kont	Eksp	Kont	Eksp	Kont
\bar{X}	13,26	14,11	17,43	15,88	33,76	32,64	20,38	17,93
s	4,45	4,76	3,04	3,81	0,55	1,97	10,47	7,39
n	34	44	34	44	34	44	34	44
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	4,63		3,5		1,525		8,86	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	-0,183		0,443		0,734		0,276	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	-0,18		0,44		0,73		0,27	
$\overline{ES'_{sm}} = \frac{\sum ES'_{sm}}{4}$	0,315							
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (\overline{ES'_{sm}})^2}$	18,95							

Dobivena veličina učinka ovog sustava je $ES'_{sm} = 0.315$ i njena težina $w_{sm} = 18.95$. Svaki segment testa je posebno ocjenjen pa se za krajnju veličinu učinka koristila prosječna vrijednost.

4.4.6. KNOWLEDGE MAPPING

Knowledge mapping tj. mape znanja su definirane kao dijagramski prikaz glavnih semantičkih veza između skupa pojmova. U tom prikazu čvorovi predstavljaju pojmove a linije između njih pokazuju vezu između tih pojmova. Koriste se za poučavanje i ocjenjivanje znanja u različitim područjima znanosti.

Završni test se sastojao od dva dijela, od kojih je za analizu pogodniji drugi dio; test u obliku eseja (Tablica 3.4.). Ovdje su izdvojeni rezultati tog eseja (Tablica 4.6.), koji je najbolji pokazatelj kvalitete znanja u zadanom području. Pošto je svakom učeniku ocjenjen cjelokupni esej, i posebno sposobnost međupovezivanja sustava, stvarna veličina učinka će biti aritmetička sredina te dvije dobivene veličine učinka.

Tablica 4.6. Izračun veličine učinka i njene težine iz rezultata eseja

	CJELOKUPNI ESEJ		MEĐUPOVEZIVANJE SUSTAVA	
	Eksperiment	Kontrolna	Eksperiment	Kontrolna
\bar{X}	3,12	2,89	2,48	1,83
s	0,84	0,89	1,24	1,13
n	20	22	20	22
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	0,87		1,18	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	0,265		0,55	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	0,26		0,54	
$\overline{ES}'_{sm} = \frac{\sum ES'_{sm}}{2}$	0,4			
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (\overline{ES}'_{sm})^2}$	10,27			

Dobivena je veličina učinka kao prosječna vrijednost navedenih rezultata, i ona iznosi 0.4. Ta vrijednost predstavlja veličinu učinka ovog sustava. Nezina težina iznosi $w_{sm} = 10.27$.

4.4.7. SHERLOCK

Sherlock je inteligentni tutorski sustav koji predstavlja simulaciju okruženja u kakvom stručnjaci zrakoplovne tehnologije rade na ispitivanju ispravnosti te uklanjanju kvarova na djelovima zrakoplova.

Za vrednovanje ovog sustava provedeno je više testova koristeći dva tipa testa (VTT i NIT), a provodilo se i testiranje istim testom više puta. Najvažniji rezultati su prikazani u Tablica 3.5., to su rezultati dva VTT testa i jednog NIT testa. Zbog činjenice da su oba testa jednako kvalitetna za ispitivanje tog specifičnog znanja, i isti je broj ispitanika, veličina učinka ovog sustava biti će određena kao aritmetička sredina veličina učinka dobivenih tim testovima. U Tablica 4.7. su stoga izdvojeni rezultati sva tri testa.

Tablica 4.7. Rezultati provedenih testova i odgovarajući izračun veličine učinka i njene težine

	VTT3		VTT4		NIT	
	Ekспери	Kontroln	Ekспери	Kontroln	Ekспери	Kontroln
\bar{X}	95	59	91	58	87	75
s	5	37	7	37	12	14
n	18	23	18	23	18	23
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	27,985		28,17		13,16	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	1,286		1,17		0,91	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	1,26		1,15		0,89	
$\overline{ES'_{sm}} = \frac{\sum ES'_{sm}}{3}$			1,1			
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (\overline{ES'_{sm}})^2}$			9,11			

Dobivena je veličina učinka kao prosječna vrijednost rezultata navedena tri testa, i ona iznosi 1.1. Ta vrijednost predstavlja veličinu učinka ovog sustava. Nezina težina iznosi $w_{sm} = 9.11$.

4.4.8. ONLINE GENERAL COLLEGE CHEMISTRY COURSE

Online general college chemistry course je sustav e-učenja i poučavanja kolegija Opća kemija koji kombinira male grupe studenata, tjedne kontakte i web-temeljene materijale.

Rezultati na završnom testu ne uključuju podatke o standardim devijacijama za kontrolnu i eksperimentalnu grupu. Poznate su varijance za te dvije grupe, te zajednička varijanca pa će taj podatak biti koristan za zajedničku standardnu devijaciju potrebnu za izračun veličine učinka (Tablica 4.8.).

Tablica 4.8. Podaci vrednovanja sustava, te izračun veličine učinka i njene težine

	EKSPERIMENTALNA	KONTROLNA
\bar{X}	75,84	61,98
s	Nema podataka	Nema podataka
n	25	117
Zajednička varijanca	258,524	
$s_p = \sqrt{\text{zajed. varijanca}}$	16,08	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	0,86	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9} \right] ES_{sm}$	0,857	
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (ES'_{sm})^2}$	19,55	

Dobivena veličina učinka ovog sustava je $ES'_{sm} = 0.857$ i njena težina $w_{sm} = 19.55$.

Ovaj model poučavanja pokazuje relativno veliku učinkovitost, s obzirom da se poučavanje odvijalo na daljinu. Stoga su se dodatno ispitivali faktori koji su možda utjecali na uspješnost ovakvog poučavanja. Nije proveden inicijalni test za utvrđivanje ekvivalencije grupa na početku poučavanja, pa nije isključeno da su studenti u eksperimentalnoj grupi bili boljeg predznanja i sposobnosti. Dodatna analiza dviju grupa pokazala je da je prosječna dob u eksperimentalnoj grupi bila 24.6 godina, dok je u kontrolnoj grupi prosječna dob bila 18-19 godina. Autor je zaključio da su studenti u eksperimentalnoj grupi bili zreliji, discipliniraniji i ozbiljnije shvaćali poučavanje, pa je i to jedan od razloga veće uspješnosti. U prilog tome govori i podatak da su na satu ponavljanja za završni test, bili prisutni svi studenti eksperimentalne grupe, a samo 25% studenata kontrolne grupe.

4.4.9. DISTANCE LEARNING: BUSINESS COMMUNICATIONS CLASS

Na sveučilištu North Carolina je razvijen sustav e-učenja, i to učenja na daljinu za kolegij Business communications.

Studenti kontrolne i eksperimentalne grupe su pisali dva testa na kraju poučavanja: završni test i cjelokupni ispit znanja. Zatim im je ocjenjen izvještaj istraživanja o međunarodnom poslovanju, a na kraju su dobili i završnu ocjenu kolegija (Tablica 3.6.). Sve su te ocjene podjednako bitne u vrednovanju ovakvog načina poučavanja, pa će svi rezultati biti korišteni u dobivanju veličine učinka. Broj ispitanika u obje grupe je stalan kroz sve testove, pa je dopušteno kao krajnju veličinu učinka uzeti prosječnu vrijednost (Tablica 4.9.)

Tablica 4.9. Rezultati provedenih testova te odgovarajući izračun veličine učinka i njene težine

	Završni test		Ispit znanja		Izveštaj		Ocjena kolegija	
	Eksp	Kont	Eksp	Kont	Eksp	Kont	Eksp	Kont
\bar{X}	72,43	65,55	85,92	78,26	91,39	87,45	85,42	80,57
s	9,12	10,91	8,16	12,63	12,32	28,6	13,11	16,16
n	24	23	24	23	24	23	24	23
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	10,035		10,58		21,85		14,68	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	0,685		0,72		0,18		0,33	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	0,67		0,71		0,18		0,32	
$\overline{ES'_{sm}} = \frac{\sum ES'_{sm}}{4}$	0,47							
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (\overline{ES'_{sm}})^2}$	11,43							

Dobivena je veličina učinka kao prosječna vrijednost rezultata četiri navedene kategorije ocjena, i ona iznosi 0.47. Ta vrijednost predstavlja veličinu učinka ovog sustava. Nezina težina iznosi $w_{sm} = 11.43$.

4.4.10. COMPUTER BASED INSTRUCTION IN AN ENGINEERING CLASS

Na sveučilištu North Carolina State University, dizajnirane su CBI jedinice za poučavanje kolegija "Civil Engineering: Traffic Engineering". Proveden je eksperiment u kojem su se ispitali učinci računalno-temeljenog poučavanja na inženjering.

U Tablica 4.10. su izdvojeni rezultati testiranja koje je provedeno dva tjedna nakon završenog poučavanja. Kontrolna grupa je ostvarila malo bolje rezultate, ali nije postignuta statistička značajnost, što implicira da je računalno-temeljeno poučavanje pokazalo podjednaku učinkovitost kao i tradicionalna predavanja.

Treba napomenuti da je ovaj sustav dizajniran i testiranje je provedeno u akademskoj godini 1993/94, te da ovaj eksperiment spada među prve provjere učinkovitosti CBI- poučavanja u inženjeringu, što donekle opravdava njegovu slabiju učinkovitost.

Tablica 4.10. Rezultati vrednovanja CBI-jedinica u poučavanju kolegija Civil Engineering

	TEST QUESTION 1		TEST QUESTION 2	
	Eksperiment	Kontrolna	Eksperiment	Kontrolna
\bar{X}	9,6	10,33	3,85	3,94
s^2	42,6	36,1	3,5	3,58
n	20	18	20	18
$s_p = \sqrt{\frac{(n_E - 1)s_E^2 + (n_K - 1)s_K^2}{(n_E - 1) + (n_K - 1)}}$	39,53		3,54	
$ES_{sm} = \frac{\bar{X}_E - \bar{X}_K}{s_p}$	-0,0184		-0,0254	
$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm}$	-0,018		-0,025	
$\overline{ES}'_{sm} = \frac{\sum ES'_{sm}}{2}$	-0,02			
$w_{sm} = \frac{2n_E n_K (n_E + n_K)}{2(n_E + n_K)^2 + n_E n_K (\overline{ES}'_{sm})^2}$	9,47			

Iako nije postignuta statistička značajnost, izračunata je veličina učinka, koja ima negativan predznak zbog uspješnije kontrolne grupe, iznosi $ES'_{sm} = -0.02$. Njezina težina iznosi $w_{sm} = 9.47$.

Postoji jedinstven razlog za uključivanje ovog sustava poučavanja u analizu, s obzirom na njegovu starost i slabu učinkovitost. Zasigurno je provedeno još mnogo istraživanja u kojima je neki sustav e-poučavanja također pokazao podjednaku ili čak slabiju učinkovitost nego tradicionalna nastava, ali zbog izostanka učinka, rezultati tih istraživanja se rijede objavljuju. Stoga, da bi se uravnotežio omjer učinkovitih sustava, uključeni su ovi neutralni rezultati, i oni će utjecati na ukupnu veličinu učinka na način da će krajnja vrijednost korektnije pokazivati stvarni učinak ovakvih sustava.

4.5. IZRAČUN UKUPNE VELIČINE UČINKA

Ukupna veličina učinka i time povezani statistički računi mogu se jednostavnije izračunati kreiranjem još dvije varijable koje se onda sumiraju kroz sve zapise i koriste za izračun potrebnih podataka. Već je bilo riječi da je za svaki zapis potrebna veličina učinka i pripadajuća težina. Prethodno su napravljene sve potrebne prilagodbe i korekcije veličina učinka, posebice, u ovoj meta-analizi se koristila korekcija standardne razlike aritmetičkih sredina zbog relativno malog uzorka ispitanika. Isto tako, vodilo se računa da dobivene veličine učinka budu statistički nezavisne. Time su ispunjeni svi uvjeti, pa se može prijeći na izračun ukupne veličine učinka.

4.5.1. PROSJEČNA VELIČINA UČINKA

U Tablica 4.11. su sažeti rezultati svih sustava, tj. veličine učinka i pripadajuće težine. Koristeći navedeno, kreirane su još dvije potrebne varijable, to su: umnožak svake veličine učinka njenom težinom ($w_i ES_i$), te umnožak kvadrata svake veličine učinka njenom težinom ($w_i ES_i^2$), gdje je $i = 1$ do 10.

Tablica 4.11. Veličine učinka i težine pojedinog sustava te njihovi umnošci potrebni za daljnju analizu

		w	ES	w*ES	w*ES²
1	LBA	14,75	1,16	17,11	19,85
2	Conceptual Helper	9,97	0,635	6,33	4,02
3	SQL-Tutor	10,6	0,73	7,74	5,65
4	PsyCAL	44,46	0,85	37,79	32,12
5	Computer training	18,95	0,315	5,97	1,88
6	Knowledge Mapping	10,27	0,4	4,11	1,64
7	Sherlock	9,11	1,1	10,02	11,02
8	Online general chemistry	19,55	0,857	16,75	14,36
9	Bussiness communications	11,43	0,47	5,37	2,52
10	Engineering class	9,47	-0,02	-0,19	0,003
$\Sigma =$		158,56		111	93,063

Sada se ukupna veličina učinka računa kao mjerena aritmetička sredina svih dobivenih veličina učinka prema formuli (2.22), koristeći prethodnu tablicu. Tako prosječna veličina učinka navedenih 10 sustava iznosi:

$$\overline{ES} = \frac{\sum (w_i ES_i)}{\sum w_i} = \frac{111}{158.56} = 0.70 \quad (4.1)$$

Dobivena je prosječna veličina učinka svih sustava, koja iznosi 0.70, što znači da je prosječni korisnik zadanih sustava (učenik, student), 0.70 standardne devijacije uspješniji nego tradicionalno poučavan učenik ili student.

4.5.2. INTERVAL POUZDANOSTI

Interval pouzdanosti za prosječnu veličinu učinka temelji se na standardnoj pogrešci tog prosjeka i kritičnoj vrijednosti iz z-distribucije. Standardna pogreška dobivene prosječne veličine učinka, prema formuli (2.23) iznosi:

$$SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_i}} = \sqrt{\frac{1}{158.56}} = 0.079 \quad (4.2)$$

Sada se interval pouzdanosti od 95% oko prosječne veličine učinka računa koristeći formule (2.24). Kritična vrijednost z-distribucije za interval pouzdanosti od 95% iznosi $z = 1.96$ ($\alpha = 0.05$), a donja i gornja granica navedenog intervala pouzdanosti iznose:

$$\overline{ES}_L = \overline{ES} - z(SE_{\overline{ES}}) = 0.7 - 1.96 * 0.079 = 0.545 \quad (4.3)$$

$$\overline{ES}_U = \overline{ES} + z(SE_{\overline{ES}}) = 0.7 + 1.96 * 0.079 = 0.85 \quad (4.4)$$

Interval pouzdanosti od 95% oko prosječne veličine učinka ($0.545 < \overline{ES} < 0.85$) ne uključuje 0, pa je dobivena prosječna veličina učinka statistički značajna za $p < \alpha$.

4.5.3. ANALIZA HOMOGENOSTI

U homogenoj distribuciji, disperzija veličina učinka oko njihovog prosjeka nije veća od očekivane greške vezane uz uzorak. Drugim riječima, u homogenoj distribuciji individualna se veličina učinka razlikuje od prosječne samo za tu grešku.

Za provedbu testa homogenosti izračunava se Q vrijednost (prema formuli (2.27)), koja predstavlja distribuciju χ^2 (hi-kvadrat) sa 9 stupnjeva slobode ($k-1$ gdje je k broj sustava).

$$Q = \left(\sum w_i ES_i^2 \right) - \frac{\left(\sum w_i ES_i \right)^2}{\sum w_i} =$$

$$= 93.063 - \frac{111^2}{158.56} = 93.063 - 77.7 = 15.36 \quad (4.5)$$

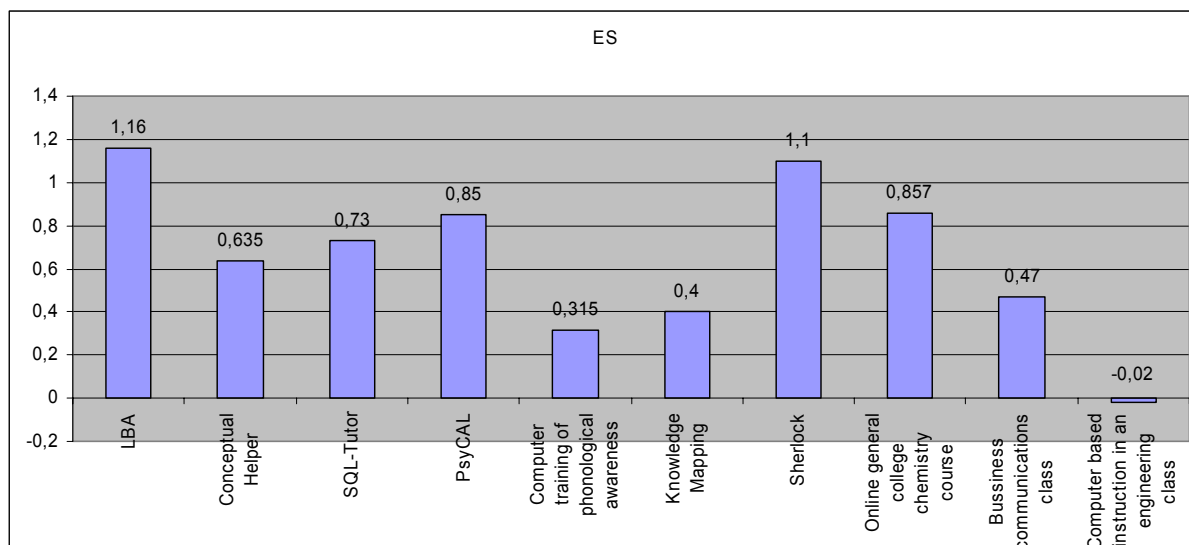
Dobivena vrijednost ne prelazi kritičnu vrijednost χ^2 za $\alpha = 0.05$, koja za 9 stupnjeva slobode iznosi 16.92. Nul hipoteza homogenosti se prihvaća, dana distribucija je, dakle, homogena. To znači da varijabilnost u ovom uzorku veličina učinka nije veća nego što se očekuje od greške vezane uz veličinu uzorka ispitanika.

Iako se sustavi i istraživanja, po karakteristikama i metodologiji, međusobno uvelike razlikuju, ovo svojstvo homogenosti implicira da dobivena prosječna veličina učinka jasno i smisleno reprezentira distribuciju veličina učinka danih sustava.

4.5.4. ANALIZA REZULTATA META-ANALIZE

Dobivena ukupna veličina učinka od 0.70 sigma, pokazuje umjerenu učinkovitost opisanih sustava u procesu učenja i poučavanja prema tradicionalnim metodama. Većina analitičara to interpretira na način da primjena navedenih sustava *premješta* (eng. shift) prosječnog učenika sa 50-tog percentila na 76-ti percentil na ljestvici uspješnosti. Ili da je uspješnost tradicionalno poučavane grupe 33% dok je za grupu poučavanu nekim od sustava e-učenja uspješnost 66%.

Na Slika 4.1. je prikazana distribucija dobivenih veličina učinka svih danih sustava.



Slika 4.1. Veličine učinka opisanih sustava e-učenja

U ovu meta-analizu uvršteni su sustavi koji uključuju različite oblike e-učenja, inteligentne tutorske sustave, sustave na web i računalno temeljenom poučavanju, sustave za poučavanje na daljinu itd. Međutim, sustavi koji spadaju u navedene kategorije nisu zasebno analizirani pa se ne mogu donositi zaključci o njihovoj učinkovitosti. Neki sustavi, zaista pokazuju veću učinkovitost, i u ovoj meta-analizi su to inteligentni tutorski sustavi čija je veličina učinka u korištenim istraživanjima iznosila 0.635, 0.73 i 1.1 sigma. Veliku su učinkovitost pokazali i web-temeljeni sustavi te sustavi sa interaktivnom multimedijom koji su pokazali veličinu učinka redom od 0.857 i 1.16 sigma. Značajni učinak su pokazali i sustavi za poučavanje na daljinu od 0.47 i 0.857 sigma, a najmanji su učinak pokazali računalno-temeljeni sustavi gdje su dobivene veličine učinka -0.02, 0.315 i 0.4. Ti se rezultati podudaraju sa rezultatima dosadašnjih meta-analiza, koje također na prvo mjesto po učinkovitosti stavljaju inteligentne tutorske sustave, a na posljednje računalno-temeljene sustave [FLET2003].

Međutim, nije moguće generalno, iz ove meta-analize donositi zaključke o veličini učinka za svaku od navedenih kategorija. Korišten je skroman broj istraživanja, po jedan do dva-tri istraživanja iz određene kategorije, što nisu reprezentativni uzorci da bi se analizirala pojedina kategorija. U prilog tome govore i rezultati poučavanja na daljinu koji u ovoj meta-analizi pokazuju veliku učinkovitost. Naime, prilikom analize jednog od tih sustava, utvrđeno je da je veliki učinak iz tog sustava dijelom posljedica velike motiviranosti i ozbiljnosti tako poučavanih studenata, što nije uvijek slučaj. Zato bismo počinili grešku, kad bismo donijeli zaključak da takav način poučavanja generalno postiže veliki učinak.

5. ZAKLJUČAK

Meta-analiza predstavlja sintezu istraživanja, stoga su dobiveni rezultati toliko kvalitetni koliko su kvalitetna istraživanja koja uključuje. Postupak odabira odgovarajućih istraživanja, je zato ključni korak u provedbi meta-analize. Potencijalna istraživanja se moraju detaljno proučiti a krajnji odabir mora biti rezultat dobro definiranih kriterija. Ako u bazi istraživanja te meta-analize ne postoje istraživanja sa visokom metodološkom kvalitetom, teško se može očekivati da će udruživanje njihovih otkrića dati valjane i korisne rezultate, a time dolazi u pitanje i smisao meta-analize. S druge strane, kvalitetno i pažljivo provedena, meta-analiza predstavlja korisno sredstvo kojim se ispituje učinkovitost neke metode ili tretmana, i ona daje smjernice za njihovo daljnje korištenje, njihovo unaprijeđenje ili pak odbacivanje u slučaju neučinkovitosti.

Sustavi e-učenja koji su opisani u trećem poglavlju, prema provedenim istraživanjima povećavaju učinak i kvalitetu procesa učenja i poučavanja. Premda je njihov učinak manji od spomenutog učinka individualnog poučavanja, stručnjaci se slažu da se u buduću, zbog velikog napretka tehnologije, očekuju sve bolji i učinkovitiji sustavi. Isto se tako i uvjeti za njihovu integraciju u svakodnevni proces obrazovanja iz dana u dan poboljšavaju. Pored prednosti sustava e-učenja pred tradicionalnim poučavanjem, a to se prvenstveno odnosi na povećanje instruktivskih interakcija u jedinici vremena i uštedu vremena potrebnog pojedincu za savladavanje određenog nastavnog sadržaja, dodatnu prednost predstavlja i pozitivno mišljenje učenika i studenata o takvim sustavima. Oni su u korist sustava e-učenja naveli sljedeće: da su beskrajno strpljivi, nisu nikad umorni, da se nikad ne uznemire niti se naljute, dopuštaju učenicima privatnost u učenju, nikad ne zaboravljaju ispravak i pohvalu, zabavni su i zanimljivi, individualiziraju učenje, omogućavaju učeniku učenje vlastitim tempom, ne omalovažavaju učenika kad pogriješi, u trenutku pružaju povratnu informaciju, objektivniji su od nastavnika i profesora, a njima ostavljaju vremena za kvalitetniji kontakt sa učenicima (za razgovor, davanje savjeta, pomoć u rješavanju problema vezanih i nevezanih uz nastavu...), nemaju predrasude o rasnoj i etničkoj pripadnosti, dobri su motivatori, poučavaju u malim pomacima, a rade izuzetnom brzinom - sličnom brzini ljudskih misli, itd...

Meta-analiza koja je provedena u četvrtom poglavlju, ispitala je učinkovitost tih sustava prema tradicionalnom načinu poučavanja. Kriteriji po kojima se vršio odabir istraživanja uključivali su ona koja odgovaraju temi i daju uvid u sve rezultate potrebne za izračun veličine učinka.

Meta-analizom je dobivena je ukupna veličina učinka od 0.7 sigma, koja predstavlja umjerenu učinkovitost sustava e-učenja prema tradicionalnoj nastavi. Taj učinak se još interpretira kao "skok" prosječnog učenika sa 50-tog percentila na 76-ti percentil ljestvice uspješnosti, nakon korištenja sustava e-učenja. Rezultati dobiveni provedenom meta-analizom u skladu su sa rezultatima prijašnjih meta-analiza, i pokazuju da su sustavi vrlo korisni u poučavanju. Oni kod učenika oblikuju pozitivan stav prema učenju i potiču motivaciju, što su glavni preduvjeti za uspješno provođenje učenja i poučavanja. Ako učenik primjeti da s lakoćom uči i postiže sve bolje rezultate, što je slučaj sa primjerima navedenih sustava, dobiva dodatnu motivaciju za daljnji rad i učenje. Zato bi dugoročno poučavanje pomoću tih sustava, po mom mišljenju,

pokazalo daleko veći uspjeh i učinkovitost od navedene, koja je rezultat primjene nekog od sustava na kraći rok. Bilo bi korisno za provjeru navedene pretpostavke, pratiti npr. eksperimentalni razred koji će na duži rok, možda od par godina, uz tradicionalnu nastavu biti redovito poučavan nekim od sustava e-učenja. Zatim istražiti uspješnost takvog razreda i usporediti s tradicionalno poučavanim razredom. Smatram da bi tako dobivena razlika bila daleko veća od prije navedene.

Sve prethodno predstavlja više nego dovoljne razloge za nastavak razvoja sustava e-učenja, njihovog unaprijeđivanja, poboljšavanja ali i veću implementaciju u redovni proces poučavanja. Nastavnici i profesori su zatrpani nastavnim sadržajem kojeg trebaju obraditi u određenom broju sati, pa je često slučaj da u nedostatku vremena svaki sat poučavaju novo gradivo, a ne provode redovite provjere znanja. Učenici obično ne uspiju pratiti takav tempo, ili im ponestane motivacije, baš zbog toga što znaju da profesor nema vremena provjeriti da li se nastavni sadržaj prati. Rezultat je nekvalitetno i neadekvatno znanje i slaba prolaznost na ispitima znanja, jer se velika količina gradiva uči neposredno pred ispit. Svrha aktivnije implementacije sustava e-učenja u redovito obrazovanje zato nije u zamjenjivanju živog učitelja, jer je to nemoguće izvesti u potpunosti. Oni bi korisno nadopunjavali nastavnike i profesore u poučavanju i tako olakšali njihove satnice, te im ostavili više vremena za kvalitetniji kontakt s učenicima, diskusiju, učestalije provjere znanja, razgovor itd.

6. LITERATURA

- [ALBA2000] Albacete, P. L., & VanLehn, K. (2000). Evaluating the effectiveness of a cognitive tutor for fundamental physics concepts. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society* (pp. 25-30). Mahwah, NJ: Erlbaum.
- [BLOO1984] Bloom, B.S.: The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13, (1984) 4-16
- [BUCH2000] T. Buchanan: The efficacy of a World-Wide Web mediated formative assessment. (2000) Blackwell Science Ltd, *Journal of Computer Assisted Learning* 16, 193-200.
- [CASA2001] Casanova, R. S. (2001). *Student performance in an online general college chemistry course*. An on-line conference on chemistry CONFICHEM "On-Line Teaching Methods"
- [COHE1969] Cohen, J.(1969) *Statistical Power Analysis for the Behavioral Sciences*. NY: Academic Press.
- [COTT1991] Cotton, K. (1991) *Computer-assisted instruction* (SIRS Close-up No. 10). Portland, OR: Northwest Regional Educational Laboratory, 23.11.2006. URL:www.nwrel.org/scpd/sirs/5/cu10.html
- [FLET2003] Fletcher, J.D.: Evidence for Learning from Technology-Assisted Instruction. In: *Technology applications in education: a learning view*, (H.F. O'Neal, R.S. Perez (Ed.)), Mahwah, NJ: Lawrence Erlbaum Associates, (2003) 79-99
- [FLET2005] J. D. Fletcher, R. Johnston (2005). "Chapter Seven: Instructional Technology: Effectiveness and Implications for the Intelligence Community", in R. Johnston ed. *Analytic Culture in the U.S. Intelligence Community: an Ethnographic Study*. Washington, DC: Central Intelligence Agency, Center for the Study of Intelligence
- [GOTT1995] Gott, S. P. (1995). *Cognitive technology extends the work environment and accelerates learning in complex jobs*. Brooks Air Force Base, TX: Armstrong Lab, Human Resource Directorate (NTIS No. ADA 303 597/9/XAB).
- [HUMM1996] Joseph E. Hummer, John W. Baugh Jr., Bhavani Konuru, and Steven M. Click: A controlled comparison of traditional classroom instruction with computer based instruction in an engineering class. In *1996 ASEE Annual Conference Proceedings*, 1996.

-
- [LEAR2006] Learnframe: e-Learning Management System,
URL:http://isp.webopedia.com/TERM/E/e_learning.html (19/11/06)
- [LIPS2001] Mark W. Lipsey, David B. Wilson: *Practical meta-analysis*. Sage Publications, Inc., Thousand Oaks, California, 2001.
- [MITR1999] Mitrovic, A. and Ohlsson, S.: Evaluation of a constraint-based tutor for a database language. *International Journal of Artificial Intelligence in Education* (1999), 10, 238-256.
- [OSMU1999] Osmundson, E., Chung, G. K., Herl, H. E. et Klein, D. C. (1999). *Knowledge mapping in the classroom : A tool for examining the development of students' conceptual understandings* (Technical report 507). Los Angeles : CRESST/University of California.
- [SEGE2005] Segers, E., Verhoeven, L. (2005). Long-term effects of computer training of phonological awareness in kindergarten. *Journal of Computer Assisted Learning*, 21, 17.
- [SMIT1977] Smith, M.L., Glass, G.V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.
- [TUCK2001] Tucker, S. (2001), "Distance education: better, worse, or as good as traditional education?", *Online Journal of Distance Learning Administration*, Vol. 2 No.4,
- [WIKI2006] E-learning from Wikipedia, the free encyclopedia,
URL:<http://en.wikipedia.org/wiki/E-learning> (19/11/06)
- [ZANG2006] Zang, D., Zhou, L., Briggs, R.O., and Nunamaker, J.F.Jr.: Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management* 43, (2006) 15-27.