

S AIAKUNTZA ETA SASISAIKUNTZA DISEINUAREN BALIDEZIA PSIKOLOGIAN



Nekane Balluerka Lasa
Xabier Isasi Balantzategi

SAI AKUNTZA ETA SASISAI AKUNTZA DISEINUAREN BALIDEZIA PSIKOLOGIAN

**Nekane Balluerka Lasa
Xabier Isasi Balantzategi**

UDAKO EUSKAL UNIBERTSITATEA
Bilbo, 1996

© Nekane Balluerka, Xabier Isasi

© Udako Euskal Unibertsitatea

ISBN: 84-86967-75-9

Lege-gordailua: BI-2077-96

Inprimategia: RALI S.A. Particular de Costa, 8-10, 7. planta 48010 BILBO

Azalaren irudia: Iñaki Balluerka Lasa

Banatzailleak: UEU. General Concha 25, 6. BILBO telf. 4217145

Zabaltzen: Igerabide, 88 DONOSTIA

AURKIBIDEA

Aitzinsolasa	1
--------------------	---

1.- Ikerkuntzari lotutako zenbait kontzeptu.....3

1.1.- Hizkuntza matematikoa psikologian	5
1.2.- Errealitatea	7
1.3.- Adierazgarritasuna eta esanguratsutasuna	8
1.4.- Zorizkotasuna	9
1.5.- Aldagaiak	11
1.5.1.- Aldagai azalgarriak	11
1.5.2.- Aldagai arrotzak	11
1.6.- Jomugaren teorema	12
1.7.- Fisher-Snedecorren legea	15
1.8.- Askatasun graduak eta efektuaren tamaina	16
1.8.1.- Askatasun graduak zorizko talde anitzeko diseinuan	18
1.8.2.- Efektuaren tamaina eta testaren indarra	20

2.- Balidezia kontzeptua eta balideziaren sailkapena ikuspegi desberdinen arabera.....27

2.1.- Balidezia kontzeptua	29
2.2.- Balideziaren sailkapenari buruzko ikuspegi ezberdinak	30
2.2.1.- Campbell eta Stanleyren eredua	30
2.2.2.- Cook eta Campbellen eredua	31
2.2.3.- Kruglanski eta Kroyren eredua	32
2.2.4.- Cronbachen eredua	34
2.2.5.- Sailkapen klasikoez bestelako bi ikuspegi: Marken eredua eta Chen eta Rossiren eredua	35
2.3.- Saiakuntza eta sasisaiakuntza diseinuaren eremuan sistematizapen hedatuena: Cook eta Campbellen eredua	40

3.- Ondorio estatistikozko balidezia	43
3.1.- Ondorio estatistikozko balideziaren definizioa eta ezaugarri nagusiak	45
3.2.- Ondorio estatistikozko balideziaren aurkako mehatxuak	49
3.2.1.- Eredu estatistikoaren baldintzen betekizun eza	49
3.2.1.1.- Normaltasun aurretikoa	51
3.2.1.2.- Bariantzen homogeneotasun aurretikoa.....	52
3.2.1.3.- Behaketen arteko askatasun aurretikoa.....	57
3.2.2.- Test estatistikoaren indar eskasa.....	59
3.2.3.- Efektuaren tamaina	61
3.2.4.- Laginaren tamaina.....	64
3.2.5.- I. motako errore maila	65
3.2.6.- Menpeko aldagaiaren neurketan fidagarritasun eta balidezia eskasa	66
3.2.7.- Tratamenduak ezartzerakoan fidagarritasun eskasa eta egiaztapen indar gutxiko tratamenduen aplikazioa.....	67
3.2.8.- Saiakuntza unitateen arteko ausazko heterogeneotasuna.	68
3.2.9.- Ausazko aldakortasunak saiakuntza egoeran	69
3.2.10.- "Outliers" edo ibiltartez kanpoko balioak.....	70
3.2.11.- "Missing" edo balio ez-agertuak.....	70
4.- Barne balidezia.....	75
4.1.- Barne balideziaren definizioa eta ezaugarri nagusiak.....	77
4.2.- Barne balideziaren aurkako mehatxuak	81
4.2.1.- Historia.....	81
4.2.2.- Heltzea	82
4.2.3.- Testen aplikazioa	83
4.2.4.- Neurketa baliakizuna.....	84
4.2.5.- Tratamendu anizkoitzen arteko elkarreagindakoa	85
4.2.6.- Erregresio estatistikoa	86
4.2.7.- Subjektuen hautapen diferentziala	88
4.2.8.- Hilkortasun esperimentalak edo subjektuen galera selektiboa	88
4.2.9.- Subjektuen hautapen diferentziala eta beste zenbait mehatxuren arteko elkarreagindakoa	90
4.2.10.- Kausazko eraginaren norabideari buruzko zalantzaraginitasuna	92
4.2.11.- Egoerak eragindako efektu erreaktiboak	92
4.3.- Barne balideziaren aurkako mehatxuak kontrolatu ahal izateko prozedurak	98
4.3.1.- Ezabapena	99

4.3.2.-	Konstantzia	99
4.3.3.-	Orekatzea subjektuarterko diseinuetan.....	100
4.3.3.1.-	Zorizkapena	100
4.3.3.2.-	Parekatzea	101
4.3.3.3.-	Doikuntza estatistikoa.....	103
4.3.4.-	Kontrabalantzeoa edo ekiponderazioa	105
5.-	Konstruktua balidezia	109
5.1.-	Konstruktua balideziaren definizioa eta ezaugarri nagusiak	111
5.2.-	Konstruktua balideziaren aurkako mehatxuak.....	113
5.2.1.-	Azalpen operazional desegokia.....	114
5.2.2.-	Operazionalizazio bakar baten erabilerak eragindako erroreak	114
5.2.3.-	Operazionalizazio metodo bakar baten erabilerak eragin- dako erroreak	115
5.2.4.-	Konstruktua eta konstruktua mailen arteko nahasketa	116
5.2.5.-	Tratamendu anizkoitzen arteko elkarrenergindakoa	117
5.2.6.-	Tratamenduaren eta alde aurretik aplikatutako testen arteko elkarrenergindakoa	117
5.2.7.-	Konstruktua batetik bestera orokortzeko aukera eskasa ..	118
5.2.8.-	Egoerarekin eta elkarrengintza sozialarekin zerikusia duten faktoreetatik eratorritako mehatxuak	118
5.2.8.1.-	Hipotesien sorkuntza testuinguriko ezaugarrietan oinarrituz	119
5.2.8.2.-	Saiakuntza subjektuek hartzen dituzten errol desberdinak.....	120
5.2.8.3.-	Esperimentatzaileak eragindako erroreak.....	121
5.2.8.4.-	Subjektuak eta esperimentatzaileak eragindako erroreen kontrola	124
6.-	Kanpo balidezia.....	129
6.1.-	Kanpo balideziaren definizioa eta ezaugarri nagusiak.....	131
6.2.-	Kanpo balideziaren aurkako mehatxuak	132

6.2.1.-	Tratamenduaren eta kanpo balideziaren oinarrizko dimentsioen arteko elkarreragindakoak	133
6.2.2.-	Kanpo balidezia zalantzan jartzen duten beste zenbait faktore	137
6.2.2.1.-	Tratamendu anizkoitzen arteko elkarrera-gindakoa	137
6.2.2.2.-	Testaurrearen sentsibilizazio efektua.....	138
6.2.2.3.-	Testostearen sentsibilizazio efektua	138
6.2.2.4.-	Saiakuntza egoerarekiko eta ebaluazio prozesuarekiko erreaktibitatea	139
6.2.2.5.-	Saiakuntza egoeraren berritasunak eragindako efektuak	139
 7.-	 Balidezia mota desberdinak ikerketaren eremuan: elkarren arteko harremanak eta lehentasunak.....	 143
 8.-	 Diseinuaren balideziaren azterketa ikerketa eremu zehatzean: ikuspegi aplikatua	 149
 9.-	 Psikologian erabiltzen diren diseinuen balideziari buruzko planteamendu berriak	 157
9.1.-	Marken eredua.....	159
9.2.-	Chen eta Rossiren eredua.....	161
9.3.-	Campbellen iradokizun berriak barne balidezia eta kanpo balidezia kontzeptuen inguruan eta Cook eta Shadishen hausnarketak diseinuaren balideziaz testuinguru sozialean	164
	BIBLIOGRAFIA AIPAMENAK.....	167

AITZINSOLASA

August Comtek psikologoa *fisikari soziala* gisa definitu zuenetik gaurdaino, Psikologian aurrerapen franko eman izan da. Psikologia gizarte zientzien esparru zabalean apurka-apurka dagokion lekua hartzen ari da. Eginkizun horretan, psikologiaren arlo desberdinetan garatutako formalizazio maila gero eta handiagoa da. Hala ere, L. Vigotskirekin bat gentozke Psikologiak oraindik berezko metodori ez duela esaterakoan. Alde handia baitago psikologiaren alor desberdinen artean; nekez gaindi daiteke Gizarte Psikologia eta Psikologia Dinamikoaren arteko hutsunea, zaila da Psikofisiologia eta Nortasunaren Psikologiaren arteko lotunerik aurkitzea.

Askotan, psikologoaren prestakuntzak *puzzle* osa ezinaren itxura hartzen du. Enpirismo gordina eta idealismo gorriaren arteko lehia bailitzan, psikologoak ez daki kontzientzia datu matrize baten determinantea den, hipotalamoaren azpiko osagai txiki bat ote den edo adiera jakinik gabeko hitz hutsa den, besterik gabe. Guk ere ez dugu argituko, baina lantzen duguna baliagarri gerta dakiguke ikertze prozesu horietan. Lewinek teoria on bat baino hoberik ez dagoela adierazi zuen, bat gatoz eta areago erantsiko genuke, Psikologiaren batasuna eta aurrerapena metodologiaren eskutik datoz, guk hori uste dugu behintzat.

Edozelan ere, liburu honetan psikologia ikasi baino, psikologiaz ikas daiteke, metodologiari dagokionaz alegia. Liburuaren muina eta ardatza balidezia da. Balideziari buruzko azterlan sakona eta gaurkotua burutu da. Eskola liburu gisa idatzita dago, irakurtzeko baino ikasteko, zirriborratzeko, azpimarratzeko, hitz batez, aztertzeko. Lagun batek behin esan zuen eleberria ona denean hiru kakalditan aisa irakurtzen dela, ziur asko horrela da. Zoritarrez, hau ez da eleberria, ez dugu uste irakurlerik saiaturik ere liburu honi duen guztia atera diezaiokeenik hiru tirritilo luzeetan. Haatik, ez bedi inor izu, ez baitugu inor horrelako atakan jarri nahi.

Liburua bederatzi ataletan eratu dugu. Lehena sarrera gisa idatzita dago, ondorengo ataletan agertzen diren zenbait kontzeptu argitzeko. Funtsean, oinarritzko metodologia eta analisi estatistikoei buruzko kontzeptuak. Bigarrenetik aurrera balideziaren azterketari lotzen gataizkio. Balideziaren kontzeptua (2. atala), ondorio estatistikozko balidezia (3. atala), barne balidezia (4. atala), konstruktua balidezia (5. atala) eta kanpo balidezia (6. atala) zehatz-mehatz aztertu eta garatzen dira bost atalen zehar, aipamen bibliografiko eta erreferentzia ugari eskainiz. Azkeneko 3 ataletan (7., 8. eta 9. ataletan) balideziaren ikuspegi aplikatua nabarmentzen da. Beraz, atal hauek aurreko atalen osagarri har daitezke.

Liburuaren erabilera erraztu nahian, laburpen gisa kontzeptuen taulak eratu ditugu. Ildo beretik, amaieran izen eta kontzeptuen aurkibidea erantsi ditugu, irakurleak nahi dituen erreferentziak ahalik eta azkarren aurki ditzan.

Edozelan ere, arestian esan dugunez, liburua eskola liburu modura eratuta dago. Psikologiako ikasleentzat egina, beraz. Zuek, ikasle nekagaitzak, izango zaituztegu irakurle. Gure itxaropena eta nahia, liburu hau zuentzat erabilgarria izatea da. Irrikitan gaukazue zuen iritziak, kritikak eta ekarpenak jasotzeko. Bai edukiaren aldetik, balizko hutsuneak, okerrak edo gaizki ulertuak; zein hizkuntza aldetik, terminologiaren arazoak; nahiz itxura aldetik, aurkibideak edo atalen arteko barne antolakuntza, besteak beste. Izan ere, liburugintza kontuan, gu ere ikasle baikara.

Nekane Balluerka, liburuaren benetako egilea, lankide zintzoa da, langile amorratua (ai Bittor!), batzutan laguna ere badudala pentsatzen dut. Beste inori baino, berari zor diozue eraikitzen ari garen Euskal Unibertsitatean jarritako adreilu hau. Neronek ere eskerrak eman niezazkioke igeltsero lanak egin izan ez banitu.

Aitzinsolasa amaitzeko beste aipamen bi egin nahi nituzke. Aurrena euskarari dagokio; euskara tekniko ulerkorra erabili nahi izan dugu, azkeneko arauetara irizpideei lotuta. Hala ere, gure laguna den J.R. Etxebarría (1994) Euskal Unibertsitateko irakasleak zioenez, *gai zientifikoa lantzean kontutan izan behar da zientzilarien hizkera edo "argota", arlo horretako hitzak zientzilariek erabiltzen dituzten moduan ulertu behar direlarik, eta ez hitzen esangura literal edo arrunta besterik gabe hartuz*. Euskararen egoera dela eta, izugarritzko lana egiteke dago, eginkizun horretan zuek, ikasle/irakurle, baduzu zer eginik, zure iradokizunak eta, batik bat, zure erabilera funtsezkoak baitira.

Atzeneko aipamena unibertsitateari doakio. Liburu hau Udako Euskal Unibertsitatearen (UEU) eskutik datorkizu. Egileok, batez ere Nekane Balluerkak, lan eskerga egin behar izan dugu testua argitara ekartzeko, urteetako lana esango genuke. Egia da, Eusko Jaurlaritzak euskararen kalitatea jagon eta zerbait eman digula liburua zure eskuetan jartzeko, Autonomi Erkideko unibertsitate publikoak ostera, ezer ez. Gure unibertsitate publikoaren lehentasunak beste batzuk omen dira. Esku artean duzuna, bada, lan militantearen emaitza da. Horrela hartu behar duzu, gure etekinik oparoena euskararen normalkuntzan aurreratzea da, gure ekimenetik besterik ez dugu jaso.

Aitzinari gibela eman orduko, liburua egiten lagundu diguten guztiei eskerrak eman nahi dizkiogu. Horretarako aurreko liburuan antzera amaiera emango diogu aitzinsolas honi:

*Ulertzen gaituztenei mila esker,
ulertzen ez digutenei berriz, deus ez.*

Donostian, 1996eko urriaren 13an

-Xabier Isasi Balanzategi-

1.- Ikerkuntzari lotutako zenbait kontzeptu

1.1.- Hizkuntza matematikoa psikologian

1.2.- Errealitatea

1.3.- Adierazgarritasuna eta esanguratsutasuna

1.4.- Zorizkotasuna

1.5.- Aldagaiak

1.5.1.- Aldagai azalgarriak

1.5.2.- Aldagai arrotzak

1.6.- Jomugaren teorema

1.7.- Fisher-Snedecorren legea

1.8.- Askatasun graduak eta efektuaren tamaina

1.8.1.- Askatasun graduak zorizko talde anitzeko
diseinuan

1.8.2.- Efektuaren tamaina eta testaren indarra

1.- IKERKUNTZARI LOTUTAKO ZENBAIT KONTZEPTU

1.1.- HIZKUNTZA MATEMATIKOA PSIKOLOGIAN

Aspalditik gizarte zientzietan, oro har, eta haien artean Psikologian ere, matematika gero eta gehiago erabiltzen dela gauza jakina da. Matematikak berez, formalizazio maila handia dakar, zehaztasuna eta orokortasuna aldi berean. Edozein zientzitan gertatzen den moduan, matematikak badu eduki oso betekizun garrantzitsua psikologiaren garapenean. Psikologia zientzia gisa nahikoa berria bada ere, Thurstonek (1959) psikologiaren ezinbesteko matematizazioa beharrezkoa zela aldarrikatzen zuen. XX. mendearen hondarrean horretan gaude oraindino. Liburu honen irakurleak, oro har, matematikaren eta, bereziki estatistikaren oinarritzko ezagutza behar du.

Horregatik guztiagatik, hizkuntzaren zehaztasunak berebiziko garrantzia du liburu honetan. Sarri, zenbait hitzen adiera mugatzen dugu harik eta hizkuntza *arrunta* behartzeraino. Hitz horiek irakurleak horrela ulertu behar ditu. Batzuetan arrotz eta nekagarri gerta dakioke baina ez dago beste biderik argitasuna lortzeko.

Saiakera honetan, euskara lanabes eta ikasbidea izanda ere, ulergarritasunari men eginez maiz maileguetara jo izan dugu. Beste zenbaitetan oster, ez, euskarak berak dituen baliabideei heldu diegu eta. Egiari zor, ez gara hizkuntzalariak eta euskararen egoera ez da samurra; horrek badakarkigu ekarri arazo ugari. Edozelan ere, euskararen bidetik hizkuntzaren kalitatea ahalik eta estuen zaindu nahi izan dugu. Geroak eta liburuaren erabiltzaileak esango du gure ekarpena noraino iristen den. Bien bitartean edukiari lotuko gatzazkio.

Euskara *normaldu* ahala eremu desberdinetan hizkera bereziak agertzen dira. Hori da gure kasua hain zuzen ere. Eginkizun horretan matematika lagungarria izaten dugu. Dena den, matematikaren ekarpena ez da alderdi horretara mugatzen. Gaur egungo gizarte psikologian, psikometrian, garapenaren psikologian, psikodiagnostikoan eta, oro har, psikologian, matematikaren arrastoa nonahi suma daiteke. Azken urteotan psikologiak aldaketa asko eta sakonak ezagutu ditu, etengabeko aldatze prozesuan dagoela dirudi, baina, gure ustez, aurrerapenik nabarmenena alderdi metodologikoan ari da ematen. Test estatistikoak eta hizkera estatistikoa arruntak bilakatu dira psikologiaren alor askotan. Egungo psikologoak estatistikaren oinarritzko ezagutzak inoiz baino gehiago behar ditu test psikometrikoak eta beste zenbait diagnostiko tresnak egoki erabili ahal izateko. Are gehiago behar ditu, ereduak eraiki nahi baditu, psikologia esperimentalean murgiltzen bada eta, jakina, ikerkuntzan iharduten badu.

Hizkuntza matematikoa psikologian garrantzizkoa izanagatik, psikologoaren funtsezko eginkizuna psikologiaren esparruari dagokio. Hau da, psikologia alorreko ikertzailearen helburua izaera psikologikoko gertaeretatik

abiatu eta ondorio psikologikoetara iristea da. Matematika portaera zientzietan baliagarria da baldin eta psikologiaren garapena ekartzen badu. Edozein ikerketaren euskarri metodologikoa erabakigarria izan daiteke baina ikerketaren funtsa, beti, adierazpenetan eta ondorioetan datza. Euskarri metodologiko konplexua eta landua erabiltzeak ez dakar, besterik gabe, psikologiaren inongo aurrerapenik planteatutako arazoari konponbiderik ematen ez badio. Ikerkuntzan, arazoaren ebazpenean dago gakoa. Tamalez, zenbaitetan euskarri metodologikoa halako erretolika antzu bilakatzen zaigu, non psikologiaren aurrerapena behin baino gehiagotan haren itzalean geratzen zaigun.

Esanak esan, testuliburu honek psikologiaren alderdi metodologikoari heltzen dio bete-betean. Balidezia arazoa zabal eta sakon jorratuko dugu. Haatik, horretan murgildu baino lehen, testuliburuan zehar agertzen diren zenbait kontzeptu agertu nahiko genituzke.

Gizarte zientzietan, oro har, eta portaera zientzietan, bereziki, estatistika berebiziko tresna da erabakiak hartzeko. Berez, erabakigarria ez bada ere, gero eta garrantzi handiagoa dauka. Psikologiaren garapenean bere eragina maiz agerikoa izan gabe, sakonean oso nabarmena da.

Testuinguru horretan diseinu estatistikoek nahi eta ahal denaren arteko esparruak lantzen dituzte. Ikerketa proiektua bera izendatzerakoan, bere lan eremua, ikergaia, hedapena eta tamaina, besteak beste, diseinu estatistikoen mugatzaile eta eragileak dira aldiberean. Portaera zientzietako ikerkuntzan ez daukagu saihesterik beraz.

Liburu honen edukia saiakuntza diseinuari edo diseinu esperimentalari lotzen zaio zuzenean, nahiz ageriko gaia *balidezia* izan. Hori horrela, diseinuaz hitz egin behar dugu eta, diseinuen artean esparru esperimental (saiakuntza eta sasisaiakuntza) eta ez-esperimental bereiztu behar ditugu. Oro har, alderdi esperimentalean *saiakuntza eta sasisaiakuntza* diseinuak bereizten ditugun bitartean, alderdi ez-esperimentalean *behaketa azterketen diseinua* eta *inkesta laginketaren diseinua* bereizten ditugu.

Kasu bakoitzean, diseinu mota bat edo bestea hautatzean, zientzia enpirikoetako oinarritzko auzi filosofikoak aurki ditzakegu. Nola burutu inferentziak populazio handiekiko, mugagabeko unibertsoekiko edo kausa sistemekiko baldin eta gure abiaburua behaketen laginak besterik ez badira? Are gehiago, lagin horiek errore eta zorizko aldaketen menpe badaude? Ezin bazter ditzakegu arazo horiek beraz. Alderdi honi heldu eta bertan sakondu dutenak, besteak beste, David Hume (1740), Popper (1959), Salmon (1967), Burks (1977), Fisher (1935), Neyman (1934), etab. ditugu.

Egia esan, test estatistikoen teknika hutsalak ez du gehiegi argitzen. Teknika horiek erabili aurretik ikertzaileak ezkutukoarena argitzeari heldu behar dio. Hau da, diseinuaren hautapena ez da nolanahiko erabakia.

1.2.- ERREALITATEA

Demagun euskara zuzpertzeko plangintza bi edo gehiago ditugula udal mailan ezartzeko. Gure helburua, jakina, posibleen artean egokiena eta eraginkorrena aukeratzea da. Horretaz, plangintzak ebaluatu eta (elkarren artean) erkatu behar ditugu; ildo horretatik lan egitasmo bat edo diseinu bat behar dugu. Lantegi horretan aurreneko irizpidea **errealitatea** da. Hau da, **aldagai azalgarrien benetakotasuna**, bai aldagai erabilgarriak edo manipulagarriak (*tratamendua*), zein ondorioztatutakoak (*erantzuna*).

Ezinbestean plangintzak definitu eta **operazionalizatu** behar ditugu. Maila honetan adostasuna bilatu behar dugu, sarri legozkioken analisi teknikak ezin erabil baitaitezke. Ikerketak erabateko egokitasunaren eta **bideragarritasunaren** arteko oreka puntua aurkitu behar du. Abiaburu abstraktuegia jarritz gero, arazo zehatzek eta aplikazio erabilgarriek ihes egingo digute. Zein da *euskara zuzpertzeko plangintzaren* helburua? Euskaldunen kopurua handitzea? Euskararen erabilera zabaltzea? Euskararen estatusa hobetzea?... Itau horien erantzunetatik eratortzen da plangintzaren ebaluaketarako gakoa. Horrela, **operazionalizatzen** dugu gure diseinua, **barne edo menpeko** aldagaiak definituz alegia.

Zer neurtu behar (nahi) dugun zehaztu beharra dago beraz. Gure gogoeta oster, neurtzen dugun hori, hein batean behintzat, plangintzaren *ondorioa* delakoan datza. Atzeman egin nahi dugu, **pronostikatu** hain zuzen ere. Plangintzaren arrakasta edo porrota neurketa horien gora-beherei lotzen zaie. Horra gure diseinuaren proposamena beraz. *Euskara zuzpertzeko plangintza* hobesten dugu bere *emaitzetan* emankorra dela ikusten dugunean. Hau da, plangintzaren arabera jarritako baliabideak eta bitarteak direla eta *euskara, maila jakin batean, zuzperten denean*. Baina, nola jakin *zuzpertze maila* hori erabilitako plangintzaren *ondorioa* den? Edozein plangintzak, berez, badakar *zuzpertze maila adierazgarririk*? Plangintzarik gabeko *zuzpertzetik* eman badaiteke, nola bereiz berezkoa dena eta plangintzari egotzitakoa?

Tankera horretako galderek gure auziaren muineraino garamatzate. Laburbilduz, bat, ez dago erabateko ziurtasunik edozein plangintzak euskararen zuzperketarako balio duenik eta, bi, plangintzaren arrakasta edo porrota **faktore** askoren menpe dago. Azken buru, plangintza guztiak ez direla baliokideak adierazi nahi dugu. Adibide honen bitartez azaldutakoa, portaera zientzietan aztergai gehienetara orokor daiteke. Zoritzarrez, zorionez agian, gure zientzia honetan oraindik ez daukagu *lege* bat ezartzerik portaera auresateko. Edozelan ere, legerik badago oraindik ez dugu ezagutzen. Proposamen bat edo beste egin dezakegu, egin ohi dugu hurbilpen gisa baina ez besterik.

1.3. ADIERAZGARRITASUNA ETA ESANGURATSUTASUNA

Testuliburuaren hariari jarrai diezaiogun. Bi edo plangintza gehiago genituela esaten genuen. Dena den, plangintzak **populazioan** ezartzekoak dira, gizartean alegia. Plangintza berak ez luke emaitza berdinik ekarriko udalerri euskaldun edo erdaldun batean. Hona hemen beste arazo bat bada, plangintzaren **subjektuak** dei genezakeena. Plangintzak elkarren artean erkatu ahal izateko **subjektu eta baldintza** baliokidetan ezarri behar ditugu. Nola berma dezakegu baliokidetasun hori? Bide bat ezaugarri bereko udalerri edo giza multzoak aukeratzea izan daiteke. Baina, posible ote da, zentzu hertsian ezaugarri bereko bi edo giza multzo gehiago aurkitzea? Biztanleriaren tamaina, industrializazioa, euskaldunen proportzio eta banaketa, adina... baliokideak direla esaterik ote dugu? Zinez zaila dela. Hala ere, bada izan bide bat nolabaiteko *baliokidetasuna* aurkitzeko, *zorizko hautaketa*.

Zorizko hautaketa esaterakoan probabilitate legeen eremuan murgiltzen gara ezinbestean. Ez da *edonolako hautaketa*, prozedura zehatz eta arautua baizik. Funtsean, subjektu guztiek eta bakoitzak gure laginan izateko probabilitate maila bera daukatenean *zorizko laginketa* (aukeraketa) dela esaten dugu. Horren guztiaren itzalean *adierazgarritasunaren* arazoa dago. *Adierazgarria* hitza zentzu hertsian erabiliko dugu testuan zehar. Oro har, lagina populazioarekiko *adierazgarria* dela esaten dugu populazioa **egoki** isladatzen duenean; neurri batean bederen, lagina populazioa **ordezkatz**en duenean alegia. *Adierazgarri* hitzaren esanahia, adiera hain zuzen, mugatu nahi dugu beraz. *Lagina adierazgarria* populazioa zuzen eta egoki **ordezkatz**en duen azpimultzo gisa zehazten dugu; *aldagai adierazgarria* berriz, aldagai posibleen artean eraginkorra, nabarmena edota garrantzitsua den aldagaiari esaten diogu.

Gure eginkizunetarako *adierazgarria* hitzaren adiera zabalegia da eta maiz *test estatistikoen* esparruan *signifikazio estatistikoa* adierazteko erabiltzen da. Gure aburuz, hortik nahasketa asko etorri ohi da. Izan ere, *test estatistikoen* ondorioz **esanguratsu** har daitezkeen erlazio guztiak ez baitira, ezinbestean eta berez, **adierazgarriak**. Are gehiago, subjektu asko edo lagin handiekin lan eginez gero erraz gerta dakiguke erlazio estatistiko gehienak **esanguratsuak** izatea, kalkulu prozesuan laginaren tamainak eragin erabakiorra duelako hain zuzen.

Hori horrela, hiru aukera genituen, hots, maileguetara jo eta *signifikazioa* hitzari eutsi edo, euskararen baitan baliabiderik izanik, *adierazgarria* adiera guztietarako erabili, edo, hautatu duguna, bi hitz aukeratu kontzeptu desberdinak adierazteko. Azken horri atxikiz, **adierazgarria** zentzuz eta edukiz betetakoa dela diogu, ezagutza arruntaren arabera jakin daitekeena, ikuspegi teorikoaren arabera trinkoa dena, bere horretan zilegi eta bidezkoa dena, hain zuzen ere. **Esanguratsua**, berriz, *test estatistikoen* ondorioei legokieke, **erlazio esanguratsuak** estatistikoki onargarriak har daitezkeen horiek izango lirateke, **adierazgarriak** izan edo izan gabe.

1.4. ZORIZKOTASUNA

Ikuspegi estatistikotik *zoriak* ez du *patuaren* esangura bera. Probabilitatearen kontzeptua ezinbestekoa da *zoriaren* adiera estatistikoa ulertzeko. *Zorizkoa* araurik gabeko orori esan ohi zaion arren, portaerari lotutako jazoerak, neurri batean behintzat, denboraren jarioan nola gauzatzen diren, zeren arabera gertatzen diren jakiterik badago. Gure alorrean, portaera zientzietan, gertaera gehienak, ia guztiak esan genezake, *zorizko gertaerak* dira. Hau da, jazo aurretik nola, noiz, zein neurritan eta zergatik gertatzen diren ez dakigu, nahiz gerta behar dutela badakigun. Edozelan ere, bada izan *probabilitate* maila jakin bat gertaeraren agerpen/ez-agerpena zehaz dezakeena. Hori baino harantzago jakiteak eredu deterministara eramango gintuzke; hots, jazoera lege jakinari lotzen zaionean, lege horrek erabat zehaztutako (*determinatutako*) eredia definituko luke.

Zorizkotasuna probabilitate legeen ikuspuntutik oinarritzko kontzeptua da. **Zorizko aldagaia**, oro har, aldeztu aurretik nola aldatzen den jakiterik ez dago, baina berau aztertuz gero **behatutakoa** nolakoa den atzeman eta, geroari begira, horrela izan daitekeela itxaron dezakegu. **Zorizko laginak** berriz, populaziotik *zoriz* aukeraturiko azpimultzoak direlarik, populazioaren edozein elementu hauta daiteke; zehatzago, probabilitate maila berberaz aukera daiteke edozein elementu. Laginaren tamaina nahikoa bada, aipatu *zorizkotasunak* damaie laginei **populazioarekiko adierazgarritasuna**.

Hala ere, diseinuetan *zorizkotasunak* badu beste zehaztapen zorrotzagoa, ikerketetan, *tratamendua*(*k*) ezarri baino lehen, *zorizkotasuna* taldeen arteko baliokidetasunaren berma baita. Hots, erkatu nahi ditugun taldeak *zorizkoak* badira, baliokideak direla erdiesten dugu. Beste era batera esan dezakegu: populazioaren bi edozein talde, tamainaz antzekoak eta nahikoak, *zoriz* aukeratuak izan badira, probabilitate legeen arabera beren ezaugarri nagusienetan baliokideak dira.

Bada *zorizkotasunak* dakarren beste abantaila: *saiakuntza errorea* kontrolatu ahal izatea, hain zuzen. *Errorea*, berez, ohiko arazoa da edozein neurketa prozesuan, zientzia zehatzetan ere, eta are ohikoagoa portaera zientzietan. Nolanahi ere, *neurketa errorea zorizkoa* da, eta, *zorizkoa* delako hain zuzen, banaketa normalari doitzen zaio, non berorren *itxaropen matematikoa* zero den.

Saiakuntza diseinuetan, beste alderditik, *zorizkotasunak*, arestian aipatutakoaz gain, beste bi adiera ditu. Bat, lehen esan duguna: *saiakuntza taldeen zorizko eraketa* esan genezakeena. Hau da, populaziotik *ausaz* hautatutako subjektuak *talde esperimentaletan zoriz* banatzeari esaten diogu. Bestea, taldeei ezarri beharreko *tratamendu* desberdinak *zoriz* egokitzea zehaztutako talde esperimental desberdinei.

Arestian jarritako adibideari berriro heltzen badiogu, honela zehatz genitzakeen eman beharreko urratsak:

1. Behin aztertu beharreko populazioa zein den mugatu ondoren, subjektu esperimentalak *zoriz* aukeratu. Subjektu kopurua, populazioaren tamaina eta, batez ere, erabili nahi dugun diseinuaren araberakoa izango da. Gure aztergaia euskara zuzpertzeko plangintza izanik, euskaldun arruntak hautatu beharko genituzke.

2. Subjektuak *zoriz* aukeratuak izan ondoren, zehaztutako *saiakuntza talde* desberdinetan *ausaz* banatzen ditugu. Horrela, taldeak populazioarekiko **adierazgarriak** eta elkarren artean **baliokideak** dira. Haiek populazioa ordezkatzeko dutela eta haien artean diferentziarik ez dagoela esan dezakegu, hain zuzen ere.

3. Azkenik, gure *tratamenduak* -euskara zuzpertzeko plangintza desberdinak- eratutako talde desberdinei *zoriz* egokitzen dizkiegu.

Neurri batean behintzat, esandako hiru urrats horiek esperimentaltasunaren oinarria ezartzen dute.

Urratsez urrats, ikerketa diseinuetarako oinarritzko hiru irizpide ekarri dizkizuegu: *errealitatea*, *adierazgarritasuna* eta *zorizkotasuna*. Bati edo besteari lehentasuna emateak diseinu mota desberdinak aholkatzea dakar. Hala nola, behaketa kontrolatutako diseinuetan *errealitateari* eman behar diogu lehentasuna. Laginketa bitarteko ikerketetan (inkestak, e.a.) ostera, *adierazgarritasuna* izan behar du funtsezko irizpide. Azkenik, saiakuntza diseinuetan edo diseinu esperimentaletan *zorizkotasunari* emango diogu garrantzirik handiena.

1.5. ALDAGAIAK

Irakurleak dagoeneko badaki aldagaia zer den. Ikerkuntzan, jakina, aldagaiak ezinbesteko lanabesa ditugu, beraien bitartez informazio adierazgarria eskuratzen dugu eta. Testuinguru honetan aldagaiak definitzeko eta azaltzeko bide anitz dago, honetan diseinuei ematen diegularik garrantzia. Horrela, aldagaiak bi kategoria nagusi hauetan bana ditzakegu:

1. *Aldagai azalgarriak.*

2. *Aldagai arrotzak.*

1.5.1. ALDAGAI AZALGARRIAK

Aldagai azalgarriak ikerketa diseinuaren abiaburu eta xede dira. Diseinu desberdinetan izen desberdinak hartzen dituzte. Oro har, bi multzotan bana ditzakegu, *aldagai azalgarriak* (X) eta *aldagai azalduak* (Y), hain zuzen ere. Eredu estatistikoetan, laginketa bitarteko ikerketetan *aldagai azalgarriek*, *kanpo aldagaiak* (X) esaten diegu; *aldagai azalduak* berriz, *barne aldagaiak* (Y). Saiakuntza diseinuetan, testuliburu honetan lantzen ditugunetan hain zuzen ere, *aldagai azalgarriak aldagai askeak* (X) dira, eta *aldagai azalduak menpeko aldagaiak* (Y). Saiakuntza diseinuetan *aldagai askeak* ikertzailearen kontrolpekoak dira, ez dira *zorizko aldagaiak* beraz.

Esanak esan, *aldagai azalgarriak*, batzuk zein besteak, ikerketa helburuei dagozkie. Jakintza alorreko ezagutza eta ulerkuntzatik datozkigu, eta edozein ikerketan haien definizioa eta *operazionalizazioa* aurreneko eginkizuna izaten da.

1.5.2. ALDAGAI ARROTZAK

Oro har, ikerketaren eginkizun eta helburuetatik kanpo geratzen den aldagai oro *aldagai arrotza* da. *Aldagai arrotzek* ikerketa diseinuetan duten eragina alborapen gisa har daiteke. Ikertzailearen eginkizuna, besteak beste, *aldagai arrotzen* kontrola da. Kontrol mailaren arabera, hiru kategorietan bil ditzakegu aldagai hauek:

- *Kontrolatutako aldagaiak*: Diseinuaren bitartez kontrolaturiko aldagaiak dira, edo haien eragina ezabaturik izaten da. Kontrol prozedura esperimentalak nahiz estatistikoa izan daiteke. Edozein kasutan, kontrol prozeduren helburua *zorizkatutako aldagaiek* sortutako *zorizko erroreak* edo *aldagai nahasgarriak* eragindako *alborapena* ekiditean datza.

- *Aldagai nahasgarriak*: Kontrolatu gabeko aldagai arrotzak izanik, *aldagai azalgarriekin* nahas daitezke. Ikerketaren alborapena ekar dezakete.

- *Zorizkatutako aldagaiak*: Zehaztutako aldagaiak dira, ezagunak alegia. Hala ere, ez dira *aldagai azalgarri* gisa hartzen. Diseinuetan *zorizkatzea* kontrol prozedura jakin bati esaten zaio.

Diseinu eraginkorretan aldagai arrotzak ahalik eta gehienetan *kontrolatuak* izan behar ditugu. Testuliburu honetan bertan *aldagai arrotzen* eragina ekiditeko edo kontrolatzeko zenbait prozedura azaltzen dugu.

Atal honetan darabilgun adibideari berriro lotuz, honela adieraz genitzake zenbait aldagai:

1.1.- Taula: *Diseinuetarako aldagaien sailkapena*

ALDAGAI AZALGARRIAK		ALDAGAI ARROTZAK		
Aldagai azalgarriak	Aldagai azalduak	Kontrolaturiko aldagaiak	Aldagai nahasgarriak	Zorizkatutako aldagaiak
- Euskaldunen proportzioa - Biztanleko diru etekinak - Publizitatea - Teknikari kopurua - Teknikarien prestakuntza maila ... / ...	- Euskaldungoa-ren hazkuntza - Euskararen erabilera - Euskararen ezagutza /	- Kultur maila - Ikasketa maila ... / ...	- Euskararekiko jarrerak - Hizkuntza gaitasuna ... / ...	- Adina - Sexua ... / ...

Aldagai zerrenda horrek, jakina, ez du deus ere argitzen diseinu jakin bati lotzen ez bagatzaizkio. Ikusten denez, zenbait aldagai kategoria desberdinetan errepikaturik ager dakizkiguke. Ikertzaileak zehaztu behar du darabilen eredu eta, horri atxikiz, diseinuaren *aldagai azalgarriak* zein diren adierazi. Irakurlearen esku uzten dugu, ariketa gisa, 1.1. taulako aldagaiak kategoriaz aldatzea, beste zenbait aldagai gehitzea edo dauden batzuk kentzea. Edozelan ere, ikerketaren esparruak eta diseinuak berak erabili beharreko aldagaiak eta haien izaera zehazten dute.

1.6. JOMUGAREN TEOREMA

Portaera zientzietan gertaera zehatza alde zurretik atzematea ezinezkoa dela esan beharrik ez dago, beharbada guztion zorionerako. Ez dakigu nork egingo duen bere buruaz beste, ez egoera etsigarrian nor gaixotuko den, ez eta nortasuneko nahasketak nori ekarriko dion portaera desegokia ere. Hala ere, psikologok bere buruaz beste egin ohi dutenen *batazbesteko profila* ezagutzen dugu, etsipenari eusten dionaren ezaugarri nabarmenenak badakizkigu eta zenbait portaera desegokiren itzalean nortasun nahasketa jakinak ezkututzen direla ikasiak gaude. Azken buru, kasu bakanduen pronostikoa egiterik ez badugu, multzo handietan edo, zehatzago, kasuen zenbaki handietan zer gerta daitekeen esateko moduan gaude.

Zenbaki handien legeak auzi hori azaltzen digu. Gertaera jakin baten probabilitatea eta gertaera horren *maiztasun erlatiboan* (proportzioa) arteko diferentziak zerora jotzen du. Horra *zenbaki handien legearen* muina.

Saiakuntzan, oro har, zenbat eta behaketa gehiago erabili, **benetako balioetatik** orduan eta **gertuago** ibiliko gara. Bildutako neurketa edo balio multzoaren batazbesteko aritmetikoak benetako baliorantz jotzen du neurri kopurua edo multzoaren tamaina handitu ahala. Are gehiago, behin eta birritan neurtuz gero,

behatutako neurketaldietako batazbesteko puntuazioak era jakinean banatzen direla frogatu izan da. Banaketa hori kanpai itxurako irudiaz adierazi ohi da; XIX. mendean Karl Friedrich Gausssek deskribatu zuen banaketa normalaz ari gara, hain zuzen.

Normaltasunaren kontzeptua bada, hein handian behintzat, oinarri estatistikoa duen kontzeptua da. Psikologia eremuko aldagai gehienak banaketa normalari doitzen zaizkio, banaketa normalaren probabilitate-funtzioari darraiote alegia.

Esanak esan, **jomugaren teorema** eta banaketa normalaren ekarpena ez dira deskribapen mailara mugatzen. Gaur egun, horren erabilera funtsezkoa da estatistika inferentzialean. Izan ere, psikologian darabiltzagun neurketa eta neurgailu guztien oinarria baita. Hau da, izaera psikologikoko aldagaiak, gehienak behintzat, ez daukagu zuzenean neurtzerik; zeharkako neurketak darabiltzagu beraz. Are gehiago, izaera psikologikoko aldagai gehienak banaketa, *banaketa normalaren probabilitate-funtzioari* lotzen zaizkiola jakin baino, beraiek neurtu ahal izateko egiten dugun suposizioa da. Izan ere, psikologian arazo larriak baititugu, adimena, oroimena, herstura, etab. tankerako aldagaiak neurtzeko. Ditugun neurgailuak egunez egun hobetzen badira ere, ezin erabat zehatzak izan. Neurketa errorea ohikoa da beraz. Hala ere, *neurketa errorea zorizkoa* dela jakinik, errorearen tamaina eta eragina kontrola ditzakegu.

Jomugaren teorema zertan den gogoratzeko, berau jarraian aurkeztu dizuegu. *Bitez $X_1, X_2, X_3, \dots, X_i$, Bernouilli motako $\{X_i\}$ aldagai sorta, non aldagai guztiek np batazbesteko eta npq bariantza mugatuak dauzkaten; era berean, B_n deitutakoa zorizko aldagai batura daukagu edo, $\{X_i\}$ aldagai sortaren zorizko batura esan genezakeena.*

Hortik abiatuta, Y zorizko aldagai tipifikatu berria defini dezakegu honela:

$$Y = \frac{B_n - np}{\sqrt{npq}}$$

Zorizko aldagai tipifikatu hori (Y), *jomugan*, $N(0,1)$ banaketa normal estandarizatuari doitzen zaio. Zehatz-mehatz, bere banaketa-funtzioa honako hau da:

$$F(z) = \int_{-\infty}^z \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt$$

Jomugaren teorema, oro har, honela planteatu dezakegu: zorizko aldagaiak elkarrekiko askeak direnean eta banaketa-funtzio berari jarraitzen diotenean, $\mu =$

$E(X_i)$ eta $\sigma = V(X_i)$ batzbesteko aritmetiko eta desbidazio estandarra dutenean, hurrenez hurren, orduan, i edozein baliotarako ondorengoa betetzen da:

$$\lim_{n \rightarrow \infty} \frac{B_n - a_n}{b_n} \rightarrow N(0, 1)$$

non, a_n eta b_n , hurrenez hurren,

$$\begin{aligned} a_n &= n\mu \\ b_n &= \sqrt{n\sigma^2} \end{aligned}$$

Beste era batera, edozein Z aldagai normal tipifikaturen kasuan hurrena betetzen da:

$$\lim_{n \rightarrow \infty} P\left(\frac{B_n - n\mu}{\sqrt{n\sigma^2}} \leq Z\right) = F(z) = \frac{1}{2\pi} \int_{-\infty}^z \exp\left(-\frac{1}{2}t^2\right) dt$$

Azken buru, aldagaiaren banaketa-funtzioa normala bada, laginaren tamaina (n) zenbat eta handiagoa izan, zentru joerako estatistikoak eta populazioaren parametroak gero eta gertuago izaten dira.

Ikusten denez, jomugaren teorema eta banaketa normalaren ekarpena ez dira deskribapen mailara mugatzen. Estimazio teorian jomugaren teorema funtsezko teorema da, hots, estatistika inferentzialaren oinarrietariko bat da. Horrexegatik, hain zuzen, saiakuntza eta sasisaiakuntza diseinuetako analisietan oinarritzkoa da. Are gehiago, saiakuntza eta sasisaikuntza diseinuen errorearen kontrolerako funtsezkoa da; gure behaketetako erroreak (berez, *zorizko aldagai normalak izanik*) ez balitzaizkio banaketa normalari doituko, alborapenen bat badagoela erdietsi beharko genukeen ezinbestean.

1.7. FISHER-SNEDECORREN LEGEA

Sir Ronald Aylmer Fisher (1890-1962) matematikariak beste urrats bat eman zuen. Psikologian azterketa estatistikoetan ekarpen erabiliena: bere deitura daraman banaketa (askotan Snedecorren F banaketa deitua ere), funtsean, saiakuntza eta sasisaiakuntza diseinuetako analisien test estatistikoa da.

Jakina denez, homozedastizitatea (bariantzen arteko homogenotasuna) bariantzen arteko baliokidetasun adierazlea da, hots, homozedastizitatea izatekotan populazio bereko taldeak direla esan genezake. Bariantzen arteko erkaketa hori *bariantza analisiaren* muina da.

Orokorrean, honela dio Fisher-Snedecorren legeak: *bitez X eta Y bi zorizko aldagai elkarrekiko askeak, zeintzuk χ^2 legeari darraioten, p eta q askatasun graduekin hurrenez hurren. Horrela, $F = \frac{X/q}{Y/p}$ aldagaiaren probabilitate legea*

Fisher-Snedecor legea edo F legea da, non berorren dentsitate-funtzioa hurrengo den:

$$f(x) = \frac{\Gamma\left(\frac{p+q}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)} p^{\frac{q}{2}} q^{\frac{p}{2}} \frac{\chi^{\frac{p}{2}-1}}{(px+q)^{\frac{p+q}{2}}} \quad \text{non } x > 0$$

Funtzio horrek zehazten ditu gure artean horren ezagunak diren *F motako banaketak*. Funtzioa ezagutzeak ohiko (*zorizkoa* esan beharko genuke) banaketa ezagutzea dakarrelarik, erkaketarako irizpide gisa erabil dezakegu. Hau da, guk behatutako Fak elkarren baliokideak diren ala ez ebazteko erabil daiteke.

Azken buru, guri gagozkiolarik, Fisherren F bariantzen bi estimazioen arteko zatidura besterik ez da. Bere bitartez, test sinplenean, bi talde elkarren artean erkatzen ditugunean beraien bariantzak jatorri bereko populazioarenak ote diren jakin nahi dugu.

Ildo beretik, bariantza analisiari (ANOVA) atxikitako testa F motakoa da. Saiakuntza eta sasisaiakuntza diseinuetan tratamenduak azaldutako bariantza esanguratsua denentz jakiteko erabili ohi da. ANOVAREN ohiko test estatistikoa honakoa da:

$$F = \frac{MC_{arte}}{MC_{barne}} \quad (1.1)$$

MC (*medias cuadráticas*) bariantza gisa ulertu behar da. Zehatzago, MC_{arte} taldearteko berreduren baturen batzbestekoa da, eta MC_{barne} , berriz, taldebarneko berreduren baturen batzbestekoa. Esan bezala, biak dira bariantzak, lehena, tratamenduak azaldutakoari egozten dioguna (zenbatzailea) eta, hurrena, erroreari egotzitakoa, hondakina deitua, edo tratamenduak azaltzen ez duen bariantza (izendatzailea).

Horrela, Fisher-Snedecorren legeari jarraituz, diseinuen analisisetan kalkulaturako bariantzen arteko zatidurak F banaketari doitzen zaizkio. Behin zenbatzaile eta izendatzailearen *askatasun graduak* ezaguturik, dagokion probabilitate-funtzio zehatza zein den jakin dezakegu. Hau da, probabilitate-funtzio horren arabera ${}_a F_{v_1/v_2}$ balio zehatza ezaguna da. Balio horrek hipotesi nulupeko testaren onartze eta ukatze barrutiak bereizten ditu, hots, muga kritikoa zehazten du. Gogoan izan ANOVARI atxikitako testa norabide bakarrekola dela (eskubialdekola) eta, horregatik, muga kritiko bakarra dugula.

Saiakuntza eta sasisaiakuntza diseinuetan behatutako F-a, gure datuetatik abiatuz kalkulatzeko duguna alegia, muga kritikoa baino txikiagoa denean hipotesi nuluari (H_0) eusten diogu eta, aitzitik, berdina edo handiagoa denean H_0 baztertu eta, ondorioz, ordezkoko hipotesia (H_1) onartzen dugu.

1.8. ASKATASUN GRADUAK ETA EFEKTUAREN TAMAINA

Arestian aipatu dugunez, bariantza analisisia da diseinuetan erabakiak hartzeko erabiltzen den tresna. Esan dugu baita ere, analisi horren muina aldakortasunaren azterketa dela; are zehatzago, aldakortasun orokorraren deskonposaketa.

Menpeko aldagaia(k) aztertzen d(it)ugunean talde esperimental desberdinetako puntuazioak erabiltzen ditugu. Oro har, menpeko aldagaiak, beste edozein aldagaia bezala, badu bariantza bat. Bariantza hori, jakina, darabilgun laginari dagokio, ez populazioari; estatistiko bat da alegia. Beste alderditik, badakigu lagineko bariantza populazioko parametroaren estimatzaile gisa alboratua dela, hots, zorizko lagina zenbat eta txikiagoa izan, bariantza parametrotik orduan eta nabarmenago urruntzen da.

Gertaera hori erraz antzeman daiteke bariantza datuen sakabanatze neurketa dela ulertzen bada. Populazioko parametroa estimatzeko lagineko *kuasibariantza* erabili ohi da, bariantza doitua nolabait esatearren.

$$s^2_{n-1} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}$$

Kuasibariantzaren formularen, izendatzailean, unitate bateko zuzenketa ikus daiteke ($n-1$); zuzenketa horrek, jakina, ez dakar berez, estimazioaren erabateko zehaztasuna. Egia da, lagineko s^2_{n-1} edo kuasibariantza populazioko σ_x^2 edo bariantzaren estimagailu alboragabekoa dela, baina ezin esan estimagailu *alboragabekoaren* eraldaketa ez-linealak erabateko zehaztasuna dakarrenik. Estimazioaren alborapena, besteak beste, populazioaren banaketak zehazten du. Populazioa normala bada lagineko kuasibariantzaren, s^2_{n-1} , laginketa banaketako batzbesteko aritmetikoa populazioko bariantza (σ_x^2) baino zertxobait txikiagoa da. Zehatz-mehatz:

$$E(s) = \mu_s = \left(\frac{4n-4}{4n-3} \right) \sigma$$

Ikusten denez, n handia denean *alborapena* txikia izaten da, $\left(\frac{4n-4}{4n-3} \right)$ koefizienteak dakarrenaren araberakoa, hain zuzen.

Nolanahi ere, jakin nahi dugun balioarekiko *hurbilpena* besterik ez dugu. Aztertutako kasuan dimentsio bakarra erabili dugu, datu multzo jakin baten *sakabanatze neurriarena*, kasurik sinpleena, beraz. *Askatasun graduen* kontzeptua berriz, fisika zientzietatik datorrigu. Zientzia horretan, gurean bezala, dimentsio bat edo gehiago azter daiteke. Esaterako, mugimenduak aztertzen ditugunean, fisika zientzietan *askatasun gradu* desberdinak egozten zaizkie mugimendu mota desberdinei. Hau da, mugimendua lerro zuzeneko edo dimentsio bakarrekoa denean, *askatasun gradu* bakarrekoa da. Plano batetik bada, bi dimentsio ditu eta bi *askatasun gradu*, beraz. Azkenik, pilotalekuan pilota hiru dimentsioetatik ibil daiteke eta, horregatik, hiru *askatasun graduko* dela esaten dugu. Hortik hartua da estatistikarako, hain zuzen, *askatasun graduen* kontzeptua.

1.8.1. ASKATASUN GRADUAK ZORIZKO TALDE ANITZEKO DISEINUAN

Demagun hezkuntza psikologiaren eremuan ikaskuntza estrategia desberdinek testu baten oroimenean izan dezaketen eragina aztertu nahi dugula. Behin testuen ikaspide prozedura desberdinak zeintzuk diren aztertu eta gero, ondorengo hiru *estrategia* hauek hautatu ditugu¹:

Y_1 , testua azpimarratzea,

Y_2 , edukiari lotutako eskemak burutzea eta

Y_3 , testuan agertzen diren ideia nagusien zerrenda egitea.

Hiru *maila* horien arabera *ikaskuntza* gauzatu ondoren subjektu guztiei *oroimen froga (irizpide aldagaia)* betearazten zaie. Jarraian dagoen taulan, *maila* desberdinetan subjektuek gogoratutako ideia kopuruak beha daitezke:

MAILAK (Tratamenduen arabera)				
	Y_1	Y_2	Y_3	
	15	30	32	
	10	32	34	
	4	25	40	
	8	27	37	
	12	35	28	GUZTIRA
Baturak	49	149	171	369
Batazbestekoak	9,8	29,8	34,2	24,6
Behaketa kopurua	5	5	5	15

Ariketa honetan eragin finko eta faktore bakarreko bariantza analisisa ezarri behar dugu. Hau da, aldakortasun orokorra (SCT) bi osagaitan adieraz daiteke: *taldearteko* aldakortasuna (SCA) eta *taldebarneko* aldakortasuna (SCR). Hiru maila eta hiru talde ditugula kontutan harturik, oro har, hurrengo bete behar da:

$$\frac{\bar{Y}_1 + \bar{Y}_2 + \bar{Y}_3}{k} = \bar{Y}_{..} \quad \text{gure ariketari ezarritz,} \quad \frac{9,8 + 29,8 + 34,2}{3} = 24,6$$

¹ Isasi, X. eta Balluerka, N. (1995). *Datu Analisisa eta Saiakuntza Diseinuak Portaera Zientzietan*. Bilbo: UEU (274-277 or.).

Beste era batera, taldeen batzbestekoen batzbesteko aritmetikoa datu guztien batzbesteko aritmetiko orokorra da. Behin batzbesteko aritmetiko orokorra kalkulaturik, taldez taldeko batzbesteko aritmetikoak kalkulatzeari lotuko bagintzaizkio, azken taldeko batzbesteko aritmetikoa kalkulatu aurretik bere balioa finkaturik legoke. Izan ere, lehen adierazitako ekuazioa bete beharra dago. Diseinu hauetan *taldearteko askatasun graduak aldagai askearen maila* kopuruaren arabera ezarrita daude. Hots, *k maila* dagozkien batzbesteko aritmetikoetatik abiatuz, aurreko ekuazioa bete dadin, gutxienik bat aldeztu aurretik finkatuta dago. Ondorioz, *taldearteko askatasun graduak* maila kopurua ken bat ($k-1$) eragiketaren bidez kalkula ditzakegu. Adierazitako irizpide hori bera ezar dakioko *taldearteko berreduren baturari* (SCA), zeinaren arabera *taldearteko berreduren batzbestekoa* (MCA) kalkulatu behar dugun. Formuletara ekarrita:

$$SCA = \left[\frac{1}{n} \sum_{j=1}^k \left(\sum_{i=1}^n Y_{ij} \right)^2 \right] - \left[\frac{1}{kn} \left(\sum_{j=1}^k \sum_{i=1}^n Y_{ij} \right)^2 \right] \quad (1.2)$$

$$MCA = \frac{SCA}{k-1} \quad (1.3)$$

Mailaz mailako azterketa eginez gero, talde bakoitzaren batzbesteko aritmetikoa kalkulatu beharko genuke. Honela hain zuzen,

$$\frac{Y_{11} + Y_{21} + \dots + Y_{i1} + \dots + Y_{n1}}{n} = \bar{Y}_{.1} \text{ lehenengo talderako,}$$

$$\frac{Y_{12} + Y_{22} + \dots + Y_{i2} + \dots + Y_{n2}}{n} = \bar{Y}_{.2} \text{ bigarren talderako eta,}$$

$$\frac{Y_{13} + Y_{23} + \dots + Y_{i3} + \dots + Y_{n3}}{n} = \bar{Y}_{.3} \text{ hirugarren talderako.}$$

Kasu honetan, arestian esandako irizpide bera ezar dezakegu. Hau da, taldeen batzbesteko aritmetikoak ezagutzen ditugularik ($n=5$), kalkuluak egiteko mailako lau ($n-1$) osagai *aske* egin genitzake. Nolanahi ere, talde kopurua kontutan hartu beharra dago. Baldin eta hiru talde baditugu eta taldeko unitateko zuzenketa egiten badugu, darabiltzagun *askatasun gradu* kopurua $k(n-1)$ izango da edo, beste

era batera esanda, *orotariko behaketa kopurua ken maila kopurua*. Hori da, hain zuzen, *taldebarneko askatasun gradu kopurua*.

Taldearteko kasuan bezala, aldakortasunaren eta bariantzaren kalkuluetarako formulak ekartzen dizkizuegu:

$$SCR = \left[\sum_{j=1}^k \sum_{i=1}^n Y_{ij}^2 \right] - \left[\frac{1}{n} \cdot \sum_{j=1}^k \left(\sum_{i=1}^n Y_{ij} \right)^2 \right] \quad (1.4)$$

$$MCR = \frac{SCR}{k(n-1)} \quad (1.5)$$

Honaino bada, *askatasun gradu*ei buruzko azalpenak. Irakurleak ariketaren urratsez urratseko ebazpena nahi izango balu, bietako bat edo irakurleak berak egin lezake edo, bestela, lehen aipatutako liburuan (ikus 1. oinoharra) ebatzita aurki lezake.

Gainerako diseinuetan, orokorrean, bertan adierazitako irizpideak ezar daitezke unean uneko *askatasun graduak* zenbat diren jakiteko. Auzi honetan gehiago sakontzea testuliburu honen xedetatik kanpo geratzen delarik, ezarritako ildoari lotuko gatzazkio aurrerantzean.

1.8.2. EFEKTUAREN TAMAINA ETA TESTAREN INDARRA.

Saiakuntza diseinuetarako analisi estatistikoetan badira beste zenbait kontzeptu oinarritzkoak. Testuliburu honen edukiarekin jabetzeko funtsezkoak direlakoan, hipotesi estatistikoak, hipotesi testa, erroreak, testaren indarra eta efektuaren tamaina kontzeptuak, hurrenez hurren, hona ekartzeari komenigarri iritzi diogu.

Hipotesi estatistikoak bi dira: *hipotesi nulua* (H_0) eta *ordezko hipotesia* (H_1).

HIPOTESI NULUA (H_0): funtsean onartu edo ukatu nahi dugun hipotesi estatistikoa da. Test estatistikoak normalean *kontserbadoreak* direlarik, *hipotesi nulua*k aztergaian aldaketarik ez dagoela adierazi ohi du.

Estimatzailaren laginketa banaketa H_0 -aren arabera eratzten da, testa gero gauzatzeko.

ORDEZKO HIPOTESIA (H_1): hipotesi nuluari kontrajartzen zaion eta aldi berean beraren osagarria den hipotesi estatistikoa da. Honela idatz dezakegu:

$$p(H_0) + p(H_1) = 1$$

HIPOTESI TESTA: hipotesi estatistikoaren inguruan *laginketa banaketari* atxikiz eta probabilitate legeen menpe egiten den eraketa estatistikoari deritzo.

Hipotesi testaren oinarria definitutako hipotesietan datza beraz. Aurrenik hipotesi nulua (H_0) zehazten dugu, hipotesi nulua arabera ordezeko hipotesia (H_1), eta azkenik, *I. motako errore maila* (α) finkatu ondoren testari ekiten diogu.

Probabilitate legeen menpe gaudelarik, bada *zoriak dakarren arriskua edo errorea*. Oro har, hipotesi testei atxikitako erroreak bi motatakoak dira: I. motako errorea eta II. motako errorea. I. motako errorea H_0 -a benetakoa izanik berau baztertzean datza. II. motako errorea, berriz, H_0 -a faltsua izanik berau onartzean datza. Batzuetan erroreak direla eta, aho korapilotan gabiltzala ematen du. Hori ekidin nahiean, hurrengo kontingentzia taulan kontzeptu hauek argiago aurkezten dizkizuegu:

1.2. Taula: Hipotesi estatistikoak eta erabakiari atxikitako erroreak

HIPOTESIEN IZAERA

ERABAKIAK	H_0 -a BENETAKOA	H_1 -a BENETAKOA
H_0 -ari eutsi (d_1)	ERABAKI <u>ZUZENA</u>	ERABAKI <u>OKERRA</u> II. MOTAKO ERROREA (β)
H_0 -a ukatu (d_2)	ERABAKI <u>OKERRA</u> I. MOTAKO ERROREA (α)	ERABAKI <u>ZUZENA</u>

Ohikoa ez izanagatik *erabakiak* Hipotesi Nuluari lotuta agertu ditugu, izan ere, erabakia hipotesi horren inguruan hartzen baita. Ordezko Hipotesia onartzea, Hipotesi Nuluaren ukazioaren ondorio gisa ulertu behar dugu beraz. Gogoan izan

hipotesiak elkarren kontrakoak eta osagarriak direla eta, ondorioz, bat betetzen ez bada besteak bete behar du ezinbestean.

Edozelan ere, errore bakoitzari *probabilitate maila* jakin bana dagokio. Taulan agertzen diren α eta β , baldintzazko probabilitateak besterik ez dira:

$$\alpha = p(d_2 | H_0)$$

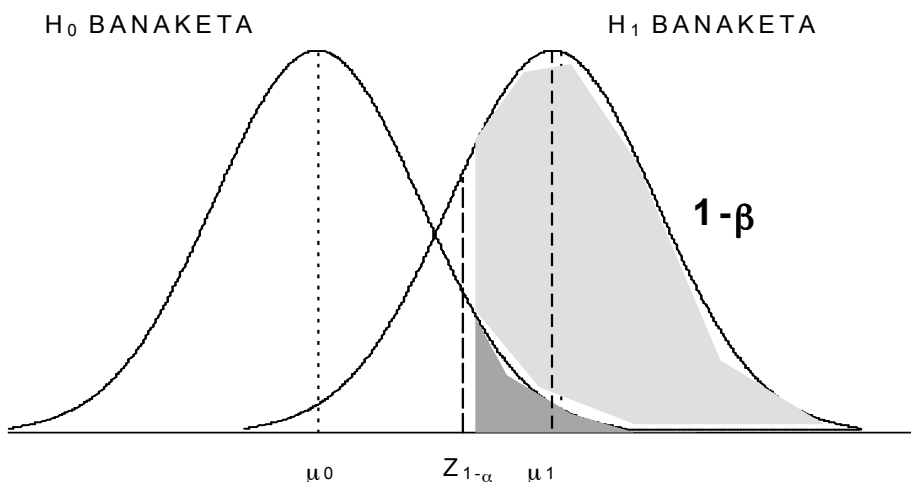
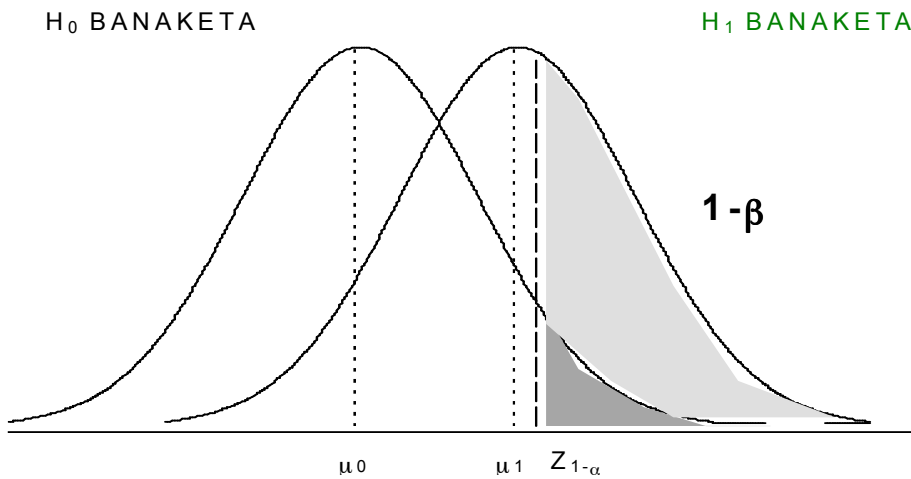
$$\beta = p(d_1 | H_1)$$

I. motako erroreari dagokion probabilitateari (α) *esangura maila* esaten diogu. Probabilitate eremu hori Hipotesi Nulupeko *ukatze barrutiari* dagokio. Analisisetan, ahalik eta txikiena izatea daukagu abiaburu. II. motako erroreari dagokion probabilitatea (β), berriz, Ordezko Hipotesipeko banaketari dagokio.

I. eta II. motako erroreei dagozkien probabilitateak (α eta β) eta laginaren tamaina (n) elkarren artean erlazionaturik daude.

Ordezko hipotesia onartzeari, benetako hipotesia delakoan, *testaren indarra* esaten diogu, berari dagokion probabilitatea ($1-\beta$) delarik.

Demagun *alde bakarreko testa* dela, eskubialdekoa hain zuzen ere, orduan batzbesteko aritmetikoaren inguruan (μ_0 , μ_1) hipotesi bakoitzari dagokion irudia marraz dezakegu.



1.1. Irudia: Testaren indarra ($1-\beta$) hipotesi testetan.²

Irudietan agerienez, *testaren indarra* ordezeko hipotesiaren aldeko erabakiari dagokio. Behin hipotesi nulua ukatuta, bere probabilitateak ($1-\beta$) baleko erabaki zuzenaren indarra adierazten du. Jakinaenez, probabilitate maila honetan, darabilgun *esangura maila estatistikoak* (α) eragiten du. Era berean, laginaren tamaina edo saiakuntzako subjektu kopuruak eragiten dute.

Esaten genuenez, *testaren indarrean* eragiten duen beste faktore bat *laginaren tamaina* da. Horrela, α -ren balioa aldeztu aurretik finkaturik eta μ_0 eta μ_1 balioak ezarrita, *testaren indar maila* jakin baterako behar den laginaren tamaina honela zehaz dezakegu:

$$n = \frac{\sigma^2 (z_\alpha + z_\beta)^2}{(\mu_0 - \mu_1)^2}$$

Edozelan ere, saiakuntza diseinuetan analisirik hedatuena *bariantza analisia* (ANOVA) delarik, beste kontzeptu bat erabili ohi da: *efektuaren tamaina*. *Efektuaren tamaina* aldagai askeak eragindako aldakortasun azaldua da (η^2). Bariantza analisiaren testuinguruan, oro har, honela kalkulatzen da:³

$$\eta^2 = 1 - \frac{SCR}{SCT} \quad (1.6)$$

Bariantza analisisian erabiltzen den testaren (F) eta *efektuaren tamainaren* artean erlazio zuzena dago. Ikus dezagun aurrenik nola kalkulatzen den F-a:

$$F = \frac{\frac{SCA}{lg_{arte}}}{\frac{SCR}{lg_{barne}}} \quad (1.7)$$

Ikus daiteke *berreduren baturak* (SCA eta SCR) eta *askatasun graduak* (lg_{arte} eta lg_{barne}) formularen oinarriak direla. Bestalde, *efektuaren tamaina* kalkulatuz:

$$\eta^2 = \frac{SCA}{SCT} \Rightarrow SCA = \eta^2 SCT = \eta^2 (SCA + SCR) \Rightarrow$$

$$SCA = \eta^2 (SCA + SCR) \Rightarrow SCA - \eta^2 (SCA + SCR) = 0$$

$$SCA - \eta^2 SCA - \eta^2 SCR = 0 \Rightarrow$$

² Irudiak eta aurreko pasarteak *Datu Analisia eta Saiakuntza Diseinuak Portaera Zientzietan* (UEU, 1995) testuliburutik jasoak dira.

³ Ikus aurreko (1.1), (1.2), (1.3), (1.4) eta (1.5) formulak.

$$\begin{aligned} SCA(1 - \eta^2) - \eta^2 SCR &= 0 \Rightarrow SCA(1 - \eta^2) = \eta^2 SCR \\ \frac{SCA}{SCR} &= \frac{\eta^2}{1 - \eta^2} \end{aligned} \tag{1.8}$$

Azkenik, (1.7) formularen (1.8) formularen emaitza ordezkatzuz⁴:

$$F = \frac{\eta^2}{1 - \eta^2} \frac{lg_{barne}}{lg_{arte}} \tag{1.9}$$

Honaino bada, *balideziaz* hitz egin aurretik laburbildutako kontzeptuen zerrenda.

⁴ Alderdi honetan gehiago sakontzeko ikus: Pascual, J.; García Pérez, J.F.; Frías Navarro, M.D. (1995). *El diseño y la investigación experimental en psicología*. Valencia: C.S.V.

2.- Balidezia kontzeptua eta balideziaren sailkapena ikuspegi desberdinen arabera

2.1.- Balidezia kontzeptua

2.2.- Balideziaren sailkapenari buruzko ikuspegi ezberdinak

2.2.1.- Campbell eta Stanleyren eredua

2.2.2.- Cook eta Campbellen eredua

2.2.3.- Kruglanski eta Kroyren eredua

2.2.4.- Cronbachen eredua

2.2.5.- Sailkapen klasikoez bestelako bi ikuspegi:

Marken eredua eta Chen eta Rossiren eredua

2.3.- Saiakuntza eta sasisaiakuntza diseinuaren eremuan sistematizazio hedatuena: Cook eta Campbellen eredua

2.- BALIDEZIA KONTZEPTUA ETA BALIDEZIAREN SAILKAPENA IKUSPEGI DESBERDINEN ARABERA

2.1.- Balidezia kontzeptua

Bai saiakuntza eta bai sasisaiakuntza eremuan ere, ikerketa orok jarraitzen duen helburua inferentziak egin eta ondorio zehatz, fidagarri eta, batez ere, baliagarriak lortzean datza. Gure eskuhartzearen eraginkortasuna eta garrantzia, neurri handi batean behintzat, balideziaren arabera ebaluatzen denez gero, gai hau berau ikerketa diseinuaren edo plangintzaren alderdi nagusienetariko bat bezala har daiteke.

Balidezia kontzeptua ikerketa diseinuaren testuinguruan modu serio eta sistematikoan landu zuten lehen autoreak Campbell eta bere laguntzaileak izan ziren (e.b. Campbell, 1957, 1969; Campbell eta Stanley, 1963, 1966; Cook eta Campbell, 1976, 1979). Egile hauek guztiak honela definitu zuten *balidezia*: egiazko adierazpenak egiteko daukagun hurbilpenik egokiena.

Campbell eta bere laguntzaileek proposatutako definizioan agertzen den "hurbilpen" hitza zorrotzasun gehiagoz edo gutxiagoz hartzen badute ere, balideziaren gaia aztertu duten ikertzaile gehienek adiera berberari heltzen diote funtsean. Honela, Gottsdankerren (1978) ustez, esperimentuak are eta baliagarriagoak dira esperimentu bikainari gehiago hurbiltzen zaizkion neurrian, hots, egiazko hipotesiaren eta gezurrezko hipotesi lehiakideen artean inolako errorerik gabe bereizteko aukera ematen duen esperimentuari gehiago hurbiltzen zaizkion heinean (Keppel, 1973). Egungo zenbait autorek, balidezia edozein ikerketatan lortutako emaitzak egiazkoak diren ala ez jakiteko izan dezakegun konfidantza mailarekin lotzen dute (Ato, 1991), edo, esperimentuak ikuspegi teknikitik aurkezten duen sendotasun mailarekin (Arnau, 1990), edo, beste zenbait adieren artean, ikerketaren diseinuak ondorio etsigarriak eta sinesgarriak lortzeko eskaintzen duen aukera mailarekin (Maxwell eta Delaney, 1990).

Ikerketa eremu honetan murgildurik dauden autore gehienek balidezia kontzeptuari buruz adostasun handia agertzen badute ere, ez da gauza bera gertatzen balideziaren sailkapenari dagokionez. Campbell eta Stanleyk (1963, 1966) proposatutako aurreneko sailkapenak aldaketa gehiago edo gutxiago jasan ditu bai barne ematen dituen balidezia mota desberdinei dagokienez bai eta balidezia mota

bakoitzaren esanahiari dagokionez ere. Hurrengo atalean egile desberdinek balideziari buruz aurkeztutako sailkapenen berrikusketa aurkezten dizuegu. Berrikusketa honen helburua ez datza sailkapen bakoitza xeheki aztertzean, ikerketa diseinuaren inguruan aurkeztu diren sailkapen nagusienak oro har aurkeztean baizik. Ondoren, psikologiaren esparru metodologiko gehienetan onartua izan den sailkapena edo eredua, Cook eta Campbellen (1976, 1979, 1986) eredua alegia, sakonki aztertuko dugu.

2.2.- Balideziaren sailkapenari buruzko ikuspegi desberdinak

2.2.1.- Campbell eta Stanleyren eredua

Ikerketa diseinuaren balideziari buruzko lehen sistematizazioa Campbell (1957) hasitako eta berorrek bakarrik (Campbell, 1963, 1969) nahiz bere laguntzailea izandako Stanleyrekin batera (Campbell eta Stanley, 1963, 1966) garatutako lanetan koka daiteke. Ikertzaile hauek W.A. McCallen (1923) "How to experiment in education" izeneko testu klasikoa hartu zuten haien lanen abiaburu. Liburu horretan McCallek honela adierazten zuen: *fisheriar* korrontearen eraginez analisi estatistikoaren esparruan gertatutako garapena guztiz garrantzitsua izan zen bitartean, datuak modu egokian biltzeko prozedurek jorratu gabe jarraitzen zutela oraindik. Premia honetatik abiatuz eta Fisher (1925) eta McCallen (1923) ikusmolde batzuk osatzeko asmoz, Campbell eta Stanleyk (1966) balideziaren aztertzeari ekin zioten eta ikerketaren alderdi nagusi hau sakonki aztertu zuten saiakuntza-aurreko, saiakuntza eta sasaisiakuntza diseinu mota desberdinak oinarri gisa hartuz. Haserako planteamendu honetan, egileek bi balidezia mota nagusi bereiztu zituzten, hots:

- 1.- *Barne balidezia*, hau da, ustezko zergati gisa jokatzeko duen aldagai baten edo gehiagoren (tratamendua delakoaren) eta ustezko eragindako gisa jokatzeko duen aldagai baten edo gehiagoren (jokabide behagarria delakoaren) arteko erlazioa ikertzaileak proposatutako moduan interpretatzeko daukagun konfidantza maila. Tratamendua ezik, jokabide behagarrian eragiten duten efektu sistematiko guztiak kontrolatuz lortzen da barne balidezia eta ikerketaren *sine qua non*-a bezala har daiteke.
- 2.- *Kanpo balidezia*, hots, aldagaien arteko ustezko kausa erlazioa eragile eta eragindako beste konstruktutaz batzutaraz eta pertsona, testuinguru nahiz epe desberdinetara zabaldu edo orokortu daitekeen maila.

2.2.2.- Cook eta Campbellen eredua

Geroko lanetan, Cook eta Campbell (1976, 1979, 1986) bi balideziaz eratutako eskema hau zabaldu zuten eta Campbell eta Stanley (1963, 1966) proposatutako balidezia mota bakoitzean beste azpisailkapen bat egin zuten. Lehenik eta behin, bi edo aldagai gehiagoren artean kausazko erlazio funtzionalak inferitzeko beharrezko baldintza elkarren arteko kobariazioa izan beharko litzatekeela planteatu zuten. Horrela, kobariazioari buruzko ondorioak ezagutza estatistikoetan oinarritzen direnez gero, *ondorio estatistikozko balidezia* edo *balidezia inferentziala* izeneko balidezia mota berri bat sortu zuten. Balidezia mota honekin zera adierazi nahi da: esanguratsutasun estatistikozko testetatik abiatuz zein konfidantza mailaz baieztu daitezkeen, bi edo aldagai gehiagoren artean kobariazioa badagoela egon. Balidezia mota hau sarritan erlazionatu izan da *diseinuaren zorrotasunarekin*, hots, tratamenduek eragina izanez gero, diseinuak eragin hori sumatzeko duen *indarrarekin* edo zehaztasun mailarekin.

Estatistikaren arabera, ikertzailea, bi edo aldagai gehiagoren artean kobariazioa ezartzeko gai bada, bere buruari galdetu behar dio ea erlazio horrek izaera kausala ote duen eta sumatutako efektuak berorrek eragile gisa planteatu duen aldagaiari soil-soilik ote dagozkion. *Barne balideziaren* funtsa galdera hauetan kokatzen da eta honela defini daiteke: X eta Y aldagaiak operazionalki manipulatuak edota neurtuak izan diren modua kontuan hartuz, elkarren artean eragile-eragindako motako erlazioa ezartzeko izan dezakegun konfidantza maila bezala.

Barne balidezia operazionalki manipulatuak eta neurtuak izan diren aldagaien arteko erlazioari dagokio, baina ez da arduratzen aldagai horiek kontzeptualki edo teorikoki adierazten dutenaz. Alabaina, behar beharrezkoa da daitekeen eragile-eragindako erlazioa testuinguru teoriko batean kokatzea. Cronbach eta Meehle (1955) neurketa eremuan erabilitako terminologiari atxikiz, Cook eta Campbell honela definitzen dute *eragile eta eragindakoen konstruktua balidezia*: bai manipulaturiko eta bai neurturiko aldagaien definizio operazionaletatik edo indikatzailetatik abiatuz, goiko mailako konstruktuetara orokortzeko daukagun konfidantza maila bezala. Aurreko zenbait lanetan, bai Campbell (1969) eta baita Campbell eta Stanley (1966) ere, eragile eta eragindakoen konstruktua balidezia kanpo balideziaren barnean sartu zuten, aipatutako balidezia hori indikatzaile zehatzetatik berauek errepresentatzen dituzten konstruktua teorikoenganako orokortzeari baitagokio.

Azkenik, X konstruktua Y konstruktuen eragilea dela suposatuz, ikertzaileak erlazio horren hedadura azter dezake. Zentzu honetan, Cook eta Campbell *kanpo balidezia* hitza hurrengo hau adierazteko erabiltzen dute: aldagaien arteko ustezko kausa erlazioa beste pertsona, testuinguru eta une desberdinetara orokortzeko daukagun konfidantza maila, hain zuzen ere.

Cook eta Campbellek (1979) lau balidezia mota bereizten dituen sailkapena planteatzeko arrazoiak funtsean praktikoak direla baieztatzen dute. Zehazki, sailkapen hau ikertzaile orok bere buruari galdetu beharko lizkiokeen lau galdera nagusiei dagokie. Ikus dezagun:

- 1.- Bi edo aldagai gehiagoren artean erlaziorik ba ote dago?
- 2.- Erlazioa baldin badago, kausa izaerakoa da edo ustezko eragilea edo tratamendua ez dagoenean ere ager daiteke?
- 3.- Erlazioa kausazkoa bada, zeintzuk dira bertan parte hartzen duten eragile eta eragindako konstruktua?
- 4.- X (eragile) konstruktua eta Y (eragindako) konstruktua artean kausa erlazioa egonez gero, zein neurritan orokor daiteke beste subjektu, testuinguru eta une batzutar?

Galdera hauetako bakoitza inpliziturik zegoen Campbell eta Stanleyk (1963, 1966) aldeztatik proposatutako balideziaren sailkapenean, zeinean ondorio estatistikozko balidezia eta barne balidezia bera egileek barne balidezia deitutakoaren barnean sartuko lirakekeen eta, modu berean, eragile eta eragindako konstruktua balidezia eta kanpo balidezia Campbell eta Stanleyren ereduaren kanpo balideziaren barne egongo lirakekeen.

2.2.3.- Kruglanski eta Kroyren eredu

Balidezia mota desberdinen sailkapena eskaini ordez, Kruglanski eta Kroyk (1976) saiakuntza eremuaren balideziari buruzko zentzu bakarreko adiera planteatzen dute. Honela, esperimentuak balidezia badu eduki baldin eta egoera esperimentaleko elementuek hipotesi esperimentalean zehaztutako kontzeptuak eta erlazioak egokiro adierazten badituzte. Saiakuntza egoeraren alderdiren batek egiaztatu nahi den proposamen teorikoan edo hipotesian dagokion elementua adierazten ez duenean, alderdi hori "garbiarazi" egiten da, hots, aldagai nahastatzaile posibleak kontrol esperimentaleko prozeduren bitartez iraitzen dira. Ondorioz, balideziaren funtsezko arazoa edozein mailatan *errepresentazio*arena da eta balidezia lortzeko modua, kontrol esperimentaleko tekniken bidez *garbiarazpena* egitean datza.

Egileek kasu honetan proposatzen duten sailkapena egokiro errepresentatu beharreko elementuei dagokie, hau da, saiakuntza hipotesiaren egiturazko osagaiei edo *balidezia erreferenteak* deiturikoei. Honela, "edozein X-entzat, G-ren berdina bada, orduan X M-ren berdina izango da" adierazburuan Kruglanski eta Kroyk planteatzen dituzten lau osagaiak honako hauek dira:

- 1) Aplikazio eremua,
- 2) Saiakuntza hipotesiaren aitzindaria edo aldagai askea (G),
- 3) Saiakuntza hipotesiaren ondorioa edo menpeko aldagaia (M), eta
- 4) "bada ... orduan" juntagailua edo G eta M-ren arteko kausa erlazioaren adierazpena, hots, efektu experimentalak.

Esperimentu batek *unibertso balidezia* dauka bertako subjektu edo objektuek saiakuntza hipotesian zehaztutako populazioa egokiro adierazten dutenean.

Aldagai askearen balidezia bi mailatan planteatu daiteke. Batetik, operazional mailan aldagai askeak balidezia dauka baldin eta tratamendu experimentalak ikertzaileak manipulatu nahi duen baldintza edo tratamendua benetan adierazten badu. Bestetik, bigarren mailan, edo maila teorikoan, aldagai askeak balidezia dauka ikerketa zehatz batean aplikatutako tratamendu experimentalak goi mailako konstruktua teorikoa egokiro errepresentatzen duen neurrian.

Menpeko aldagaiaren balideziari dagokionez, planteamendua aurrekariaren balideziarenaren antzekoa da. Honela, menpeko aldagaiaren neurketa indikatzaileak neurtu nahi den portaera edo objektua modu egokian adierazten badu, menpeko aldagaiak balidezia dauka operazional mailan. Bestalde, erabilitako indikatzaileak goi mailako edo elaborapen kontzeptual gehiagoko konstruktua egokiro adierazten duen neurrian, menpeko aldagaiak balidezia dauka maila teorikoan.

Azkenik, *efektu experimentalaren balideziaz* ondorengoa ulertzen da: efektu experimentalak zein punturaino adierazten duen aldagai askearen eta menpeko aldagaiaren arteko benetako erlazioa, hots, ausazkoa ez den erlazio sistematikoa.

2.2.4.- Cronbachen eredua

Kruglanski eta Kroyren (1976) antzeko ikuspegitik abiatuz Cronbachek (1982) ere, ikertzaile orok edozein inferentzia egiterakoan kontuan hartu beharreko elementuak oinarri gisa hartuz definitzen du balidezia. Zehazki, autoreak lau elementu mota bereizten ditu: unitateak ("units") (e.b. subjektuak), tratamenduak ("treatments"), eragiketa behagarriak ("observing operations") (e.b. menpeko aldagaiak) eta esparrua ("setting") (hau da, ikerketa zein unetan eta zein baldintza kulturaletan burutzen den). Lau elementu hauetan oinarrituz Cronbachek UTOS laburdura proposatzen du.

UTOS laburdura, larritan idatzia, ikerketa burutzen den interes guneari dagokio, hots, ikertzaileak T (tratamendua) eta O (eragiketa behagarriak) elementuak definitzen dituen modua eta U-ren (unitateren, normalki pertsonen) populazioaren zehaztapena barnean ematen ditu. Bestalde, utoS laburdura, txikitan idatzia, ikerketa zehatz bateko adibide bereziei dagokie, hau da, ikerketa horretan aztergai diren manipulazioari (t), neurriari (o) eta subjektuari (u). Haatik, Cronbachek ez du s letra txikitan erabiltzen fisheriar zentzuan finkotzat hartzen duelako, bere ustez, ez baitu adierazten ikertzaileak mugatu nahi duen esparru zehatza. Azkenik, ikerketa burutzeko erabili denetik desberdina den aplikazio eremua adierazteko, Cronbachek asteriskoa gehitzen dio UTOS laburdurari (*UTOS).

Cronbach bi inferentzia mota interesatzen zaizkio. Lehena, *barne inferentzia* deitua, utoS-etik abiatzen da UTOS-era iristeko (hots, ikerketaren eragiketa zehatzetatik berauek adierazten duten ikerketa eremuraino doa). Bigarrena, *kanpo inferentzia* delakoa, utoS-etik *UTOS-era igarotzen duena da (ikerketaren eragiketa zehatzetatik ikerketa burutzeko erabili denetik desberdina den aplikapen esparrura, alegia). Cronbachek *barne balidezia* eta *kanpo balidezia* terminoak erabiltzen ditu barne eta kanpo inferentzien egitasun maila hurrenez hurren adierazteko. Beraz, Cronbachek barne balidezia terminoari egozten dion esanahia ez dator bat Campbell eta Stanleyk edo Cook eta Campbellek egozten diotenarekin, zeren eragiketa zehatzetatik (uto) maila kontzeptualetara (UTO) orokortzea suposatzen baitu. Azken batean, barne balideziak Cronbachen ereduan, Cook eta Campbellek kanpo balideziaz (u-tik U-ra) eta konstruktua balideziaz (t-tik T-ra eta o-tik O-ra) ulertzen dutena barne ematen ditu.

Lehenago aipatu den bezala, **klasikotzat hartzen diren autoreak ikerketa eremuaren barnean balideziari buruz planteatu dituzten sailkapen desberdinek desadostasunak agertzen dituzte** zenbait alderditan. Desakordio nabarmenetariko bat barne balideziaren inguruan ematen da. Honela, Campbell eta Stanleyren (1963, 1966) eta Cook eta Campbellen (1979) ereduetan, barne balidezia tratamendu esperimental zehatz batean eta pertsona konkretuetan eta testuinguru berezian behatutako portaera baten arteko kausa erlazioa ahalik eta zehazkien ezartzeari

dagokio. Hau da, barne balidezia orokorpenik gabeko inferentzia gisa ulertzen da. Aitzitik, bai Kruglanski eta Kroyren (1976) eta baita Cronbachen (1982) aburuz ere, barne balidezian inpliziturik dagoen kausalitate kontzeptua ezin daiteke espazio-denborazko muga zorrotzen barnean sartu eta, gainera, tratamendu esperimentalak eratzten duten gertaera zehatzek kontzeptuzko kategoria orokorragoren partaide diren neurrian soilik hartzen dute zentzua.

Bestalde, bai Campbell eta Stanleyk (1963, 1966) eta baita Cook eta Campbellek (1979) ere ikerketa prozesu berari izugarrizko garrantzia ematen dioten bitartean, Cronbachi (1982) gai aplikatuak interesatzen zaizkio batik bat. Ondorioz, lehenengoei barne balidezia ikerketaren *sine qua non*-a bezala hartzen dute eta bigarrenak, berriz, *kanpo balideziaren* funtsezko ekarpena aldarrikatzen du. Hori horrela izanik, Campbell eta bere laguntzaileen eredu oso egokia suerta daiteke tratamendu esperimentalaren eraginak populazio nahiz testuinguru desberdinetan bere horretan dirauen bitartean, baina inferentzia kausala aldagaien arteko erlazio faltsuen arriskupean dagoenean. Cronbachen eredu, bestalde, erabilgarriagoa gertatzen da erlazio faltsu horiek gertagaitzak direnean, baina tratamenduaren eragina ikerketaren baldintzak aldatu heinean erraz desagertu edo alda daitekeenean.

2.2.5.- Sailkapen klasikoek bestelako bi ikuspegi: Marken eredu eta Chen eta Rossiren eredu.

Arestian aipatu autore klasikoek proposatutako eredu nagusien aurrean, beste zenbait planteamendu agertu da. Eredu klasikoetan ez zaie balidezia mota desberdinei garrantzi bera egozten, baina azken hauek berriz, ikuspegi eklektikoagoak dira. Aurreragoko atal batean (5. puntua) planteamendu horien azalpenari lotzen gatzazkiolarik, orain aurreko ereduaren esplizituki aztertu ez diren zenbait arazo interesgarritaz jabetzeko aukerari helduko diogu. Horrela, proposamen berrien artetik, bi aipamen labur dakarzkizuegu.

Marken eredu

Ikerketa diseinuaren testuinguruan aurreko sailkapenen ekarpenak zalantzan jarri gabe, Markek (1986) orain arte behar bainateko garrantzia jaso ez duten zenbait alderditara jotzeko beharra azpimarratzen du. Alderdi horiek diseinuaren balidezia areagotzeko aukera ematen dute eta bost puntutan labur daitezke:

- 1.- *Goi mailako orokorpenak* egiten saiatzea. Marken ustez, barne balidezia indartzeaz gainera, ikertzaileak behe mailako definizio operazionalatik edo

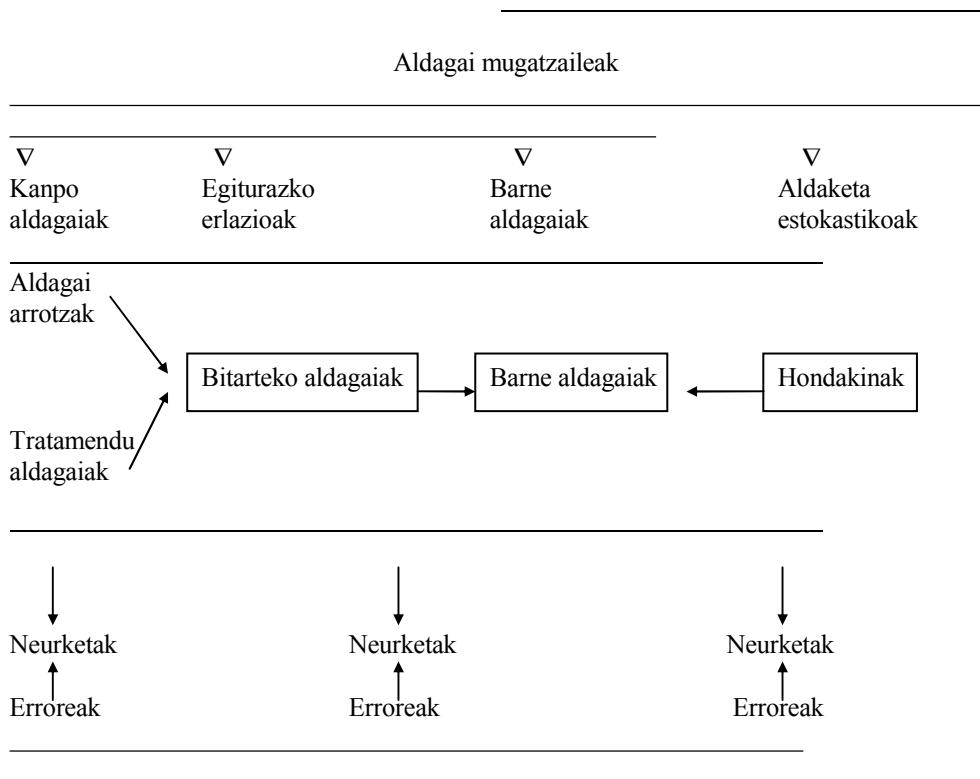
indikatzailleetatik abiatuz goi mailako orokorpenak modu zehatz eta fidagarrian egiteko beharrezko prozedurak erabili behar ditu.

- 2.- *Kausazko prozesuak aztertzea*. Egilearen aburuz, diseinuaren balidezia neurri handian areagotu daiteke. Horretarako, eragin zehatz baten azpiko mekanismoa edo tratamendu baten eragina azal dezakeen sekuentzia kausala aztertu beharko litzateke.
- 3.- *Balideziaren sailkapenak dogmatikoki ez onartzea*. Eredu honen arabera, balideziaren sailkapenak aldagai arrotzak edo ikerketa diseinuaren aurkako mehatxuak sumatzeko baliagarri gerta daitezke baina, era berean, kaltegarriak izan daitezke ikertzaile beraren analisi kritikoa eta logika ordezkatzan badute. Kasu honetan, ikertzailea erabaki desegokiak hartzera edo premiagabeko kontrol teknikak erabiltzera eramán dezakete, diseinuaren balidezia narriatuz.
- 4.- *Balidezia eta garrantzia ez nahastea*. Proposamen honetan Markek hurrengoa azpimarratu nahi du: diseinuaren balidezia egokia izateak ez dakar, besterik ezean, ikertzaileak burututako inferentziak edo lortutako ondorioak aplikazio eremu zehatz baten barnean garrantzitsuak izaterik.
- 5.- *Ez ulertzea balidezia "guztia/deus ez" arazoa izango bailitzan*. Planteamendu klasikoetatik aldenduz, non inferentziak kausazko erlazioei buruzko erabaki dikotomikoren (bai ala ez) baliokideak diren, Marken ereduan balidezia Reichardtén (1986) ikusmoldearen arabera ulertzen da, hots, efektuaren tamaina mugatua edo segurtasun maila mugatua daukaten inferentzia multzo gisa.

Chen eta Rossiren eredua

Chen eta Rossiren (1987) planteamendua psikologian burutzen diren interbentzio programen ebaluazio eremuan sortzen da, eta Campbell eta bere laguntzaileen hurbilpenak bezala barne balideziari garrantzia emateaz gain, Cronbachen proposamena, kanpo balidezia handitzera zuzendua, kontuan hartzen du ere, elkarren arteko sintesia eratu nahiean. Edozein ikerketan balidezia mota guztiak garrantzitsuak izan daitezkeela abiapuntutzat hartuz, autoreek teoriarin oinarritutako eredua ("theory-driven approach") proposatzen dute diseinuaren balidezia hobetzeko. Izan ere, edozein interbentzio programaren ebaluazioa eredu teoriko batean oinarritu behar dela argudiatzen dute. Azken buru, programen ebaluazioa ikerketa diseinuaren araberrakoa izan ordez, ikertzailearen ezagutza teorikoaren menpe egon beharko lukeela uste dute.

Idea horretatik abiatuz, Chen eta Rossik balideziari dagokion zenbait arazo ebaluatzeko eta Cook eta Campbellrek definitutako lau balidezien aurkako mehatxuak atzemateko esparru kontzeptuala definitzen dute (ikus 2.1. irudia).



2.1. Irudia.- Balidezia ebaluatzeko esparru kontzeptuala

Aurreko koadroaren ezker aldean bi *kanpo aldagai* mota beha daitezke: (1) *Tratamendu aldagaiak*, hots, esperimendatzaileak manipulaturako edo banaturako aldagaiak, tratamenduak alegia; eta (2) *aldagai arrotzak*, hau da, *barne aldagai* direlakoetan eragiten dutenak eta tratamendu aldagaiekin koerlazioa edo koerlazio- eza izan dezaketenak.

Bi *kanpo aldagai* mota hauek *bitarteko aldagai* desberdinen bidez *barne aldagaiei* eragiten diete. *Kanpo aldagai*, *bitarteko aldagai* eta *barne aldagaien* arteko *egiturazko erlazioak* *eredu teoriko* baten barnean finkatzeak aukera dakargu *kanpo aldagaien* eta *barne aldagaien* arteko erlazioen itzalean dauden prozesu kausalak antzemateko. Halaber, *barne aldagaiak* *hondakinak* deritzoten ausazko aldaketen eraginpean aurkitzen dira.

Aurreko irudiaren (2.1. irudia) goikaldean *prozesu kausalean parte hartzen duten elementuak aldagai mugatzaile* batzuegandik baldintzatuak ikus daitezke. Behekaldean, sisteman zehaztutako aldagai teorikoen *neurketak* edo *definizio operazionalak* agertzen zaizkigu, zeintzuk *errore* pean dauden. Erdialdean tratamendu aldagaien eta barne aldagaien arteko erlazioa azal dezaketen *prozesu kausalak* eta baita erlazio horri eragin diezaioketen beste zenbait elementu ere adierazten dira.

Eredu honek barne balidezia, kanpo balidezia, konstruktua balidezia eta ondorio estatistikozko balidezia nola areagotzen dituen egoki ulertu ahal izateko oraindik aztertu ez diren zenbait kontzeptu ezagutu behar dugu. Dena den, hori guztia ondoko atal baterako utziko dugu. Oraingoz, aski zaigu jakitea Chen eta Rossiren planteamenduak aldagai aipagarrien eta bitarteko prozesuen arteko erlazioak zehazki eredutzeko aukera ematen duela. Hau da, tratamendu esperimentalen eta hauek eragindako ondorioen arteko erlazioaren itzalean dauden mekanismo kausalen ulertzeari ekiten diola.

Laburbiltzeko, 2.2. atalean zehar aztertutako sailkapen desberdinen alderdi nagusiak 2.1. Taulan hurrengo orrialdean aurkezten dira.

2.1. Taula.- Balideziaren sailkapena ikuspegi desberdinen arabera

AUTOREAK	SAILKAPENA
1.- CAMPBELL ETA STANLEY	Balidezia motak: - Barne balidezia - Kanpo balidezia
2.- COOK ETA CAMPBELL	Balidezia motak: - Ondorio estatistikozko balidezia - Barne balidezia - Eragile eta eragindako konstruktua balidezia - Kanpo balidezia
3.- KRUGLANSKI ETA KROY	Saiakuntza hipotesiaren egiturazko osagaiak edo balidezia erreferenteak: - Unibertso balidezia - Aldagai askearen balidezia - Menpeko aldagaiaren balidezia - Efektu esperimentalaren balidezia
4.- CRONBACH	Inferentziak burutzerakoan kontuan hartu beharreko elementuak eta inferentzia horiei dagozkien balidezia motak: - Barne inferentzia (utoS→UTOS): Barne balidezia - Kanpo inferentzia (utoS→*UTOS): Kanpo balidezia
5.- MARK	Ikerketa diseinuaren balidezia areagotzeko aukera ematen duten prozedurak: - Goi mailako orokorpenak egin - Kausazko prozesuak aztertu - Balideziaren sailkapenak dogmatikoki ez onartu - Inferentzia baliagarriak eta garrantzitsuak burutu - Ez ulertu balidezia "guztia-deus ez" arazoa izango bailitzan
6.- CHEN ETA ROSSI	Teorian oinarritutako eredua ("theory-driven approach"): - Eredua teoriko egoki bat edozein balidezia mota areagotzeko oinarri bezala har daiteke

2.3.- Saiakuntza eta sasisaiakuntza diseinuaren eremuan sistematizapen hedatuena: Cook eta Campbellen eredia

¡Error! Marcador no definido.

Bigarren atal honetan zehar ikusi dugun bezala, balideziaren arazoa ikuspegi desberdinetatik aztertua izan da. Dena den, ikerketa diseinuaren barnean balidezia landu duten autore gehienek Cook eta Campbellen (1976, 1979, 1986) planteamenduari jarraitu diote edo, gutxienez, berau hartu dute abiapuntu bezala. Ikertzaile batzuk (e.b. Judd eta Kenny, 1981; Krathwohl, 1985) eredu hori sakonki aztertu dute eta Cook eta Campbelllek proposatutako balidezia motei buruzko eztabaida bikainak eskaini dizkigute. Beste batzuk (e.b. Pedhazur eta Pedhazur Schmelkin, 1991) lau balidezia mota hauek ikerketa diseinuaren barnean kokatu ordez, konstruktuko balidezia neurketa psikologikoaren eremuan eta ondorio estatistikozko balidezia hipotesi testaren eta laginketaren testuinguruan aztertu dituzte. Dena den, zenbait aldaketa edo ñabardura aurki daitezkeela onartuz ere, eztabaidea da Cook eta Campbellen eredia psikologiaren eremu metodologiko gehienetan zabalki erabilia izan dela eta gaur egun ere zabal erabiltzen dela. Hori dela eta, hurrengo lau ataletan, **Cook eta Campbelllek balidezia aztertzeko proposatutako ereduari** jarraituko diogu.

Azalpenaren hasieran baieztatu den bezala, ikerketa ororen helburu nagusia ondorio zehatz, fidagarri eta, batez ere, baliagarriak lortzea da. Cook eta Campbellen iritziz, ikerketaren balidezia ezinhobeia izan beharko litzateke bai ondorio estatistikozko balideziari (analisi mailan), bai barne balideziari (ezarrera mailan), bai konstruktuko balideziari (maila teorikoan) eta baita kanpo balideziari (emaitzen orokorpen mailan) dagokionez ere. Lau balidezia mota hauek ikertzaileak edozein ikerketa planifikatu aurretik kontuan hartu behar dituen alderdi anitzak barnean ematen dituzte. Arestian esan bezala, balidezia mota bakoitzak galdera nagusi bati erantzuten dio ikerketa diseinuaren barnean (Kazdin, 1992):

- 1.- *Ondorio estatistikozko balidezia:* Zein da efektuaren tamaina, eta zein zorroztasun maila dauka ikerketak bi edo aldagai gehiagoren arteko kobariazioa antzemateko kobariazio hori ematen bada?
- 2.- *Barne balidezia:* Emaitzak, aldaketak edo taldeen arteko desberdintasunak tratamenduaren ondorio bezala soilik azal daitezke? Hau da, bestelako azalpen edo hipotesi lehiakide posible guztiak bazter daitezke?
- 3.- *Konstruktuko balidezia:* Tratamendua aldaketaren eragilea dela suposatuz, tratamenduaren zein alderdi zehatzek sortzen du eragina? Beste era batera, zein da eraginaren itzalean dagoen erreferente kontzeptuala (konstruktuko teorikoa)?

- 4.- *Kanpo balidezia*: Emaizak beste subjektu, testuinguru, une eta jatorrizko ikerketatik desberdinak diren beste ezaugarri batzutara orokor daitezke?

Galdera hauek erreferentzia puntu direla, balidezia mota bakoitza aztertuko dugu: ikerketaren zein alderdiri dagokio, berorren aurkako mehatxuak eta ikerketan aldagai aske bezala ez jokatzeko arren lortutako emaitzak azal ditzaketen faktore edo elementu hauek kontrolatu ahal izateko prozedurak ere.

3.- Ondorio estatistikozko balidezia

3.1.- Ondorio estatistikozko balideziaren definizioa eta ezaugarri nagusiak

3.2.- Ondorio estatistikozko balideziaren aurkako mehatxuak

3.2.1.- Eredu estatistikoaren baldintzen betekizun eza

3.2.1.1.- Normaltasun aurretikoa

3.2.1.2.- Bariantzen homogeneotasun aurretikoa

3.2.1.3.- Behaketen arteko askatasun aurretikoa

3.2.2.- Test estatistikoaren indar eskasa

3.2.3.- Efektuaren tamaina

3.2.4.- Laginaren tamaina

3.2.5.- I. motako errore maila

3.2.6.- Menpeko aldagaiaren neurketan fidagarritasun eta balidezia eskasa

3.2.7.- Tratamenduak ezartzerakoan fidagarritasun eskasa eta egiaztapen indar gutxiko tratamenduen aplikapena

3.2.8.- Saiakuntza unitateen arteko ausazko heterogeneotasuna

3.2.9.- Ausazko aldakortasunak saiakuntza egoeran

3.2.10.- "Outliers" edo ibiltartez kanpoko balioak

3.2.11.- "Missing" edo balio ez-agertuak

3.- ONDORIO ESTATISTIKOZKO BALIDEZIA

3.1.- Ondorio estatistikozko balideziaren definizioa eta ezaugarri nagusiak

Aurreko 2.2.2. atalean aipatu den bezala, *ondorio estatistikozko balidezia* esanguratsutasun maila zehatza kontuan hartuz eta lortutako bariantzen arabera, bi edo aldagai gehiagoren arteko ustezko kobariazioari buruzko inferentziak egiteari dagokio; maila estatistikoan hipotesiaren inferentzia egokiro burutzeari, alegia (Vallejo, 1991a). Hipotesi estatistikoari buruzko eta, azken batean, tratamenduaren eraginkortasunari buruzko inferentzia egiterakoan ikertzailea bi errore motatan eror daiteke. *I. motako errorea (positibo faltsuak)* aldagaien artean benetako erlaziorik ez dagoenean, erlazioa badagoela erdiestean datza. *II. motako errorea (negatibo faltsuak)*, berriz, populazioan benetan dagoen erlazioa aintzat ez hartzeari dagokio. Argi geratu behar da, barne balidezia zalantzan jartzen denean tratamenduaren eragina desitxuratzen den bitartean, ondorio estatistikozko balideziaren aurkako faktoreen kontrol ezatik eratorritako alborapenak ez duela tratamenduaren efektuen gainbalorapena edo infrabalarapena eragiten, efektu horien esanguratsutasun estatistikoarena baizik.

Ondorio estatistikozko balideziak edo bi edo aldagai gehiagoren arteko kobariazioa antzemateak hiru funtsezko galderai erantzutea eskatzen du, hots:

- 1.- Ikerketa, aldagaien arteko kobariazioari buruzko ondorio adierazgarriak atera ahal izateko adina zorrotza al da?
 - 2.- Behar adina zorrotza izanez gero, ustezko eragile-eragindako kobariazioa inferitzeko beharrezko datu enpirikoak ba al daude?
 - 3.- Datu enpiriko hauek egonez gero, zein da X eta Y-ren arteko kobariazio maila?
-
- 1.- Lehenengo galdera *ikerketa diseinuaren potentzia edo indar estatistikoari* dagokio, hau da, eragina dagoela suposatuz, eragin hori sumatzeko diseinuak duen probabilitateari. Autore gehienek potentzia kontzeptua diseinuaren zorrotzasun kontzeptuarekin parekatzen duten arren, badira bi kontzeptu hauek bereizten dituzten zenbait ikertzaile (e.b. Lipsey, 1990). Lipseyk honela definitzen du *potentzia estatistikoa*: tratamendua benetan eraginkorra denean esanguratsutasun estatistikoa lortzeko probabilitatea da. *Zorroztasuna*, berriz, tamaina esanguratsuko efektua sumatu ahal izateko gaitasunarekin parekatzen du. Tratamenduaren eragina antzemateak kontrol taldearen eta talde esperimentalaren puntuazioen artean estatistikoki esanguratsuak diren diferentziak aurkitzea eskatzen duenez, Lipseyk dio diseinuaren zorrotzasuna azaltzerakoan diseinu horrek potentzia estatistikoan duen eragina aipatu behar dela funtsean. Honela, egilearen aburuz, diseinua zorrotza izango da berorren baldintzak test estatistikoaren potentzia

areagotzeko egokiak direnean. Lipseyren antzeko planteamendutik abiatuz, Cohenek (1988) dio zorrotzasun kontzeptua tratamenduen eragina sumatzeko posibilitatea mugatzen duten parametro anitzei lotzen zaien bitartean, potentzia kontzeptuak askoz parametro gutxiago ukitzen dituela.

Edozein ikerketa planifikatzen denean, *potentzia analisia* bariantza zehatzetatik abiatuz, nahi dugun tamainako efektua sumatu ahal izateko beharrezko subjektu kopurua mugatu egiten da. Aitzitik, ikerketa burutu denean, potentzia analisiak funtzio desberdina betetzen du. Kasu honetan, laginaren bariantza eta lagina osatzen duten subjektu kopurua populazioan egon litekeen efektuaren tamaina finkatzeko erabiltzen dira. Honela, efektuaren tamainaren estimazioa balio txikikoa bada eta datu estatistiko anitzez adierazita badago, potentzia handia dela pentsa daiteke eta, ondorioz, hipotesi nulua behin-behinean onar daiteke. Datu enpiriko adina ez badago, hipotesi nulua onartzea arriskutsua da, emaitza hori *lehen mailako bariantza sistematikoa* behar bezala gehitu ez delako ager baitaiteke. Bestalde, efektuaren tamainaren estimazioa handia bada eta esanguratsutasun estatistiko handia badaukagu, ikerketa kobariazioa sumatzeko adina zorrotza dela behin behinean ondoriozta dezakegu. Esanguratsutasun enpirikorik ez badago, ezin genezake jakin kobariazio ezak populazioan aldagaiak elkarren artean askeak direla adierazten ote duen edo erkaketa irizpide gisa hartzen den zarata hondoan benetako efektua sumatzeko ikerketaren zorrotzasun eskasa isladatzen ote duen. Azken kasu honetan, *bariantza errorea* behar beste ez dela txikiagotu pentsa genezake.

- 2.- Bigarren galdera, ustezko eragile baten eta ustezko eragindako baten artean kobariaziorik ote dagoen erabakitzeak aukeratu behar diren test estatistikoei dagokie batik bat. Hautapen hau zenbait irizpide objektiboren menpe dago (Sheskin, 1984; Siegel eta Castellán, 1988), hots: (a) *Testaren potentzia estatistikoa*. Irizpide honen arabera, gainerako faktoreak egonkor daudelakoan, I. motako errorea txikiagotzen duen eta faltsua denean hipotesi nulua baztertzeko probabilitatea handitzen duen testa aukeratu behar dugu; eta (b) *Proposatutako eredu estatistikoa*. Ereduak finkatutako aurretikoak edo baldintzak betetzen badira, eredu matematiko murrizgarrietan oinarritzen diren test estatistikoak, test parametrikokoak alegia, test ez-parametrikokoak baino egokiagoak dira hipotesi nulua baztertu behar denean. Test parametrikokoak ausartagoak, *indartsuagoak* direla esan genezake.
- 3.- Hirugarren galdera aurkitutako kobariazio mailari edo *efektuaren tamaina* deritzonari dagokio. Efektuaren tamainaren estimazioa ez dago, esanguratsutasun test klasikoak bezala, laginaren tamainaren menpe, horregatik kobariazioa egokiro ondoriozta ote daitekeen estimatzeko abantaila asko ditu.

Atal honetan zehar azalduak izan diren arren, **ondorio estatistikozko balideziarekin zerikusia duten zenbait kontzeptu garrantzitsu** esplizituki definitzea, balidezia mota honek ikerketaren zein alderdi ukitzen dituen argiago ulertzeko lagungarri gerta daitekeela uste dugu. Hori dela eta, kontzeptu horiek hurrengo taulan laburtzen ditugu (ikus 3.1. Taula).

Oro har, bi edo aldagai gehiagoren arteko kobariazioa sumatzeko edo tratamenduaren eragina inferitzeko posibilitatea *aldakortasun erlatiboaren* menpe dago batik bat, hau da, baldintza esperimentalen barnean menpeko aldagaiaren puntuazioei dagokien *errore bariantzaren tamainaren* eta baldintza esperimental horien arteko diferentziaren tamainaren edo *efektuaren tamaina delakoaren* artean dagoen desberdintasunaren menpe.

3.1. Taula: Ondorio estatistikozko balideziari lotutako zenbait kontzeptu garrantzitsu.

KONTZEPTUA	DEFINIZIOA
I. motako errorea Alpha (α)	Hipotesi nulua benetakoa denean berau ukatzea. I. motako erroreari dagokion probabilitate maila.
II. motako errorea Beta (β)	Hipotesi nulua faltsua denean berau onartzea. II. motako erroreari dagokion probabilitate maila.
Indarra ($1-\beta$)	Hipotesi nulua faltsua denean berau baztertzeari dagokion probabilitatea edo ordezkio hipotesia onartzeko daukagun probabilitatea berau egiazkoa denean.
Efektuaren tamaina	Bi edo baldintza esperimental gehiagoren arteko (edo baldintza esperimental eta kontrol baldintzen arteko) diferentziaren tamaina neurketa eta ikerketa desberdinentzat neurbide edo metrika berberaz adierazia.
Desbidazio tipikoa	Puntuazioek batazbestekoarekiko agertzen duten aldakortasun neurria.

(Kazdinen lanetik moldatua, 1992)

Aldakortasun kontzeptua ondorio estatistikozko balideziaren oinarri gisa har daiteke, efektu esperimentala haitemateko funtsezko zeregina betetzen baitu. Demagun, esate baterako, distantzia luze batetik Santiagoko bidea burutzen ari diren bi erromes talde Santo Domingo de la Calzadako ostatura hurbiltzen ikusten ditugula eta bien artean, A taldearen eta B taldearen artean alegia, lehena zein doan jakin nahi

dugula. Talde bakoitzaren barnean aldakortasunik ez balego, 3.1. Irudiaren goialdean errepresentatu denaren antzeko irudia ikusiko genuke. Kasu honetan, taldeen arteko gonbaraketa erraza gertatzen da eta inolako arazorik gabe ondoriozta daiteke B taldea Aren aurretik doala. Aitzitik, 3.1. Irudiaren erdialdean beha dezakegun bezala, taldebarneko aldakortasuna handia bada, ostatutik gertuago dagoen taldea zein den erabakitzea oso zaila gertatzen da, nahiz eta bi taldeen distantzien arteko batzbesteko diferentzia goialdekoaren berdina izan. Alabaina, aldakortasunaren eragina erlatiboa da, antzeman nahi dugun efektuaren tamainaren menpe baitago. Gure adibidearekin jarraituz, erromesek 3.1. Irudiaren erdialdean agertzen duten adina aldakortasuna aurkeztuz gero, baina bi taldeen artean distantzia handiagoa egonez gero (3.1. Irudiaren behealdean), diferentzia erraz hauteman daiteke berriro ere.

			ERROMESAK	JOMUGA ETA BEHAKETA PUNTUA	
Goialdea			A B A B A B A B A B	Santo Domingo de la Calzadako ostatua	
Erdialdea	A B B A A B	B A B	Santo Domingo de la Calzadako ostatua		
Behealde	A A A A	B B B B	Santo Domingo de la Calzadako ostatua		

3.1. Irudia.- Aldakortasun erlatiboak efektu esperimentalaren hautematean betetzen duen funtzioa ikusi ahal izateko adibidea

Ikerketa diseinuaren eremuan laginketatik, prozedura esperimentaletik, neurketatik edo beste iturri askotatik erator daitezkeen eta ondorio estatistikozko balidezia zalantzan jar dezaketen aldakortasun eragile anitz daude. Jarraian, bi edo aldagai gehiagoren arteko kobariazioari buruzko inferentzia eragotz dezaketen faktoreak eta mehatxu hauen aurrean erabili beharreko kontrol teknikak edo prozedurak deskribatzen ditugu.

3.2.- Ondorio estatistikozko balideziaren aurkako mehatxuak

3.2.1.- Eredu estatistikoaren baldintzen betekizun eza

Test estatistikoek datuek aurretiko batzuk bete ditzaten eskatzen dute, horren ezean ikertzaileak inferentzia desegokiak egin baititzake. Baldintzen betekizun ezak bai I. motako erroreari eta bai II. motako erroreari eragin diezaiokie (Stevens, 1992).

Menpeko aldagaia gutxienez tarte neurketa eskalan emana izatearen aurretikoaz gain (*menpeko aldagaiaren neurketa baldintza*), bariantza analisiak oinarritzko beste hiru aurretiko eskatzen ditu: menpeko aldagaiaren puntuazioak edo behaketak normaltasunez banatzea (*normaltasun aurretikoa*), talde guztien bariantzen arteko homogeneotasuna edo homozedastizitatea (*Bariantzen homogeneotasun aurretikoa*) eta behaketen arteko independentzia.

Diseinuan aldagai aske bat baino gehiago izanez gero, aurreko baldintzez gainera, dagokion F testak beste bi betekizun egiztu behar ditu: populazioen batzbestekoak, normalak eta homogeneoak, diseinuko aldagai askeek eragindako efekturen arteko lerro konbinaketak izan behar dira (*efektuen gehikortasun aurretikoa*) eta ikertzaileak proposatutako eredu matematikoak planifikaturiko ikerketarentzat garrantzitsuak diren aldagai guztiak barne eman behar ditu (*ereduaren zehaztapen egokiaren aurretikoa*).

Kobariantza analisirako F testaren kasuan, aipatutako betekizunez gainera, ondoko aurretikoak bete behar dira: tratamenduek ez dute aldagai kobariantetan eraginik izan behar (*tratamendu-kobarianteen arteko askatasun aurretikoa*), erregresio-koefizienteak berdinak izan behar dira tratamendu bakoitzarentzat (*erregresio-koefizienteen arteko berdintasun aurretikoa*), menpeko aldagaiaren eta aldagai kobariantearen artean lerro ereduaren arabeko erlazioa egon behar da (*menpeko aldagaiaren eta aldagai kobariantearen arteko lerrotasun aurretikoa*) eta aldagai kobariantea errorerik gabe neurtu behar da (*aldagai kobariantearen neurketaren fidagarritasun aurretikoa*).

Diseinu esperimentalaren eremuan gehien erabiltzen den testak, bariantza analisiak alegia, eskatzen dituen funtsezko hiru betekizunak zehazkiago aztertu aurretik, aurretiko horiek ez betetzeak F, Welchen W eta Kruskal-Wallisen KW estatistikoetan izan ditzakeen ondorioak laburtzen dituen taula aurkezten dizuegu (ikus 3.2. Taula).

3.2. Taula - ANOVAk betebeharreko aurrekariak ez betetzeak F, Welchen W eta Kruskal-Wallisen KW estatistikoetan eragindakoa

Betekizun eza	Taldea orekatuak		Orekatu gabeko taldeak	
	Eragina α -n	Eragina indarrean	Eragina α -n	Eragina indarrean
Behaketen arteko independentzia eza	Behaketen arteko independentzia ezak bai F, W eta KW testen indarrari eragiten dio, taldeek subjektu kopuru berbera (orekatuak) edo desberdina (orekatu gabeak) izanik ere			
Normaltasun eza	Garrantzirik gabeko eragina F eta KW testetan. Eragina zertxobait handitzen da W testaren kasuan	Garrantzirik gabeko eragina F-n. Gutxitze xumea KW-n. Gehiegizkoa ez den gutxitzea W-n	Garrantzirik gabeko eragina KW-n. Handitze xumea F-n. Handitze garrantzitsua W-n	Garrantzirik gabeko eragina F-n. Gehiegizkoa ez den gutxitzea KW-n. Gutxitze garrantzitsua W-n
Bariantzen heterogeneotasuna	Gehiegizkoa ez den handitzea F-n, garrantzi gehiagohartzen duena bariantzak handitzen diren neurrian. Eragin aldakorra KW-n. Gehiegizkoa ez den eragina W-n 8:1 en gainetik dauden proportziotan	Garrantzirik gabeko eragina F eta W-n. Gehiegizkoa ez den gutxitzea KW-n	Eragin garrantzitsua F-n eta zertxobait txikiagoa KW-n. Handitze xumea W-n	Garrantzirik gabeko eragina F eta KW-n. Handitze xumea W-n

(Harwell, Rubinstein, Hayes eta Oldsetik moldatua, 1992)

3.2.1.1.- Normaltasun aurretikoa

Aurretiko honen arabera, menpeko aldagaian lortutako puntuazioak kurba normalari jarraituz banatu behar dira. *Jomugaren Teorema* delakotik abiatuz, hots, elkarren arteko behaketa askeen batura, hauen banaketa edozein izanik ere, lagingaren tamaina handitzen dugun neurrian banaketa normalari hurbiltzen zaiola dioen teoriatik abiatuz, laginketa unitate kopuru handia erabiltzen den guztietan normaltasun aurretikoa betetzen da. Gainera, datuen normaltasun ezak eragin txikia dauka F testean (e.b. Glass, Peckham eta Sanders, 1972).

Talde bakoitzean **menpeko aldagaiaren puntuazioen normaltasun aurretikoa betetzen ote den egiaztatzeko prozedura** anitz dago (Chambers, Cleveland, Kleiner eta Tukey, 1983; Stevens, 1992) eta, horrez gainera, ordenagailuetarako programa estatistiko arruntenek betekizun hau ebaluatzeko testak dakarzkigute. SASak eta SPSSak, esate baterako, *Shapiro-Wilken testa* aplikatzen dute lagingaren tamaina 50 subjektukoa edo txikiagoa denean eta *Kolmogorov-Smirnov* izeneko testa, berriz, 50 subjektu baino gehiago ditugunean. Azken test honek banaketa σ^2 bariantza eta μ batazbestekoa dituen kurba normalera egokitzen ote den aztertzen du. Populazioaren parametroak ezagutzen ez direnean, hau da, sarritan gertatzen zaigun bezala, lagina batazbestekoa eta bariantza ezezagunak dituen banaketa normalatik eratortzen denean, Kolmogorov-Smirnov testan *Lilliefors testa* (1967) deritzon zuzenketa aplikatu ohi da.

Faktore bakarreko taldearteko diseinurekin lan egiterakoan **normaltasun aurretikoa errotik hausten bada**, bariantza analisia ordezka dezakeen analisi teknika ez parametrikoki batera jo dezakegu, test ez-parametrikoei teknika parametrikoei baino aurretiko gutxiago eta orokorragoak ezartzen baitituzte, baina, era berean, ez dira azken hauek bezain indartsuak (Marascuilo eta McSweeney, 1977; Roscoe, 1975; Siegel, 1956). Gehien erabilitako prozeduren arteko bat *Kruskal-Wallis test* dugu, *mailaz mailako ANOVA metodoa* deitua ere. Ikertzaile batzuen ustez, Kruskal-Wallis enereduaren pean bariantzen homogeneotasun aurretikoa besterik ez da ezartzen (Maxwell eta Delaney, 1990). Beste batzuk diotenez, berriz, ez normaltasun aurretikoa eta ez bariantzen homogeneotasun aurretikoa betetzen ez den kasuetan ere aplika daiteke test hau (e.b. Roscoe, 1975). Pascual, García eta Friasek (1994) adierazten digutenez, normaltasun aurretikoa betetzen ez denean teknika parametrikoki eta ez parametrikoren arteko hautapenaren arazoa ebazteko *mailaz mailako eraldakuntza* (Conover eta Iman, 1981) izeneko prozedura proposatu da, bi metodo moten artean zubi gisa jokatu nahi duena, hain zuzen ere. Bestalde, normaltasunaren hausketa zorrotza jatorrizko puntuazioak banaketa normalera berregoki daitezen beharrezko eraldakuntzak eginez konpon dezakegu baita ere (Stevens, 1992).

3.2.1.2.- Bariantzen homogeneotasun aurretikoa

Aurretiko honen arabera, talde esperimental desberdinen bariantzek berdinak izan behar dute. Homozedastizitate ezak bai I. motako errorean erortzeko probabilitateari eta baita test estatistikoaren potentziari ere eragin diezaioke. Bariantza analisia sendoa da bariantzen homogeneotasuna gehiegi hausten ez den kasuetan (Winer, 1971). Gainera, zenbait autorek diotenez, baldintza edo tratamendu talde bakoitza subjektu kopuru berberaz osaturik dagoenean, F banaketa ez da serioski desitxurutzen bariantzen artean desberdintasun handiak egon arren. Alabaina, autore guztiek ez dute baieztapen honekin bat egiten. Honela, zenbaitzuren ustez, bariantzen heterogeneotasuna atzemanaz gero, F testaren sendotasuna zalantzan jarri behar da talde bakoitzean subjektu kopuru berbera dagoenean ere (Clinch eta Keselman, 1982; Wilcox, Charlin eta Thompson, 1986). Azken buru, ikertzaile gehienek onartzen dutena zera da: taldeak orekatuak ez direnean, gehiegizkoa ez den heterogeneotasun mailak ere, *I. motako errorea edo benetako α eta ikertzaileak berak aurrez finkatutako α nominala* elkarren artean oso desberdinak izatea eragin dezakeela (e.b. Maxwell eta Delaney, 1990; Riba, 1990).

Normaltasun ezak ez bezala, homozedastizitate ezak zenbait kasutan F testean eragin handia izan dezakeenez gero, aurreko baldintzaren kasuan baino sakonkiago aztertuko ditugu homogeneotasunaren aurretikoa betetzen ote den egiaztatzeko aukera eskaintzen diguten formulapenak eta baita beronen hausketa zuzentzeko erabil daitezkeen prozedurak ere.

Bariantzen arteko homogeneotasun aurretikoa betetzen ote den egiaztatzeko prozedurak

Berauek guztiak datuen banaketa normalaren hausketarekiko sendoak ez badira ere (Stevens, 1992), bariantzen homogeneotasuna egiaztatzeko gehien erabilitako testak *Hartley-ren gehienezko F-a*, *Cochranen C-a* eta *Bartletten testak* dira. Normaltasunaren betekizun ezaren aurrean testik sendoena *Leveneren testa* den arren, oso gutxi erabilia izan da ikertzaileen artean.

Hartleyren testaren oinarria, errore bariantza handienaren eta txikienaren arteko zatiketaren datza. Homogeneotasun hipotesia baztertzen da, gehienezko F-ren balio enpirikoak Hartleyren edo gehienezko F-ren taulatan agertutako balio kritikoa gainditzen duenean. Estatistiko hau erabili ahal izateko, tratamendu talde guztien subjektu kopurua berdina izan behar da (Arnau, 1981a; Roscoe, 1975).

$$\text{Gehienezko } F - a = \frac{\sigma_{\text{handiena}}^2}{\sigma_{\text{txikiena}}^2} \quad (3.1)$$

Izendatzaileari eta zenbakitzaileari dagozkion gehienezko F-ren banaketaren askatasun graduak a eta $n-1$ dira hurrenez hurren.

Cochranen testa, C estatistikoaren kalkuluan oinarritua, tratamenduetan behatutako bariantza handienaren eta bariantza guztien baturaren arteko zatidura da. Homogeneotasun hipotesia baztertu egiten da, C estatistikoaren balio enpirikoak Cochranen taulari dagokion balio kritikoa gainditzen duenean. Taldeen artean desberdintasun txikiak daudenean erabilgarria izan arren (Riba, 1990), test honen aplikazio egokiak taldeak orekatuak izan daitezen eskatzen du (Roscoe, 1975).

$$C = \frac{\sigma_{handiena}^2}{\sum \sigma_j^2} \quad (3.2)$$

Izendatzaileari eta zenbakitzaileari dagozkion C estatistikoaren laginketa banaketaren askatasun graduak a eta $n-1$ dira hurrenez hurren.

Bartletten testa aplikatzerakoan kalkulatzeko den estatistikoa χ^2 -a da. Kasu honetan, homogeneotasun hipotesia baztertzeko dugu χ^2 estatistikoaren balio enpirikoa χ^2 -ren taulatan agertutako balioa baino handiagoa denean. Test honen erabilpenerako ez da beharrezkoa taldeen tamaina berbera izatea.

$$\chi^2 = \frac{2.3026}{c} \left[gl_{errorea} \log MC_{errorea} - \left(\sum (n_j - 1) (\log \sigma_j^2) \right) \right] \quad (3.3)$$

non:

$$c = I + \frac{\left[\sum_{j=1}^a \frac{I}{(n_j - 1)} \right] - \left(\frac{I}{gl_{errorea}} \right)}{3(a - 1)} \quad (3.4)$$

Chi-karratuaren laginketa banaketaren libertate graduak $a-1$ dira.

Azkenik, **Leveneren testa**, normaltasun ezaren aurrean sendoena, ordeko hipotesitik abiatuz eratortzen diren estimazio erroreen balio absolutuak kalkulatzeko eta bariantza analisian puntuazio hauekin egiteko datza. Test hau nola aplikatzen den ikusi ahal izateko Pascual, García eta Fríasen (1994) testua kontsulta daiteke.

Datuak bariantzen homogeneotasun aurretikoa betetzen ez dutenean heterozedastizitatea zuzentzeko prozedura desberdinak aplikatu daitezke. Pascual, García eta Frías (1994) jarraituz, erabilienak aipatuko ditugu.

Bariantzen homozedastizitate eza sahiesteko prozedurak

Datuek bariantzen homogeneotasun aurretikoa betetzen ez dutenean erabil daitekeen teknika bat **O'Brien prozedura** deritzona da. Prozedura hau menpeko aldagaia eraldatzean eta eraldatutako puntuazioetan ANOVA aplikatzean datza. Honela, heterozedastizitatea zuzentzen da, I. motako eta II. motako erroreak kontrolatzen dira eta testaren indarra areagotzen da. Gainera, O'Brien prozedura normaltasuna betetzen ez den kasuetan ere sendoa da. Maxwell eta Delaneyk (1990) zuzenketa hau menpeko aldagai bakarreko diseinuetan egiterakoan jarraitu beharreko urratsak adierazten dizkigute:

- 1.- Talde bakoitzaren kasuan batazbestekoa ($\bar{Y}_j$) eta alborapenik gabeko bariantza (σ_j^2) kalkulatu:

$$\sigma_j^2 = \frac{\sum_i (Y_{ij} - \bar{Y}_j)^2}{n_j - 1}$$

- 2.- Behaketa bakoitzari dagokion puntuazio eraldatua (r_{ij}) lortu:

$$r_{ij} = \frac{(n_j - 1,5) \cdot n_j \cdot (Y_{ij} - \bar{Y}_j)^2 - 0,5 \cdot \sigma_j^2 \cdot (n_j - 1)}{(n_j - 1)(n_j - 2)}$$

- 3.- Talde bakoitzaren kasuan r-ren batazbestekoa eta Y-ren jatorrizko puntuazioen bariantza berdinak direla, hots, $\bar{r}_{ij} = \sigma_j^2$ berdintasuna egiaztatu.
- 4.- r puntuazioak menpeko aldagai bezala erabiliz bariantza analisisa aplikatu.

Bariantzen heterozedastizitatea agertzen denean sarri erabiltzen den beste prozedura bat **F kontserbatzailea** deritzon testa da. Test honen funtsa, esangura estatistikoa egokitu ahal izateko askatasun graduak zuzentzean datza. Test honek I. motako errorea kontrolatzen du, F estatistikoa esanguratsutat hartua izan dadin berorrek gainditu beharreko balio kritikoa handitzen baitu. Taldeek subjektu kopuru berbera dutenean ondorengo zuzenketa burutzen da (Riba, 1990):

$$\varepsilon = \frac{1}{a - 1} \quad ; \quad \varepsilon' = \frac{1}{a} \quad (3.5)$$

F estatistikoaren laginketa banaketa F ($gl_{\text{subjektuarte}} = \epsilon$; $gl_{\text{errorea}} = \epsilon'$) parametroaren banaketari hurbiltzen zaio. ϵ eta ϵ' osagaiek tratamenduaren askatasun graduak eta errorearen askatasun graduak zuzentzen dituzte hurrenez hurren. Hipotesi testa F-ren ($MC_{\text{subjektuarte}} / MC_{\text{errorea}}$) balio enpirikoa taulatan agertutako balioarekin konparatzean datza:

$$F[\alpha, (a-1)\epsilon, a(n-1)\epsilon'] = F[\alpha, 1, n-1] \quad (3.6)$$

Talde bakoitzeko subjektu kopurua desberdina denean zuzenketa konplexuagoa gertatzen da. Prozedura hori Boxen (1954) testuan kontsulta daiteke.

Brown eta Forsytheren F*-a (1984) F testaren eraldakuntzarako askotan erabiltzen den beste prozedura bat da. Faktore bakarreko taldearteko diseinuetan test hau jatorrizko F testaren taldearteko (kasu honetan A faktore nagusiaren) berreduren baturan oinarritzen da, zeinen izendatzailean beharrezko eraldakuntzak egiten diren.

$$SC_{\text{subjektuarte}} = \sum_{j=1}^a n_j (\bar{Y}_j - \bar{Y})^2 = \sum_{j=1}^a n_j \alpha_j^2 \quad (3.7)$$

Talde bakoitzean subjektu kopuru desberdina dagoenean, F* ondorengo formularen arabera kalkulatzen da:

$$F^* = \frac{\sum_{j=1}^a n_j \alpha_j^2}{\sum_{j=1}^a \left[1 - \left(\frac{n_j}{N} \right) \right] S_j^2} \quad (3.8)$$

F* testaren banaketa F testarenari hurbiltzen zaio zenbakitzaileari $a-1$ eta izendatzaileari f askatasun gradu dagozkielarik. Izendatzailearen askatasun graduak test honetarako bereziki kalkulatzen dira, $S_j^2 \sigma_j^2$ -ren alborapenik gabeko estimatzaile gisa erabiliz.

$$f = \frac{1}{\sum_{j=1}^a \left[\frac{g_j^2}{(n_j - 1)} \right]} \quad (3.9)$$

non:

$$g_j = \frac{\left[1 - \left(\frac{n_j}{N}\right)\right] S_j^2}{\sum_{j=1}^a \left[1 - \left(\frac{n_j}{N}\right)\right] S_j^2} \quad (3.10)$$

Talde bakoitzean subjektu kopuru berbera daukagunean, F^* ondorengo formularen bidez kalkulatzen da:

$$F^* = \frac{\sum_{j=1}^a n_j \alpha_j^2}{\sum_{j=1}^a \left[1 - \left(\frac{n_j}{N}\right)\right] S_j^2} \quad (3.11)$$

Kasu honetan izendatzailearen askatasun graduak ondorengoak dira:

$$f = \frac{(n-1) \left[\sum_{j=1}^a S_j^2 \right]^2}{\sum_{j=1}^a (S_j^2)^2} \quad (3.12)$$

Brown eta Forsytheren F^* -a naharo erabiltzen den arren, autore gehienen iritziz, bariantzak oso heterogeneoak diren kasuetan ezik, F testarekiko abantaila gutxi aurkezten ditu.

Bariantzen homogeneotasun ezaren aurrean erabil daitezkeen prozeduren artean **Welchen W testa** (1951) aipatuko dugu azkenik. Test honek ere, ez du eskatzen bariantzen homogeneotasunik. Welchen W testak bere tamainaren arabera pixu berezia ematen dio talde bakoitzari eta, ondorioz, zenbakitzailea ez da jatorrizko F -arenaren berdina. Bere formulazioa ondorengoa da:

$$W = \frac{\sum_{j=1}^a W_j (\bar{Y}_j - \tilde{Y})^2}{1 + \frac{2}{3}(a-2) \wedge} \quad (3.13)$$

non:

$$W_j = \frac{n_j}{S_j^2} \quad (3.14)$$

$$\tilde{Y} = \frac{\sum_{j=1}^a W_j \bar{Y}_j}{\sum_{j=1}^a W_j} \quad (3.15)$$

$$\Lambda = \frac{3 \sum_{j=1}^a \left(1 - \frac{W_j}{\sum_{j=1}^a W_j} \right)^2}{a^2 - 1} \quad (3.16)$$

Hipotesi nulua onartzen denean W-ren banaketa F testarenaren antzekoa da, zenbakitzaileari eta izendatzaileari $a-1$ eta $1/\Lambda$ askatasun gradu dagozkielarik hurrenez hurren.

Pascual, García eta Fríasek (1994) heterozedastizitatea zuzentzeko atal honetan zehar aipatu ditugun prozedurak nola aplikatzen diren ikusi ahal izateko adibideak eskaintzen dituzte. Halaber, taldeak orekatuak izatearen edo ez izatearen arabera, bariantzen homogeneousutasun mailaren arabera eta laginaren tamainaren arabera F, F* eta W testen ezaugarrien arteko erkaketa barne ematen duen taula aurkezten digute.

3.2.1.3.- Behaketen arteko askatasun aurretikoa

Aurretiko hau ikerketako behaketen arteko koerlazio ezari dagokio. Bi behaketa elkarren artean *askeak* dira, bata ematen delarik, bestearen probabilitate baldintzatua eta ez baldintzatua berdinak direnean. Halatan, aurretiko hau betetzen bada zera esan genezake: laginean sartzeko asmoz populaziotik edozein kasu hautatzek ez diola eragiten beste edozein kasu hautatzeko probabilitateari eta, modu berean, edozein kasuri esleituriko puntuazioa beste edozeini esleitutako puntuaziotik aske dagoela baieztatu daiteke ere. Aurretiko hau betetzea guztiz beharrezkoa da bariantza analisia egoki erabili nahi bada, beronen betekizun ezak eragin handia baitauka esanguratsutasun mailan eta test estatistikoaren indarran (Pascual eta

Camarasa, 1991). Barcikowsky (1981) eta Stevens (1992) bezalako autoreek, beste zenbaitzuren artean, behaketen arteko menpekotasunak I. motako errorearen gehikuntza izugarria dakarrela egiaztatu dute. Vallejok (1986a, 1986b) dioenez, askatasun aurretikoaren betekizun ezak II. motako errorean erortzera eraman diezaguke ere. Aurretiko hau ez betetzeak ekar ditzakeen ondorioen azterketa sakona Kenny eta Judden (1986) eta Scariano eta Davenporten (1987) lanetan aurki daiteke.

ANOVAren kasuan, behaketen arteko menpekotasuna klasebarneko koerlazio-koefizientea kalkulatzuz **froga daiteke** (Stevens, 1992). Koefiziente honek saiakuntza talde bakoitzaren barneko lotura maila adierazten du eta bere formulapena honako hau da:

$$R_{KLASEBARNEKO} = \frac{MC_{taldeartekoa} - MC_{errorea}}{MC_{taldeartekoa} + (n - 1) MC_{errorea}} \quad (3.17)$$

non $MC_{taldeartekoa}$ eta $MC_{errorea}$ F zatiketaren zenbakitzaileari eta izendatzaileari dagozkien hurrenez hurren, eta n -k talde bakoitzeko subjektu kopurua adierazten duen.

Behaketen arteko askatasun aurretikoaren betekizun eza zuzentzeko prozedurak erabiltzen den diseinuaren arabera aldatzen dira. *Subjektuarteko diseinu*ekin lan egiterakoan, datuen askatasuna ziurtatzeko prozedurarik arruntena *subjektuak baldintza esperimentaletara ausaz esleitzean datza* (Keppel, 1982).

Errepikapen neurketako edo subjektubarneko diseinuetan, non subjektu bakoitza saiakuntza baldintza bakoitzean eta guztietan behatzen den, behaketen arteko askatasuna nekez gaindi daitekeen arazo nagusia bihurtzen da. Kasu honetan, behaketen arteko kobariantza egon arren, berredura batzbestekoen arteko zatidurak F banaketari doi-doi jarraitzen dio baldin eta errepikapen neurketako batzbesteko bikote guztien bariantzen arteko desberdintasunak egonkorrak badira, hots, ortonormalizaturiko aldagaien bariantza-kobariantza matrizeak *esferizitate patroia* deritzona aurkezten badu (Huynh eta Feldt, 1970; Mendoza, 1980; Rouanet eta Lepine, 1970; Vallejo, 1991b). Esferizitate mota berezi bat *simetria konposatua* deitutakoa da, zeinen arabera errepikapen neurri bikote guztien arteko koerlazioek berdina izan behar duten. Laginketa bariantzen arteko berdintasuna eta koerlazioen arteko berdintasuna eskatzen ditu (Maxwell eta Delaney, 1990; Pascual eta Camarasa, 1991). Simetria konposatuaren betekizun ezak esferizitatearen etendura halaberharrez ez badakar ere, etendura hori seguruenik gertatu dela adierazten du. Puntuazioen arteko menpekotasuna kontrolatu ahal izateko, ohiko ANOVAren ordez gutxienez bi prozedura erabiltzea gomendatzen da. Zenbait autorek *aldagai anitzeko bariantza analisiaren (MANOVA)* erabilera proposatzen dute. ANOVAren kasuan ez bezala, non errore osagaiaren batzbesteko estimazioa burutzen den, MANOVAren kasuan kontrastapen bakoitza berorri dagokion errore osagaiarekiko egiten da, beraz esferizitatea ez da beharrezkoa (Pascual eta Camarasa, 1991). Bestalde, F testaren

askatasun graduak egokitzen dituen *F kontserbatzailea* izeneko testaren erabilera gomendatzen da, baita ere.

3.2.2.- Test estatistikoaren indar eskasa

Ondorio estatistikozko balideziaren ezaugarri nagusiak aztertzerakoan ikusi denez, *indarra* II. motako errorearen osagarria da eta hipotesi nulua faltsua denean berau baztertzeko daukagun probabilitatea bezala defini daiteke, $1-\beta$ bezala hain zuzen ere. Azken buru, *indarra* erabaki egokia hartzeko probabilitateari dagokio. Portaera zientzien eremuan ondorio estatistikozko balideziaren aurkako mehatxu nagusienetariko bat indar eskasa izatean datza, hots, desberdintasunak benetakoak direnean berauek hautemateko probabilitate txikia egotean.

Izan ere, test estatistikoaren potentzia edo *indarra* nahikoa ez bada, efektu ezen ondorio baino gehiago, II. motako errorearen gehikuntzaren ondorio izan daitezkeen hipotesi nulua aldeko emaitzak, edo taldeen arteko desberdintasun eza, aurki ditzakegu. Kazdin (1992) eta Lipseyk (1990) dioten bezala, arazo hau guztiz garrantzitsua da medikuntzaren eta psikologia klinikoaren eremuetan, gizabanakoen osasun fisiko eta mentalarentzat onuragarriak izan daitezkeen tratamenduak aurkitzeko ikertzaileak duen gaitasuna eragotz baitezake.

Indar maila hiru faktoreren menpe dago batik bat: (1) hipotesi nulua eta ordezkio hipotesiaren arteko diferentziaren edo *efektuaren tamaina* deritzonaren menpe, (2) *α esanguratsutasun mailaren* menpe eta (3) *laginaren tamainaren* menpe (Cohen, 1977; Maxwell eta Delaney, 1990). Aztertzen ari garen populazioan *efektuaren tamaina* handiagoa den heinean, ikerketan behatutako efektuak errazkiago lortuko du esangura estatistikoa, beraz *indarra* handiagoa izango da. Modu berean, *I. motako errore probabilitatea* (α) areagotzen den neurrian, *indarra* ere handitzen da. Datu berberak izanik, 0,05eko α -rekin 0,01ekoarekin baino indar gehiago lortzen da. Azkenik, *laginaren tamaina* handiagoa den heinean, garaia da efektua behatu ahal izatearen probabilitatea eta, ondorioz, *indarra* ere handitu egiten da. Cohen (1988) eta Kraemer eta Thiemanek (1987) ikerketa adibide ugari eskaintzen digute, non tratamenduen eragina hautemateko gaitasuna subjektu kopurua handitzearekin zuzenean erlazionaturik dagoela egiaztatzen den.

Bestalde, oraintxe aipatu ditugun faktorez gainera, *behaketen aldakortasunean* eragiten duen orok, hala nola menpeko aldagaiaren neurketa fidagarritasunak edo laginaren heterogeneotasunak, testaren indarrean eragiten du ere. Hortaz, ikerketa planifikatzerakoan, subjektuen homogeneotasuna eta neurketa tresnen fidagarritasuna bezalako alderdiak edo, azken batez, errore bariantza kontrolatu eta murrizteko aukera ematen diguten alderdi guztiak kontuan hartu beharko ditugu.

3.1. atalean **indarraren analisiari** buruz esan denarekin bat eginez, Pedhazur eta Pedhazur Schmelkinek (1991) honela definitzen dute aipatutako

analisia: esangura estatistikozko testetan parte hartzen duten lau elementuen, hots, efektuaren tamaina, I. motako errorea (α), II. motako errorea (β) eta laginaren tamainaren arteko erlazioak aztertzen dituen estrategia bezala. Lau elementu horietatik hiru finkatuz, laugarrena kalkula daiteke. Ondorioz, lau indar analisi mota egin daitezke. Alabaina, arruntena, ikerketa planifikatzerakoan aurrera eramaten dena alegia, bariantza zehatz batzutatik abiatuz eta α eta β balioak (beraz, indarra) finkatuz, nahi dugun tamainako efektua sumatu ahal izateko beharrezko subjektu kopurua mugatzean datza. Stevensen (1992) dioen bezala, itxarondako tratamenduaren efektuaren tamaina aurrez finka badaiteke, II. motako errorea gutxitzeko aukera emango ligukeen lagin tamaina egokia zehazteko analisi indartsua burutu daiteke. Aitzitik, itxarondako efektuaren tamainari buruz adostasunik ez dagoenean, hipotesi nulua baztertu ahal izateko beharrezkoa den gutxieneko efektuaren estimaziotik abiatuz kalkulatzen da subjektu kopurua.

Vallejori (1991a) jarraituz, behin aldagai askea eta menpeko aldagaia zehazturik, saiakuntzan behar dugun subjektu kopurua ondoko elementuak kontuan hartuz kalkula daiteke: (1) esperimendatzaileak lortu nahi dituen gutxieneko tratamendu efektuak; (2) erabili behar duen tratamendu maila kopurua; (3) populazioko bariantzari dagokion errorea (σ_ϵ^2); (4) I. motako errorearen probabilitatea (α) eta (5) II. motako errorearen probabilitatea (β).

Aldez aurretik egindako saiakuntza piloto batetik abiatuz σ -ren balioa finkatuz gero, indar estatistiko maila zehatza lortu ahal izateko beharrezko subjektu kopurua kalkula daiteke. Horretarako, parametro zehatz baten, esate baterako ϕ -ren, balioa kalkulatu eta hautatutako lagin tamainak planteatutako indarra irizpideari erantzuten ote dion egiazta dezakegu. Tangek (1938) ondoko formula proposatzen du ϕ parametroa kalkulatzeko:

$$\phi = \frac{\sum \left[\frac{\sum_{j=1}^p (\mu_j - \mu)^2}{p} \right]^{1/2}}{\left(\frac{\sigma_\epsilon^2}{n} \right)^{1/2}} \quad (3.18)$$

non:

$\mu_j - \mu$ = Esperimendatzaileak lortu nahi duen gutxieneko tratamendu efektua.

p = Tratamendu maila kopurua.

n = Talde bakoitzeko subjektu kopurua.

Indar analisi mota garrantzitsuena eta erabilgarriena oraintxe aurkeztu duguna, hots, ikerketa aurretik burutzen dena izan arren, ikerketa amaitu ondoren egiten denak gure inferentzietan egon daitezkeen errakuntzak atzematen lagun

diezaguke (ikus 3.1. atala). Beraz, edozein ikerketaren kalitate metodologikoa ebaluatu ahal izateko guztiz gomendagarria da. Cohenen (1977, 1988) testuek eta Kraemer eta Thiemanen (1987) monografiak test estatistikoen indar analisia sakon- - sakonean aztertzen dute. Analisi honetan trebatu nahi den ikaslea testu horietara igortzen dugu, beraz.

3.2.3.- Efektuaren tamaina

Coheni (1988) jarraituz, *efektuaren tamainak* ikerketa objektu den fenomenoaren populazioan zein mailatan agertzen den edo hipotesi nulua zein mailatan faltsua den adierazten du. Hipotesi testen eremuan efektuaren tamainak balio zehatza hartzen du, zein hipotesi nulupean zeroren berdina eta ordezkoko hipotesipean zerotik desberdina den.

Kazdinek (1992) baldintzen arteko diferentziaren edo tratamenduak izan duen eraginaren tamaina bezala definitzen du efektuaren tamaina. Autore honen aburuz, efektuaren tamaina ondorio estatistikozko balideziaren barnean kontzeptu nagusienetariko bat bezala har daiteke, beronen ikasketak arazo eta errore metodologiko anitz, kontrolatugabeko aldakortasun iturri posibleak esate baterako, hautemateko aukera ematen baitigu.

Definizio hauetatik abiatuz, tratamendu taldeen batzbestekoen arteko edo tratamendu talde(ar)en eta kontrol taldearen batzbestekoen arteko desberdintasunak **efektuaren tamainaren estimazio** bezala hartzen dira. Beste zenbait arrazoiaren artean, psikologian erabiltzen diren neurketa unitateen heterogeneotasuna dela eta, oso zaila gertatzen da diferentzia horien "maila" edo "tamaina" modu zehatzean definitzea. Horregatik, ikertzaile gehienek **efektuaren tamainaren neurri estandarrek** erabiltzera jo dute. Batzbestekoen arteko diferentzia estandarra, *d estatistikoa* deritzona, efektuaren tamaina neurtzeko sarri erabilitako indizeen arteko bat da. Bere formula honako hau dugu:

$$d = \frac{\mu_1 - \mu_2}{\sigma} \quad (3.19)$$

non μ_1 eta μ_2 bi talde edo baldintzen (e.b. tratamendu taldearen eta kontrol taldearen) lagin batzbestekoei dagozkien, eta σ -k desbidazio estandar amankomuna adierazten duen.

Bi batzbesteko baino gehiagoren arteko efektuaren tamaina estimatu nahi denean, *f estatistikoa* izeneko batzbesteko efektuaren tamaina kalkulatu daiteke ondorengo formula erabiliz:

$$f = \frac{\sigma_m}{\sigma_\varepsilon} \quad (3.20)$$

non:

$$\sigma_m = \sqrt{\frac{\sum(\mu_j - \mu)^2}{a}} = \sqrt{\frac{\sum(\alpha_j)^2}{a}}$$

taldearteko desbidazio estandarri dagokion eta σ_ε , berriz, taldebarneko desbidazio estandarri.

Aipatutako estimatzaileez gain, aldagai askeak(ek) azaldutako menpeko aldagaiaren bariantza proportzioa edo ikerketa objektu diren aldagaien bariantza proportzio amankomuna adierazten duten beste estimatzaile batzuk proposatu dira ere. Indize horien artean r^2 , R^2 , η^2 (*etaren bigarren berreketa*), ε^2 (*epsilonen bigarren berreketa*) eta ω^2 (*omegaren bigarren berreketa*) azpimarra daitezke (Haase, Ellis eta Ladany, 1989; Murray eta Dosser, 1987; O'Grady, 1982; Sechrest eta Yeaton, 1981a, 1981b, 1982; Smith, 1982; Strube, 1988). O'Gradyk (1982) adierazi duen bezala, efektuaren tamainaren estimazioa faktore metodologiko, estatistiko, soziometriko edo teoriko desberdinen menpe egon daiteke, horregatik ikerketak efektuaren tamainaren arabera elkarren artean erkatzea oso arazo konplexua da. Maila estatistiko hutsean ere, indizeek elkarren arteko desberdintasun garrantzitsuak aurkezten dituzte. Koerlazio-koefizienteek eta batazbestekoen arteko diferentzia estandarrek, esate baterako, efektuaren tamaina eta norabideari buruzko estimazioak eskaintzen dizkigute. Aipatutako η^2 edo ω^2 bezalako neurriek, ordea, efektuaren tamainari buruzko informazioa ekartzen digute soilik. Bestalde, R^2 -ren bidez egiten den estimazioa ω^2 estatistikoak eskaintzen diguna baino liberalagoa da (Maxwell eta Delaney, 1991). Halaber, η^2 -ren bidezko estimazioa laginaren tamainaren arabera alboratua egon ohi da.

Oraintxe esan dugunatik ondoriozta daitekeen bezala, efektuaren tamaina kalkulatzeko neurri estandarren multzo zabala daukagun arren, **tamaina zehatzeko efektu batek benetan adierazten duena ezagutzea** oso zaila gertatzen da. Arazo horri erantzun nahian, zenbait autorek portaera zientzien eremuan gehienbat erabilitako testei atxikitako efektuaren tamaina konbentzionalak proposatu dituzte. Autore horien artean Cohen (1988) dugu nagusi, zeinek koerlazioen, batazbestekoen arteko diferentzien eta beste zenbait testen kasuan, tamaina txiki, ertain edo handiko efektutzat har daitekeena proposatu duen. Gure ustez, Pedhazur eta Pedhazur Schmelkinek (1991) egoki dioten bezala, hitzarmen hauek efektuaren tamainaren eta berorrek maila teorikoan edo praktikokoan izan dezakeen garrantziaren arteko nahasketa sor dezakete. Hau da, tamaina handiko efektu bezala hartzen dena, garrantzitsu bezala oker interpreta daiteke. Alabaina, emaitzen esanahi praktikoari edo benetako esanahiari buruzko erabakia ikerketa eremu mugatu batean kokatu behar da, eta ikasketa objektuaren eta berorrengan eragin dezaketen faktore mota desberdinen ezagutza sakonean oinarritu behar da.

Indarraren kasuan egin dugun bezala, **efektuaren tamainari eragiten dioten faktore nagusienak** aztertzen ditugu jarraian. Aurreko atalean azaldutakoari eutsiz (3.2.2. puntua), argi ikus daiteke esanguratsutasun maila (α) eta laginaren tamaina egonkorra izanik, efektuaren tamainaren eta *testaren indarraren* artean erlazio zuzena dagoela, hots: efektuaren tamaina handitu ahala, testaren indarra handitzen da. Era berean, I. motako errorearen probabilitatea (α) eta II. motako errorearen probabilitatea (β) egonkor eutsiz, efektuaren tamaina handitzen den neurrian, balio zehatzeko eragin esperimentala hautemateko beharrezko *subjektu kopurua* txikiagoa izango da.

Bestalde, efektuaren tamaina batazbestekoen arteko diferentzia estandarra bezala, hots, sakabanatze unitateetan oinarrituz azaltzen duen definizioari helduz, erraz ikus daiteke efektuaren tamaina *errore bariantzaren* menpe dagoela. Hortaz, errore bariantza murriztu eta lehen mailako bariantza sistematikoa gehitu ahal izateko beharrezko prozedurak erabiliz gero, efektuaren tamaina areagotu daiteke. Psikologian, errore iturri nagusienak *saiakuntza subjektuen heterogeneotasuna* eta *menpeko aldagaiaren neurketa fidagarritasun eskasa* dira (beranduago aztertuko ditugu biak), horregatik faktore hauek kontrolpean izan beharko genituzke. Lehen mailako bariantza sistematikoaren gehikuntzari dagokionez, jarraitu beharreko prozedura, menpeko aldagai(et)an jazotako aldakuntzak sumatzeko aukera emango diguten *aldagai aske(ar)en balio egokiak hautatzean datza*. Aldagai aske(ar)en eta menpeko aldagai(ar)en artean lerro erlazioa egonez gero, lehen mailako bariantza sistematikoa areagotzeko modurik egokiena aldagai aske(ar)en muturretako balioak aukeratzean datza. Aitzitik, aldagaien arteko erlazio mota ezezaguna denean, aldagai aske(ar)en balio egokien hautapena saiakuntza pilotoetan edo *optimizazio estrategia* izenekoan oinarritu behar da. Estrategia hau elkarrekintza prozesu gisa uler daiteke, non ikertzaileak aldagaien portaera enpirikoaren ezagutzatik abiatuz beharrezko ondorioak ateratzen dituen.

Azkenik, Kazdinek (1992) dioten bezala, efektuaren tamaina ikerketaren plangintzari hertsiki lotzen zaiola esan genezake, ikertzaileak erabilitako diseinu eta metodologia motaren menpe baitago hein handi batetan. Zentzu honetan, ikasketa metaanalitiko anitzek agerian utzi dute saiakuntza diseinuek (ausazko esleipenaren irizpideari jarraitzen diotenek alegia) beste diseinu motek baino tamaina handiagoko efektuak lortzen dituztela. Halaber, kalitate metodologiko goreneko azterketek ere, metodologiari dagokionez eskasagoak direnek baino efektuaren tamainaren estimazio handiagoak aurkezten dituzte.

3.2.4.- Laginaren tamaina.

Aurreko bi ataletan (3.2.2 eta 3.2.3 puntuetan hain zuzen) azaldutakotik ondoriozta daitekeen bezala, eta Pascual eta Camarasaren (1991) argudioei jarraituz, ikertzaileak laginaren tamainari buruz hartu beharreko erabakiak izugarritzko eragina

dauka bai diseinuaren indarrean, bai efektuaren tamainan bai eta zoriaren bidez lor dezakegun kontrol esperimentalean ere.

Laginaren tamainak harreman zuzena dauka efektuaren tamainarekin eta testaren indar estatistikoarekin. Horrela, laginaren tamaina handitzen den neurrian, lagin horrek berorrek jatorrizko populazioa egoki adieraztearen probabilitatea areagotzen da. Aldakortasuna gutxitzean efektuaren tamaina handitu egiten da eta, hortaz, efektu esperimentala hautemateko probabilitatea ere gehitzen da.

Efektuaren tamaina bezala, laginaren tamaina ikerketa burutu aurretik finkatu beharreko parametroa da. Cohen (1977, 1988), Kraemer eta Thieman (1987) eta García, Pascual eta Fríasek (1993) zenbait taula eskaintzen dizkigute, non indarra, I. motako errorea eta efektuaren tamaina parametroen arabera ikertzaileak erabili beharko lituzkeen behaketen kopuru egokia azaltzen duten. Pascual, García eta Fríasek (1994) dioten bezala, **laginaren tamainaren eta oraintxe aipatutako elementuen arteko erlazioak** aztertzen direnean, lau baieztapen egin daitezke: (1) *esanguratsutasun maila* murrizgarriagoa den neurrian, handiagoa da balio zehatzeko eragina lortu ahal izateko beharrezko subjektu kopurua; (2) *indar estatistiko* gehiago lortu nahi izanez gero, ezinbestekoa da laginaren tamaina handitzea; (3) beharrezko subjektu kopurua gutxitu egiten da *efektuaren tamaina* handitzen den neurrian eta (4) *bi aldetako testekin alde bakarrekoekin* baino subjektu gehiago behar ditugu.

Subjektuarteko diseinuetan goian aipatutakoak baliagarri badira ere, *subjektubarneko diseinuetan* laginaren tamaina egokia zehaztea ez dago soilik, esanguratsutasun mailaren (α), indarraren, efektuaren tamainaren eta baldintza esperimentalen menpe; aurrekoek gainera, kontuan hartu beharreko beste parametro bat ere badago: *subjektuen behaketa desberdinen arteko koerlazioa*, hain zuzen. Honela, behaketen arteko koerlazioa handiagoa den neurrian, subjektuen portaera (menpeko aldagaia) egonkorragoa izango da, beraz aldakortasun gutxiago eta itxarondako efektua hautemateko probabilitate handiagoa izango dugu. Maxwell eta Delaneyk (1990) Voneshek (1983) parametro hauen elkarren arteko erlazioen inguruan errepikapen neurketako diseinuentzat sortutako taulen laburpena eskaintzen digute.

3.2.5.- I. motako errore maila

Talde batazbestekoen arteko erkaketa kopurua handitzen den neurrian, taldeen artean zoriaren ondorio diren desberdintasunak aurkitzearen probabilitatea edo kobariazioa dagoela oker ondorioztatzearena (I. motako errorea) esanguratsuki areagotzen da.

Arazo hau azaldu nahiean, zenbait autorek (e.b. Davis eta Gaito, 1984; Keppel, 1982) *konparaketa bakoitzeko errore mailaren (comparisonwise type I error rate)* eta *konparaketa sendi bakoitzeko errore mailaren (experimentwise type I error rate)* arteko bereizketa egin dute. **Konparaketa bakoitzeko errore maila (KEM)**, esperimendatzaileak ezarritako esanguratsutasun mailari dagokio ($\alpha=0,05$, $\alpha=0,01$, etab.) eta konparaketa bakarrean I. motako errorearen probabilitatea, hots, edozein konparaketa estatistikoki esanguratsua bezala erruz hartzearen probabilitatea adierazten du. **Konparaketa sendi bakoitzeko errore maila (SEM)**, berriz, saiakuntza batean burututako konparaketa multzoari dagokion errore maila da. Bi errore hauen arteko erlazioa ondorengo formularen bidez adieraz daiteke:

$$SEM = 1 - (1 - KEM)^c \quad (3.21)$$

non c ikertzaileak egiten dituen elkarren askeko konparaketa kopuruari dagokion. Hori horrela, saiakuntza batetan KEM $\alpha=0,05$ eko balioan finkatuz gero eta elkarren askeko $c=7$ konparaketa eginez gero, berauetako bat gutxienez erruz baztertzearen probabilitateak ondoko balioa hartzen du:

$$SEM = 1 - (1 - 0,05)^7 = 0,3017$$

guztiz onartezina dena, alegia.

Bestalde, ikerketa enpirikoetan ikusi ohi denez, c konparaketak elkarrekiko askeak ez direnean, SEM balorea $1 - (1 - KEM)^c$ delakotik oso hurbil dago ere (Toothaker, 1991).

Konparaketa sendi bakoitzeko errore maila doitzeko zenbait prozedura daude. Ryanek (1959) t -ren beste balio kritiko bat kalkulatzeko datzan metodoa proposatu du. Horretarako, finkatutako α balioa (e.b. 0,05) konparaketa kopuru ahalgarriaren artean zatitzen da eta, horrela, esanguratsutasun maila kontserbadoreagoa (α') lortzen da (aurreko adibidean $\alpha' = \alpha/c = 0,05/7 = 0,007$). Konparaketa multzo hortatik, α' -ri dagokion t kritikoaren balioa berdintzen edo gainditzen dutenak soilik, %5eko esanguratsutasun mailari atxikiz, hartzen dira esanguratsu bezala.

Hala ere, konparaketa anizkoitzetan errore maila ekiditeko prozedurarik egokiena *a posteriori edo post hoc izaeradun planifikatu gabeko konparaketak* erabiltzean datza, zeintzuk konparaketa bikote guztien kasuan errore mailak esangura maila berbera izan dezan eragiten baitute (Klockars eta Sax, 1985). Honi dagokionez, testarik egokienak *Tukeyren HSD testa (T metodoa)* eta *Scheffèren testa (S testa)* dira.

Beste aukera bat *Bonferroni*ren zuzenketa (*DB testa*) aplikatzean datza. Zuzenketa honek α errore maila ikertzaileak burututako konparaketa kopuruaren arabera doitzen du; Ryanek errore maila doitzeko proposatutako prozedurari jarraitzen dio beraz.

Konparaketa anizkoitzerako test nagusienak Hochberg eta Tamhane (1987), Keppel (1982), Kirk (1982) eta Winer, Brown eta Michelsen (1991) testuetan kontsulta daitezke. Ikerketa psikologikoaren eremuan sarritan erabiltzen diren konparaketa anizkoitzerako sei test oinarritzat hartuz, Harterren (1957) artikulua klasikoak lagingaren tamainaren eta konparaketa anizkoitzen kopuruaren arabera test horiei dagozkien I. eta II. motako errore mailak errepikapen neurketako diseinuen kasuan kalkulatzeko aukera ematen duten zenbait taula eskaintzen ditu.

Menpeko aldagai anitzeko diseinuekin lan egiten delarik menpeko aldagai bakoitzean faktore bakarreko analisi motaren bat aplikatuz gero, errore maila handitu daiteke ere. Kasu honetan, *α esanguratsutasun maila* goran aipatutakoenen antzeko prozeduren bidez *doi dezakegu* edo, bestela, *faktore anitzeko test estatistikoak erabil ditzakegu* (Johnson eta Wichern, 1988).

3.2.6.- Menpeko aldagaiaren neurketan fidagarritasun eta balidezia eskasa

Neurketa fidagarritasuna, interes gune den ezaugarria modu egonkorrean neurtzeari dagokio. Horrez gainera, *neurketak balidezia badauka*, interes gune den ezaugarria, eta ez neurri horren beste zenbait aldakortasun sorburu, islatuko du soilik.

Fidagarritasun eta balidezia eskaseko neurketa aldagaien erabilerak saiakuntza erroreak gehitzen ditu eta, beraz, errore bariantza areagotzen du. Hortaz, efektuaren tamaina gutxitzen da eta benetako eragina izanik, tratamenduek inolako eraginik ez dutela ondorioztatzeko probabilitatea (II. motako errorea) handitu.

Ato (1991), Judd eta Kenny (1981) eta Vallejori (1991a) jarraituz, menpeko aldagai bezala hartutako neurri (indikatzailerak) batean behatutako aldakortasuna hiru sorburutatik erator daiteke: (1) interes gune den konstruktua teorikoaren aldakortasunatik, (2) beste konstruktua batzuen aldakortasunatik eta (3) ausazko errorearen aldakortasunatik.

Tratamendua eraginkorra bada, interes gune den konstruktuari eragingo dio soilik (interbentzioaren ondoriozko bariantza); neurriaren beste zenbait aldakortasun sorburutan (errorearen ondoriozko bariantza), aldiz, ez du inolako eraginik izango. Zentzu honetan, indikatzailearen aldakortasunak interes gune den konstruktuari aldakortasuna besterik islatzen ez badu (neurketa fidagarritasun handia), II. motako errorea agertzeko probabilitatea gutxitu egiten da. Aitzitik, interes gune den konstruktuz kanpoko sorburuek sortutako aldakortasuna nahikoa handia bada edo

indikatzailearen aldakortasunak ez badu konstruktutik teorikoaren aldakortasuna modu zehatzean islatzen, II. motako errorea agertzeko probabilitatea esanguratsuki handitzen da. Nolanahi ere, aldagai nahastatzaileen kontrol ezak edo fidagarritasun eskaseko tresnen erabilerak fidagarritasun gutxiko neurketak eragiten dituzte eta lehen mailako bariantza sistematikoaren efektua ezagutzeko konparaketa irizpide gisa erabiltzen den errore bariantza handitzen dute.

Neurketaren fidagarritasun eskasa interes gune den konstruktutik kanpoko sorburuen aldakortasunak sortua bada, *kontrol teknika esperimentalak* (e.b. blokeo teknika) edo *estatistikoak* (e.b. kobariantza teknika) erabil daitezke errore bariantza gutxitzeko.

Bestalde, neurketaren fidagarritasun eskasa eragiten duena indikatzailearen zehaztasun gabezia baldin bada, **prozedura desberdinak** erabil ditzakegu **arazoa konpontzeko**. Horien artean, ondorengoak azpimarra daitezke:

- 1.- Menpeko aldagaia neurtzeko test psikometrikoren bat erabili bada, berorren fidagarritasuna hobetzeko aukera ematen duten *prozedura psikometriko* desberdinak aplikatu daitezke. Esate baterako, testaren luzera areagotzearena (Nunnally, 1978).
- 2.- Menpeko aldagaia neurtzeko test psikometrikorik erabili ez bada, *konstruktuen indikatzaile anitzekin lan egin daiteke* eta berauetatik neurri orokor bat eratorri.
- 3.- Beste aukera bat *laginketa unitate gisa unitate handiagoak erabiltzean* datza. Adibidez, subjektuekin lan egin ordez taldeak erabil daitezke, taldearen batazbestekoak banakoaren puntuazioak baino neurri egonkorragoak baitira.

3.2.7.- Tratamenduak ezartzerakoan fidagarritasun eskasa eta egiaztapen indar gutxiko tratamenduen aplikapena

Tratamenduak ez badira subjektu esperimentalen artean modu homogeneoan edo estandarrean banatzen, errore bariantza esanguratsuki handi daiteke. Honela, tratamenduak pertsona desberdinek ezartzen badituzte (*pertsonarteko desberdintasunak*), edo subjektu berberak une desberdinetan ezartzen baditu (*pertsonabarneko desberdintasunak*), aldakortasunak jasan dezakeen hazkuntzak tratamendua eraginkorra izanik tratamendu horrek berorrek eraginik ez duela ondorioztatzea eraman gaitzake (II. motako errorea). Arazo honen aurrean zera egin beharko genuke: aldakortasuna ahalik eta gehien murriztearren, *aldagai aske(ar)en mailak prozedura objektibo edo estandarren bidez banatu*, hain zuzen ere. Tratamenduak ezartzeko modua ez da ezertan aldatu behar pertsona batetik bestera, leku batetik bestera edo une desberdinetan zehar. Azken buru, baldintza esperimental

berberaren partaide diren subjektu guztiek tratatu berdina jaso behar dute, eta baldintza batetik bestera aldatu behar den gauza bakarra subjektuei banatzen zaizkien aldagai aske(ar)en maila (edo mailen arteko konbinaketa) da.

Bestalde, aldagai askearen mailen hautapena lehen mailako bariantza sistematikoa handitu ahal izateko egokia ez bada, II. motako errorea agertzearen probabilitatea handi daiteke. Kasu honetan, *zoru efektuak* (banakoen puntuazioak neurraketa eskalaren balio txikienetan metatzen dira) edo *sabai efektuak* (banakoen puntuazioak neurraketa eskalaren balio handienetan metatzen dira) deritzotenak ager daitezke (Jung, 1982). Arazo hau konpontzeko *aldagai askearen balio egokiak hautatu behar dira*, hots, menpeko aldagaian jazotako aldakuntzak egokiro hautemateko aukera ematen dituzten mailak.

3.2.8.- Saiakuntza unitateen arteko ausazko heterogeneotasuna

Subjektu esperimentalak oso desberdinak izan daitezke elkarren artean ezaugarri askotan, hala nola sexua, adina, estatus sozioekonomikoa, nortasun ezaugarriak, etab. Oro har, saiakuntza unitateen arteko heterogeneotasunak subjektuek tratamendu zein neurretan aurrean aurkeztutako erreakzioen aldakortasuna (errore bariantza) handitzen du eta, ondorioz, tratamenduaren eragina hautemateko probabilitatea murrizten du. Arazoa are larriagoa da ezaugarri indibidualen bat tratamendu aldagaiarekin herstuki harremandurik egonez gero, kasu honetan aldagai bien eraginaren arteko nahasketa eman baitaiteke eta, beraz, emaitzak zein aldagaik azaltzen dituen jakitea ezinezkoa gerta dakiguke.

Mehatxu hau kontrolatzeko prozedura desberdinak ditugu:

- 1.- Lehenik, *lagin homogeneoak aukeratu edo eratu* ditzakegu. Honi dagokionez, kontuan hartu behar diren aldagaien eta onar daitekeen aldakortasun mailaren inguruko erabakia, aldez aurreko teoriaren menpe edo kontrolatu beharreko aldagaiak interes gune den fenomenoan duten eraginari buruzko ezagutzaren menpe egongo da. Laginaren ausazko hautapena eta subjektuak tratamendu taldeetara ausaz esleitzearen teknika, saiakuntza taldeen arteko aldez aurreko baliokidetasuna lortzeko psikologian gehien erabilitako prozedurak dira.
- 2.- Bigarrenik, lagin heterogeneoekin lan egin dezakegu baina interesatzen zaizkigun ezaugarrien eragina diseinuaren bidez ebaluatzeko aukera ziurtatuz. Horretarako, ezaugarri horiek guztiak erreferentzia gisa hartuz, subjektu laginen barnean geruza desberdinak bereizten dira eta elkarren artean heterogeneoak diren blokeak eratzen dira, bloke bakoitzaren barneko subjektuak homogeneoak izanik. *Blokeo teknika* diseinuan bertan

tratamendu aldagaia eta blokeaturiko aldagaia elkarren artean konbinatzean datza, errore bariantza gutxitu edo bi aldagaien arteko elkarreragindakoa aztertzearen, hain zuzen ere. Beste alderditik, subjektuei atxekiriko errore bariantza kontrol estatistikoaren bitartez edo *kobariantza teknika* erabiliz murriz daiteke baita ere. Aipatutako teknika, menpeko aldagaiarekin koerlazio handia duen aldagai baten ezagutzatik abiatuz aurrean daitekeen aldakortasuna errore bariantzatik ebaztean datza.

- 3.- Hirugarrenik, *subjektubarneko diseinuak (edo diseinu mistoak) erabil daitezke*. Horrela, efektuaren tamaina subjektubarneko aldakortasun sorburu edo errore osagai egokiak erabiliz kalkula daiteke; subjektuen arteko desberdintasunen menpe egon ordez, subjektu berberak tratamendu desberdinak erantzuterakoan igarotzen dituen uneen arteko desberdintasunen menpe dauden errore osagaiak erabiliz, alegia.

3.2.9.- Ausazko aldakortasunak saiakuntza egoeran

Tratamenduaz gain, saiakuntza egoerako beste zenbait ezaugarriak ere menpeko aldagaiaren neurketan eragin dezakete eta errore bariantza areagotu. Hau da, badira egoera faktore batzuk subjektuaren aditasuna erakar dezaketenak, subjektuaren zereginetan akatsak sortaraz ditzaketenak, etab. eta, azken buru, emaitzen zergatien bihur daitezkeenak edo tratamenduaren efektuarekin elkarreragin dezaketenak. **Mehatxu hau kontrolatu ahal izateko**, Cook eta Campbell (1979) *aldakortasun sorburu arrotzik gabeko ikerkuntza testuinguruak*, hots, subjektuaren aditasuna tratamenduarengan eraman eta ingurune faktoreen eragina murrizteko aukera ematen duten testuinguruak *hautatzea* gomendatzen dute.

Atok (1991) dioen bezala, guztiz kontrolaturiko testuinguruak ikerketaren izaera esperimentalaren denean erabil daitezke soilik. Ikerketa esperimentalaren ez bada, erabilgarria gerta dakiguke tratamendu baldintza guztientzat amankomunak diren bariantza arrotzaren sorburuak neurtu eta kontrol estatistikoaren bidez (hondakinen analisiaren bidez, esate baterako) sistematizatzea.

Cook eta Campbell (1979) ereduari jarraituz psikologiaren eremuan ikertzaile gehienek aztertu dituzten ondorio estatistikozko balidezaren aurkako bederatzi mehatxu hauez gainera, Pascual, García eta Fríasek (1994) gure ustetan kontuan hartu beharko liratekeen beste bi mehatxu ere proposatzen dizkigute.

3.2.10.- "Outliers" edo ibiltartez kanpoko balioak

"Outliers" deritzoten puntuazioak muturretako puntuazioak dira, beste datu guztietatik desberdintzen direnak eta berauen banaketa desitxuratzen dutenak (García, Pascual eta Frías, 1992; Pascual eta Camarasa, 1991; Winer, 1971).

Muturretako balioek oso eragin negatiboa izan dezakete test estatistikoaren aplikapen eta interpretapenean, gehienetan prozedura estatistikoak balio hauen agerpenarekiko oso sentikorrak baitira.

Stevensen (1992) aburuz, "outliers" puntuazioak datuak biltzerakoan errore batzuk egiten direlako edo zenbait puntuazio gainerakoekin konparatuz egiazki desberdinak direlako ager daitezke. Balio horiek agertzen badira, autoreak *joera zentraleko neurri gisa mediana erabiltzea* gomendatzen du, neurri hau berau ez baitago muturretako puntuazioen eraginaren menpe. "Outlier" a datuak gaizki bildu direlako agertuz gero, *datua zuzendu* eta analisiak ohi bezala egin daitezke. Muturreko balioa neurketa tresnen akatsen ondorioz agertzen bazaigu, *balio hori egotz dezakegu*. Puntuazio hauen agerpena goran aipatutako arrazoi bidez ezin azal daitekeenean, Stevensek *bi analisi mota egitea* gomendatzen du, *"outliers" balioak erabiliz kasu batean eta balio horiek egotziz bestean*.

3.2.11.- "Missing" edo balio ez-agertuak

"Missing" balioak balio ahantziak edo ikertzailearen esku ez dauden balioak dira. Balio ahantzien izateak ondorio estatistikozko balideziari negatiboki eragiten dio orekatu gabeko taldeak (subjektu kopuru desberdineko taldeak) sortarazten dituelako eta, ondorioz, eredu estatistikoari dagozkion aurretikoen betekizun ezaren efektuak areagotzen dituelako. Ordenagailuetarako programa estatistiko gehienek "missing" balioak ordezkatzeko doikuntza prozedurak barne ematen dituzte. García, Pascual eta Fríasek (1992) aipatzen duten bezala, arazo honen aurrean normalki egiten dena zera da: *balio ez-agertuari dagokion batazbestekoa gainerako puntuazioetatik abiatuz estimatu*, hain zuzen ere. Elkarreagindakorik gabeko diseinu faktorialen kasuan, Winerrek (1971) zera proposatzen du: balio ez-agertuak batazbesteko orokorrari lerro eta zutaberi dagozkien batazbestekoen batura kenduz lortutako balioez ordezkatzea, alegia.

3.4. Taulan ondorio estatistikozko balideziaren aurkako mehatxu nagusien eta berauek kontrolatzeko erabil daitezkeen prozeduren laburpena aurkezten da.

3.4. Taula.- Ondorio estatistikozko balideziaren aurkako mehatxu nagusienak eta berauek kontrolatzeko erabil daitezkeen prozedurak

MEHATXUA	DEFINIZIOA	KONTROLERAKO PROZEDURAK EDO TEKNIKAK
1.- Eredu estatistikoaren baldintzen betekizun eza ANOVA-ren kasuan betekizun eza ondoko hauetan zehazten da: (a) Normaltasun eza (b) Heterozedastizitatea (c) Independentzia eza	Edozein analisi estatistiko aplikatu ondoren burututako inferentziak egokiak izan daitezen, test estatistikoak bete beharreko oinarritzko aurretikoen betekizun eza (a) M.A.-n lortutako puntuazioen banaketak ez darraio kurba normalari (b) Saiakuntza talde desberdinen bariantzak esanguratsuki desberdinak dira elkarren artean (c) Behaketen artean koerlazioa dago	(a)- Teknika parametrikokoak bezain murrizgarriak ez diren analisi teknikak (test ez-parametrikokoak) erabili - Aldagaietan eraldakuntzaren bat egin (b)- O'Brienen prozedura - F kontserbatzailea izeneko testa - Brown eta Forsytheren F^* testa - Welchen W testa (c)- Subjektuak saiakuntza baldintzetara ausaz esleitu (subjektuarterko diseinuak) - Bariantza-kobariantza matrizeak esferizitate baldintza bete behar du (subjektubarneko diseinuak) - MANOVA - F kontserbatzailea izeneko testa
2.- Test estatistikoaren indar eskasa	H_0 -a faltsua izanik berau baztertzeko edo taldeen arteko desberdintasunak benetakoak izanik berauek hautemateko probabilitate txikia	- Laginaren tamaina handitu - Efektuaren tamaina handitu errore bariantza kontrolatzeko eta murrizteko aukera ematen duten prozedurak erabiliz, hots: aldagaien balio egokiak hautatuz, aldagaiak egokiro operazionalizatuz, neurketa fidagarriak burutuz, etab.
3.- Efektuaren tamainari buruzko erabaki desegokiak hartu	H_0 -aren faltsutasun mailari buruzko edo baldintzen arteko desberdintasun mailari buruzko erabaki desegokiak hartu	- Efektuaren tamainaren neurri estandarrak erabili - Tamaina zehatzeko efektu baten adiera egokiro interpretatu ahal izateko faktore mota desberdinak kontuan hartu. Aintzartat hartu beharreko faktoreak estatistikoak (indarra, I. eta II. motako erroreak, laginaren tamaina...), metodologikoak (subjektuen arteko heterogeneotasuna, neurketa fidagarritasuna, A.A-ren mailen hautapena, diseinu mota...) eta teorikoak (efektuaren garrantzia maila teorikoan, aplikazio eremua, interes gune den fenomenoaren ezaugarri bereziak...) izan behar dira

3.4. Taula (jarraipena).- Ondorio estatistikozko balideziaren aurkako mehatxu nagusienak eta berauek kontrolatzeko erabil daitezkeen prozedurak

MEHATXUA	DEFINIZIOA	KONTROLERAKO PROZEDURAK EDO TEKNIKAK
4.- Subjektu kopuru urriaz osatutako lagina	Laginaren tamaina murrizteak eragin negatiboa dauka diseinuaren indarrean, efektuaren tamainan eta zoriaren bidez lor daitekeen kontrol esperimentalean	- Lagina osatzen duten subjektuen kopurua handitu
5.- I. motako errore maila	Talde batzbestekoen arteko erkaketa kopuruaren hazkuntzak esanguratsuki areagotzen du kobaria-ziorik ez dagoenean kobariazioa dagoela ondorioztatzeko probabilitatea (I. motako errorea)	- Ryanen metodoa - <i>A posteriori</i> edo <i>post hoc</i> planifikatu gabeko konparaketak (e.b. Tukey-ren HSD testa eta Schefféren testa) - Bonferroniren zuzenketa
6.- Menpeko aldagaiaren neurketan fidagarritasun eta balidezia eskasa	Fidagarritasun eta balidezia eskaseko neurketa aldagaien erabilerak, benetako eragina izanik tratamenduek inolako eraginik ez dutela ondorioztatzeko probabilitatea (II. motako errorea) handi-tzen du	- Kontrol teknika esperimentalak edo estatistikoak erabili - Prozedura psikometrikoak erabili - Konstruktuaeren indikatzaile anitz erabili - Laginketa unitate gisa unitate handiagoak erabili
7.- Tratamenduak ezartzerakoan fidagarritasun eskasa eta egiaz-tapen indar gutxiko trata-menduen aplika-zioa	-Tratamendu mailak ezartzerakoan pertsonartean zein pertsonabarnean eman daitekeen homogeneousotasun ezak, benetako eragina izanik tratamenduek inolako eraginik ez dutela ondorioztatzeko probabilitatea (II. motako errorea) handitzen du -A.A.-ren mailen hautapen desegokiak II. motako errorea agertzeko probabilitatea handitzen du	- Tratamenduak prozedura estandarren bidez banatu edo hauek pertsona batetik bestera, leku batetik bestera eta une desberdinetan zehar modu berean ezartzen direla ziurtatu - A.A.-ren balio egokiak hautatu
8.- Saiakuntza unitateen arteko ausazko heterogeneotasuna	Saiakuntza unitateen arteko ausazko heterogeneotasunak, tratamendu esperimentalak eragindako efektuak hautemateko probabilitatea esanguratsuki murrizten du	- Zoriaren bitartez subjektu lagin homogeneousoak eratu - Blokeo teknika - Kobariantza teknika - Subjektubarneko diseinuak edo diseinu mistoak erabili
9.- Ausazko aldagakortasunak saiakuntza egoeran	Tratamenduz kanpo, saiakuntza egoerako zenbait ezaugarrik menpeko aldagaiaren neurketan eragin eta errore bariantza handi dezakete	- Guztiz kontrolatutako eta aldakortasun sorburu arrotzik gabeko ikerkuntza testuinguruetan lan egin (ikerketa esperimentalak) - Bariantza arrotzeko sorburuak kontrol estatistikorekin bidez sistematizatu (esperimentalak ez diren ikerketak)

3.4. Taula (jarraipena).- Ondorio estatistikozko balideziaren aurkako mehatxu nagusienak eta berauek kontrolatzeko erabil daitezkeen prozedurak

MEHATXUA	DEFINIZIOA	KONTROLERAKO PROZEDURAK EDO TEKNIKAK
10.- "Outliers" edo ibiltartez kanpoko balioen agerpena	Muturretako balioen ("outliers" delakoen) agerpenak, test estatistikoen aplikazio eta interpretazioan akatsak egitera eraman diezaguke, prozedura estatistiko gehienak balio hauen agerpenarekiko sentikorrak baitira	- Joera zentraleko neurri gisa mediana erabili - Datuak gaizki bildu direlako agertutako "outliers"ak zuzenean zuzendu daitezke - Neurketa tresnen akatsen ondorioz agertutako "outliers"ak egotz daitezke - "Outliers"en agerpena ezin badaiteke aurrez aipatutako bi arrazoen arteko baten ondorio bezala azaldu, bi analisi- mota egin ditzakegu: muturretako balioekin bata eta balio horiek egotziz bestea
11.- "Missing" edo balio ez-agertuak	Balio ez-agertuen ("missing" delakoen) agerpenak orekatu gabeko taldeak sortarazten ditu eta eredu estatistikoari dagozkion aurretikoen betekizun ezaren efektuak areagotzen ditu	- "Missing" balioari dagokion batzbestekoa gainerako puntuazioetatik abiatuz estimatu

4.- Barne balidezia

4.1.- ..Barne balideziaren definizioa eta ezaugarri nagusiak

4.2.- Barne balideziaren aurkako mehatxuak

4.2.1.- Historia

4.2.2.- Heltzea

4.2.3.- Testen aplikazioa

4.2.4.- Neurketa baliakizuna

4.2.5.- Tratamendu anizkoitzen arteko
elkarreragindakoa

4.2.6.- Erregresio estatistikoa

4.2.7.- Subjektuen hautapen diferentziala

4.2.8.- Hilkortasun esperimentala edo subjektuen
galera selektiboa

4.2.9.- Subjektuen hautapen diferentziala eta beste
zenbait mehatxuren arteko elkarrera-
gindakoak

4.2.10.- Kausazko eraginaren norabideari buruzko
zalantzagarritasuna

4.2.11.- Egoerak eragindako efektu erreaktiboak

4.3.- Barne balideziaren aurkako mehatxuak kontrolatu
ahal izateko prozedurak

4.3.1.- Ezabapena

4.3.2.- Konstantzia

4.3.3.- Orekatzea subjektuarteke diseinuetan

4.3.3.1.- Zorizkapena

4.3.3.2.- Parekatzea

4.3.3.3.- Doikuntza estatistikoa

4.3.4.- Kontrabalantzeoa edo ekiponderazioa

4.- BARNE BALIDEZIA

4.1.- Barne balideziaren definizioa eta ezaugarri nagusiak

Behin bi aldagaien artean kobariazioa dagoela egiaztatu ondoren, arazoa erlazioaren izaera kausala ote den jakitean datza. Horretarako, kausalitatearen norabidea zehazteaz gainera, manipulaturako faktorez kanpoko aldagai batek edo gehiagok behatutako efektuaren zergati gisa joka ez dezan ziurtatu behar dugu.

Barne balidezia aldagai askeak(ek) menpeko aldagai(et)an izan dezake(t)en eraginari buruz ondorio egokiak lortzeko probabilitateari dagokio. Hortaz, *bigarren mailako bariantza sistematikoa* edo, aztertu nahi dena (direnak) ezik, saiakuntza subjektuen erantzunetan eragiten duen efektu sistematiko oro kontrolatzearen menpe dago.

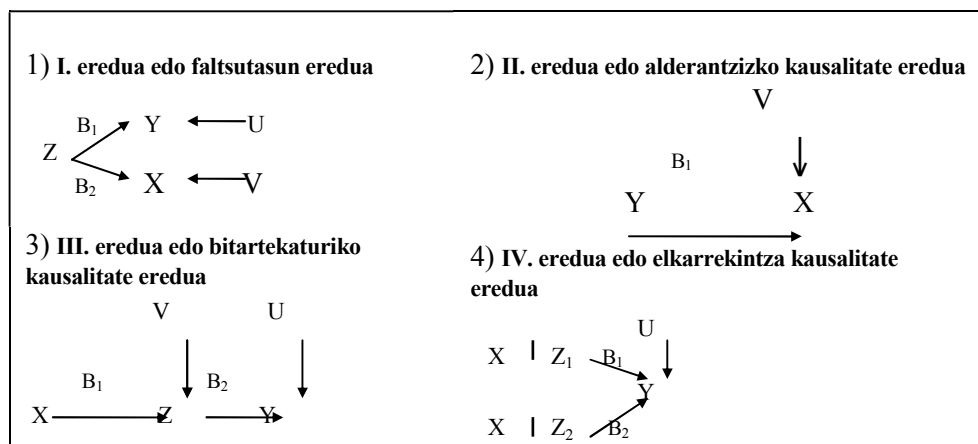
Diseinuak egokiagoak izango dira barne balidezia handiagoa lortzen duten neurrian. *Ikerketa diseinu* batek *barne balidezia* dauka, saiakuntza eta kontrol taldeen artean menpeko aldagai(et)an behatutako desberdintasunak, aldagai aske(et)an eragindako aldaketen ondorio bezala soilik azal daitezkeenean, hots, diseinu horrek berorrek hipotesi lehiakide edo ikasitako fenomenoaren ordezkari azalpen guztiak saihesteko aukera ematen duenean.

Maxwell eta Delaneyk (1990) diotenaren arabera, *barne balideziaren arazoa hirugarren aldagaiaren arazoa* bezala ezagutzen da ere. Arazo hori ondorengo galderaren bidez adieraz daiteke: X (aldagai askea) eta Y (menpeko aldagaia) aldagaien arteko ustezko erlazioaren, edo erlazio gabeziaren, erantzule den *hirugarren aldagaiaren* bat, Z delakoa alegia, ba ote dago? Campbelli (1986) jarraituz, *barne balidezia "kausazko balidezia"* bezala uler daiteke, zeren beronen helburu nagusia menpeko aldagai(et)an behatutako aldaketen eragile bakarra ikertzaileak finkatutako zergatia edo baldintza izan dadin ziurtatzean baitatza. Atok (1991) dioenez, egoera ideal horri *norabide bakarreko kausa erlazio zuzena* deritzo, erlazio horretan bi aldagaik soilik parte hartzen dutelako eta kausalitatearen norabidea aldagai asketik (X-etik) menpeko aldagaiera (Y-ra) doalako hain zuzen ere.

Psikologiaren eremuan nahiko zaila gertatzen da norabide bakarreko kausa erlazio zuzenak aurkitzea. Aitzitik, esparru honetan topatzen ditugun erlazio gehienetan *hirugarren aldagai* batek edo *Z aldagaia* delakoak parte hartzen du,

partehartze hori tankera desberdinekoa izan daitekeelarik. *Hirugarren aldagaiaren* arazoa saiakuntza diseinuari baino gehiago, sasisaiakuntzari dagokio, horregatik kausa inferentzia askoz konplexuagoa gertatzen da azken diseinu horretan lehenengoan baino. Izan ere, sasisaiakuntza ikerketaren eremuan laginketa unitateak populazio desberdinetatik lortzen dira gehienetan, eta horrek hautapen alborapena eragin eta nahasketa iturrien edo ordezeko kausazko faktore posibleen kontrola murrizten du. Arnauri (1995) jarraituz, sasisaiakuntza diseinuekin lan egiterakoan aurki ditzakegun *ordezko azalpen eredu* nagusienak aurkezten ditugu jarraian (ikus 4.1. Irudia):

- 1.- **I. eredua edo faltsutasun eredua:** aldagai askearen eta menpeko aldagaiaren arteko erlazioa kausazko hirugarren aldagai baten ondorioa da. Kasu honetan, elkarren arteko harremana *faltsua* dela diogu.
- 2.- **II. eredua edo alderantzizko kausalitate eredua:** eredu hau menpeko aldagaia aldagaia askearen eragilea denean aurkitzen dugu.
- 3.- **III. eredua edo bitartekaturiko kausalitate eredua:** kasu honetan ekintza kausala ez du X-k zuzenean eragiten, kausazko faktore sail batek baizik.
- 4.- **IV. eredua edo elkarrekintza kausalitate eredua:** eredu honen arabera, faktore baten kausazko eragina beste faktore batengatik baldintzatuta dago.



4.1. Irudia.- Ordezko azalpen eredu desberdinak sasisaiakuntza ikerketaren eremuan (Arnau, 1995)

Dwyerrek (1983) lau ordezeko azalpen eredu hauei bosgarren bat gehitzen die, “aipatu ereduaren arteko edozein konbinaketa posible” bezala definitzen duena hain zuzen ere, eta eredu guztiak *zorizkoa ez den zeharkako diseinuekin* zuzen

zuzenean lotzen ditu. Autoreak ordezeko eredu bakoitzari dagozkion path diagramak eta egiturazko ekuazioak aztertzen ditu eta zientzia sozialen eremuko adibide desberdinez hornitzen ditu. Halaber, *zorizkoak ez diren zeharkako diseinuekin* erkatuz, *zorizkoak ez diren luzerako diseinuek* eta, neurri handiagoan, *saiakuntza diseinuek* kausa inferentzia ziurtasun gehiagoz egiteko aukera eskaintzen digutela azpimarratzen du.

Cook eta Campbellek (1979) dioten bezala, kausalitatearen norabidearen zehaztapena ikertzaileak gertaeren denborazko sekuentziari buruz daukan ezagutzaren menpe dago. Ezagutza hau erraz lor daiteke ikerketa esperimentaletan, baina esperimentalak ez direnetan, berriz, nekez eskura dezakegu. Izan ere, ikerketa esperimentalak esperimentala ez denak baino barne balidezia handiagoa edukitzearen funtsezko arrazoiak ondoko biak dira: lehenik, saiakuntza diseinuak aldagai aske(k) edo ustezko zergatia(k) manipulatu ahal izateko eskaintzen duen aukera eta, bigarrenik, diseinu honetan saiakuntza unitateak edo subjektuak aldagai aske(ar)en mailetara ausaz esleitzeko aukagun aukera. Bi baldintza hauek, aldagai arrotzen batek aldagai aske(ar)en eta menpeko aldagai(ar)en arteko erlazioan eragiteko probabilitatea esanguratsuki murrizten dute.

Bestalde, sasisaiakuntza eremuan bertan, *luzerako diseinuak zeharkako* edo *taldeen arteko erkaketa diseinuak* baino barne balidezia handiagoa duela esan beharra dago. Arnauk (1995) dioen bezala, kausalitatea eta denboratasuna elkar loturik daude, horregatik denbora dimentsioa diseinuari atxiki izateak kausa inferentzia errazten du. Halaber, *luzerako diseinuak* gizabanako bakoitzaren aldakuntzak zuzenean neurtzeko aukera dakar, aldakuntza horiek subjektuaren arteko desberdintasunetatik abiatuz inferitzeko beharrik gabe alegia, eta, ondorioz, edozein aldagairen ekintza kausala zehaztasun gehiagoz finka dezakegu. Autoreak gehitzen du badirela zenbait luzerako egitura, *panel diseinuak* direlakoak esate baterako, faltsutasun hipotesia egiaztatzeko aukera ere eskaintzen dutenak eta egiaztapen hori nola egin daitekeen aurkezten du. Cook, Campbell eta Peracchiok (1990) eta Campbellek (1988) ere sasisaiakuntza diseinu mota desberdinen barne balidezia zehatz-mehatz aztertzen dute eta, beste zenbait baieztapen egiteaz gain, ondorengoa adierazten dute: zorizko esleipenik ematen ez bada ere, sasisaiakuntza eremuan kausa inferentzia lor daitekeela diseinuaren beraren zenbait ezaugarri hobetzen badira. Ezaugarri hauen artean, besteak beste, hurrenak aipatzen dituzte: *testaurre* neurketa kopuruaren gehikuntza, erkaketa talde kopurua handitzea, tratamenduak prozedura desberdinen bidez behin eta berriz egoki ezartzea, kontrol taldeen erabilera, neurketa prozedura anitzen erabilera, etab.

Bai Vallejok (1991a) eta baita Atok (1991) ere, *barne balideziaren* eta *ondorio estatistikozko balideziaren* arteko nahasketarik ez dela egin behar azpimarratzen dute. Bigarren balidezia mota ez dago alborapen sistematikoa dakarten sorburuen menpe, zorizko errorea eragiten dutenen menpe baizik, eta

interbentzioaren efektuaren ebaluapen estatistikoari dagokio funtsean. Barne balideziak, ordea, laginketa neurriak eta, beraz, tratamenduaren efektuen estimazioa alda dezaketen alborapen sistematikoko sorburuen eragin zuzena jasaten du.

Barne balidezia gainerako balidezia motetatik bereiztearren, Campbell (1986) balidezia hau berau izendatzeko "*Kausazko Balidezia Lokala Molarra (Pragmatikoa, Teoriazkanpoko)*" adierazpena erabiltzea iradokitzen du. *Psikologian erabiltzen diren diseinuen balideziari buruzko planteamendu berriei* eskaintzen diegun atalean proposamen honetan sakonduko dugun arren, ondoko lerroetan Campbell (1986) proposatutako adierazpenaren elementu nagusien esanahia laburki azalduko dugu. Lehenik, barne balidezia *kausazkoa* da, menpeko aldagai(et)an behatutako aldaketan eragileak edo zergatiak zehaztu nahi baititu. *Molar* hitzaz zera adierazi nahi da: tratamendua ez dela gauza isolatu gisa ulertu behar, ikerketaren urrats mugatu batean parte hartzen duten elementuen multzo konplexua bezala baizik. Gainera, barne balidezia *pragmatikoa* eta *teoriazkanpoko* da, bere helburu nagusia tratamenduak eragina duela edo interes gune den fenomenoan aldaketaren bat sortzen duela ziurtatzean baitatza. Azkenik, *lokaltzat* har daiteke, emaitzen generalizazio maila alde batera utziz, tratamenduak testuinguru eta une zehatzetan izan dezakeen eraginari baitagokio.

Sax eta Fineren (1981) definizioari jarraituz, *barne balideziaren aurkako mehatxua* behatutako emaitzen azalpen lehiakide bezala joka dezakeen ikerketa aldagaia da. Zentzu honetan, barne balideziaren aurkako mehatxuak aztertuz, ikertzaileak kausalitatezko inferentzia faltsuak egiteko probabilitatea murriz dezake (Arnau, 1981b). Mehatxu hauetako batzuk *subjektuarterko diseinu*ekin lan egiten denean, hots, tratamendu baldintza desberdinak talde desberdinei esleitzen zaizkienean hartu behar dira kontuan; beste batzuk, berriz, askoz garrantzitsuagoak dira *subjektubarneko diseinuak* erabiltzean, subjektu bakoitzak aldagai askearen balio guztiak jasotzen dituenean alegia, eta azkenik, badira bi diseinu motatan kontrolatu beharreko zenbait mehatxu ere. Jarraian, barne balidezia zalantzan jar dezaketen faktore nagusienak deskribatu eta mehatxu hauei aurre egiteko erabili beharreko kontrol teknikak edo prozedurak aipatzen dira. Ondorio estatistikozko balideziaren kasuan bezala, Cook eta Campbell (1979) eredu hartzen dugu oinarritzko erreferentzia gisa. Gai honetan interes berezia duten irakurleek, mehatxuen jatorrizko zerrendaren antolakuntzan zenbait aldaketa planteatzen dituen proposamen berri bat aurki dezakete Ato (1995), Ato eta Navalon (1993) eta Cook, Campbell eta Peracchioren (1990) lanak kontsultatuz.

4.2.- Barne balideziaren aurkako mehatxuak

4.2.1.- Historia

Historia, aldagai asketik desberdina izanik, saiakuntza egoeraren barnean edo egoera horretatik kanpo gertatzen den eta emaitzak azal ditzakeen edozein egitate edo gertaerari dagokio (Kazdin, 1992).

Barne balideziaren aurkako mehatxu hau deskribatzerakoan, zenbait autorek barne eta kanpo historiaren arteko bereizketa (Kish, 1987) edo historia proaktiboa eta erretroaktiboaren artekoa (Robinson eta Foster, 1979) egiten dute. *Barne historia* edo *historia proaktiboa*, subjektuak elkarren artean desberdintzera eramaten gaituzten ikasitako faktore nahiz faktore genetiko guztiei dagokie. Faktore horien artean sexua, altuera, pisua, jarrerak, pertzepzio eta mugimendu gaitasunak, buru gaitasunak, etab. aipa daitezke. Vallejoren (1991a) iritziz, *barne historia bigarren mailako bariantza sistematikoa* eragiten duten zergatien arteko garrantzitsuena da. Mehatxu hau subjektu talde desberdinak, *subjektuarterko diseinuak* alegia, erabiltzen dituzten ikerketatan agertzen da eta beronen eragina areagotzen doa subjektuak elkarren artean bereizten dituzten ezaugarriek menpeko aldagaiarekin koerlazio handiagoa aurkezten duten neurrian. Barne historiaren efektuak kontrolatzeko modurik egokiena *ausazko esleipena* erabiltzean datza.

Kanpo historia edo *historia erretroaktiboa*, tratamendu aurretik eta tratamendu ostean hartutako neurrien artean, edo aldagai askearen ezarpenaren eta menpeko aldagaiaren neurketaren artean, esperimentuan bertan edo egoera horretatik kanpo jazotako gertaera eta gora-beheren multzoari dagokio. Kontzeptu hau autore gehienek *historia* deritzon mehatxua azaltzerakoan proposatzen duten definizioarekin bat dator. Suposa dezagun, esate baterako, *heziketa sexualari buruzko programa* batek Euskal Herriko nerabeen *portaera sexualen aldakuntza* izan dezakeen eragina ikertu nahi dugula. Horretarako, galdesorten (edo antzeko proben bidez) nerabe talde baten portaera sexuala bederatzi hilabetetan zehar (ikasturtean zehar, hain zuzen) ebaluatzen da. Ondoren, aditu talde batek sexu heziketari buruzko programa ematen die eta, azkenik, beste bederatzi hilabetetan zehar eta neurketa prozedura berberaz, ikerketan parte hartu dutenen jokabide sexuala berriz ere ebaluatzen da. Horrela, subjektubarneko erkaketa egiten da, non portaera sexuala tratamendu aurreko eta osteko neurketen artean aurkitutako desberdintasunen arabera ebaluatzen den. Alabaina, behatutako ondorioa (jokabide sexualaren aldakuntza) proposatutako zergatiak (heziketa sexualari buruzko programak) eragin duela baieztatzea nahikoa arriskutsua da, testaurrearen eta testostearen artean igarotako epe luzeaz gainera, ikerketan zehar eragin dezaketen eta subjektuaren ohitura sexualak eralda ditzaketen zenbait kanpo aldagai baitaude, hala nola aditu desberdinek programa garatzeko erabiltzen dituzten estrategia desberdinak, nerabeek gaiari buruz beren gurasoeekin mantentzen dituzten elkarriketak, etab. *Kanpo historia* nahasketa sorburu izan

daiteke *subjektubarneko diseinuetan* eta, oro har, elkarrengandik denbora tarte luzez urrunduriko *tratamendu aurreko eta osteko neurriak* hartzen diren egoeratan. Neurketa hauen artean igarotako epea handitzen den heinean, mehatxua agertzeko probabilitatea ere gehitu egiten da.

Laboratorioko ikerketan, subjektuak kanpo eraginetatik urrun izanik edo itxuraz eragin horien menpe ez dauden menpeko aldagaiak hautatuz **kontrola daiteke historia**. Mehatxu honen kontrola askoz zailagoa gertatzen da lekuan lekuko ikerlanetan.

Zenbait kasutan, *historia kontrolatua izan daiteke* diseinuan *kontrol talde* bat erabiliz, zeinak tratamenduaren efektua berorretatik kanpo dauden eta prozesuen ondorio diren gertakizunetatik bereizteko aukera ematen duen. Beste kasu batzutan, mehatxu honen kontrola oso zaila suertatzen da, bi edo talde gehiagotan modu desberdinean eragiten duten gora-beherak ematen direnean batik bat. Kasu honetan, *tratamendu eta kontrol baldintzak* testuinguru berberan, denbora epe berberan eta saio bakarrean ezarri behar dira, hots, *konstantzia* izeneko teknika erabili behar da, honela historiaren eragina egonkor mantentzen baita subjektu guztietan zehar. Mehatxu honek *denbora bolada* diseinuen barne balidezari eragiten dionean, Arnauk (1990) *esperimentua fase gehiago erabiliz errepikatzea*, hau da, denbora bolada diseinu konplexuagoren bat erabiltzea gomendatzen du. Diseinu hauetan ikertzaileak interbentzio puntua bere kabuz aukeratzen duenez, jazokizun historikoa tratamenduaren ezarpenarekin bat etortzea nahikoa zaila da.

4.2.2.- Heltzea

Historia ez bezala, *heltzea* denboraren igarotze hutsaren ondorioz gertatzen den garapen biologikoa zein psikologikoa dela eta, organismoak jasaten dituen barne baldintzen aldakuntzei dagokie, eta hazkundera, ikaskuntza, adinaren gehikuntza, nekea, aspergura, gosea, etab. bezalako prozesuak barnean ematen ditu.

Prozesu hauek barne balidezia zalantzan jartzen dute diseinuak berauen eragina tratamendu esperimentalek sortutako eraginetatik bereizteko aukerarik ematen ez duenean. Mehatxu hau *historiarekin* batera agertzen da sarritan, horregatik zenbait kasutan oso zaila gertatzen da bi mehatxuen arteko bereizketa egitea. *Heziketa sexualari buruzko programak* Euskal Herriko nerabeen *portaera sexualesan* izan dezakeen eraginari buruz aurreko atalean (4.2.1. puntuan) aurkeztutako adibidearekin jarraituz, emaitzak azal ditzaketen kanpo aldagaien efektuak sortarazitako nahasketaz gain (hots, *historiaren* eraginaz gain), *heltzearekin* erlazionaturiko beste arazo nagusi bat ager daiteke ere. Seguraski, 18 hilabetetan zehar (ikerketak irautean duen denbora

tartean zehar, hain zuzen), nerabeetan heldutasun prozesu bat jazoko da eta oro har, beraien gaitasun psikologikoa programan parte hartu aurretik baino handiagoa izango da. Hori horrela, programak sexu jokabidearen aldaketan izandako eragina oso nekez bereiz genezake subjektuen garapen psikologiko handiagoak eragindako portaera aldaketatik.

Historia bezala, heldzea ere taldebarneko diseinuekin lan egiten denean eta, oro har, denbora tarte handiaz aldendutako testaurre eta testoste neurriak hartzen direnean kontrolatu beharreko mehatxua dugu. Zentzu honetan, Arnauk (1990) zera dio: heldzea epe luzeetan zehar egindako neurketa desberdinak barne ematen dituzten saiakuntza egoerei, luzerako diseinuei adibidez, dagokiela batik bat.

Historiaren kasuan bezala, kontrol taldearen erabilerak eta konstantzia izeneko teknikak mehatxu hau diseinu klasikoetan kontrolatzeko aukera ematen digute. Denbora bolada diseinuetan, berriz, Arnauk (1990) saiakuntza errepikatzea gomendatzen du, honela oinarri lerroak elkarren segidako errepikapenetan zehar finkatzen joango baitira eta, ondorioz, tratamendua aplikatu aurretik inolako aldaketarik ez dela eman ziurta genezake.

4.2.3.- Testen aplikazioa

Testen aplikapena, test bat aplikatzeak ondoko test batean lortutako emaitzetan izan dezakeen eraginari dagokio. Subjektu esperimental bereberei test bat edo gehiago behin eta berriz aplikatzen bazai(zki)e, lehen aplikapenak subjektuak, beste zenbait alderdiren artean, zeregin motari, ikerketaren helburuei, ikertzailearen itzaropenei edo aldagai aske(ar)en ezaugarriei dagokienez zorrotzagoak ager daitezen eragin dezake, edo testarekin praktika edo trebetasun gehiago lortzeko aukera eman diezaieke. Nolanahi ere, testak aldez aurretik edo behin eta berriz aplikatzeak balidezia arazo serioak planteatzen ditu, menpeko aldagaian jazotako aldaketei buruzko interpretazioak nahas baititzaie.

Demagun, esate baterako, Euskal Herrian A deritzon hizkuntza ereduan euskararen ikaskuntza hobetzeko bi metodo desberdinen eraginkortasuna aztertu nahi dugula. Edozein metodo ezarri aurretik, ikasleek euskararen ezagutza maila neurtzera zuzendutako test bati erantzuten badiote eta ikaskuntza epea amaitu ondoren test bera berrirori burutzen badute, galdera batzuk oroituko dituztela edo testa lehen aldiz erantzutean sortu zitzaizkien zenbait zalantza argitu izan dituztela ere, suposa daiteke. Hori horrela, subjektuek bigarren testan aurkeztutako hobekuntzaren zati bat, test bera aldez aurretik erantzutearen ondorio izan liteke.

Lehenago aipatu den bezala, barne balideziaren aurkako mehatxu hau subjektuen erantzunak behin eta berriz neurtzen direnean, *subjektubarneko*

diseinuetan alegia, eta, oro har, interbentzio aurretik testaurrea ezartzen denean ager daiteke.

Testak behin eta berriz aplikatzeak eragindako alborapena kontrolatu ahal izateko, tratamendurik jasotzen ez duten baina test horiei erantzuten dieten *kontrol taldeak* erabil ditzakegu.

4.2.4.- Neurketa baliakizuna

Saiakuntza subjektuen portaeren neurketa, zehatza eta egonkorra izan behar da denboran zehar. Eskakizun hau betetzen ez denean, *neurketa baliakizuna* izeneko mehatxuarekin topo egiten dugu. *Neurketa baliakizuna*, behaketa edo/ta neurketa tresnetan nahiz menpeko aldagaia ebaluatzeko prozeduretan denboran zehar jazotzen diren aldaketei dagokie.

Christensenek (1988) dioenez, errore sorburu hau behaketa *gizabanakoek* burutzen dutenean agertzen da funtsean, hauek nekea, aspergura, lekorapena, ikaskuntza, etab. bezalako barne eraginen menpe baitaude. Zentzu honetan, behatzaileak behar adina trebeak ez izatea eta, ondorioz, neurriek eraginkorrak ez diren tratamenduen aurrean ere aldaketak jasatea gerta daiteke. Bestalde, saiakuntza aurrera doan heinean esperimintatzaileek praktika edo trebetasun gehiago lor dezakete ere eta kasu honetan, *testaurre* eta *testostearen* artean, esate baterako, ager daitezkeen diferentziak, ezarritako tratamenduen ondorio baino gehiago, esperimintatzaileak berak jasandako aldakuntzen ondorio izan litezke.

Halaber, *testen* edo *neurketa tresnen* ezaugarriek beraiek aldaketak jasan ditzakete denboran zehar (kalibratze akatsak), saiakuntzaren barne balidezia zalantzan jarriz. Neurketa tresnei dagokienez, tarte edo puntu zehatz batzutan tartetzko neurketa eskalei berez atxikiriko ezaugarrien jabe ez badira, hots, neurriaren metrika probaren zenbait puntutan aldatzen bada, *eskalen gutxitze efektuak* deritzotenak ager daitezke. Efektu hauek, emaitzen interpretapena neurri handian korapilatzen dutenak, bi motatakoak izan daitezke: *sabai efektuak*, subjektuen puntuazioak neurketa eskalaren maila gorenetan metatzen direnean, edo *zoru efektuak*, puntuazio gehienak eskalaren balio baxuenen inguruan geratzen direnean. Vallejok (1991a) dioen bezala, tratamendu desberdinei dagozkien neurriak *konstruktua* beraren *indikatzaileak* izan arren, metrika desberdinak aurkeztzen badituzte nahasketa sor dezakete, neurketa eskala amankomuna lortu ahal izateko eraldakuntzaren bat egiten ez bada bederen.

Azken buru, arrazoia edozein delarik, tratamendu baldintza desberdinei dagozkien behaketak egiterakoan menpeko aldagaiaren neurri desberdinak erabiltzen badira, taldeen artean diferentzia estatistikoki esanguratsuak ager daitezke

tratamenduek inolako eraginik ez badute ere. Honela, *neurketa baliakizunak* barne balidezia zalantzan jar dezake, bai tratamendu bakoitza pertsona talde desberdin batek jasotzen duenean eta baita pertsona bakoitzari tratamendu baldintza desberdinak esleitzen zaizkionean ere. Arnauren (1990) iritziz, mehatxu hau *taldearteko erkaketa diseinuetan*, *errepikapen neurketako diseinuetan* eta, bereziki, *denbora bolada diseinuetan* agertzen da. Pascual, García eta Fríasek (1994) diotenez, berriz, *neurketa baliakizuna* batez ere *subjektubarneko edo errepikapen neurketako diseinuetan* kontrolatu beharreko mehatxua da.

Neurketa baliakizunaren efektuak *neurketa tresna estandarrak edo automatikoak erabiliz kontrola ditzakegu*. Era berean, behatzaileak edo kalifikatzaileak eragindako neurketa erroreak kontrolatzeko erabilgarriak gerta daitezkeen beste zenbait prozedura ere badaude, hala nola *kontrabalantzeoa edo ekiponderazioa*, *itsu bikoitza kontrol teknikak*, *behatzaile edo kalifikatzaileen trebakuntza*, *fidagarritasun egiaztapen anizkoitzak*, etab.

4.2.5.- Tratamendu anizkoitzen arteko elkarreragindakoa

Mehatxu hau aurreko biekin eta, bereziki, *testen aplikapenarekin* hertsiki erlasionaturik dago. *Tratamendu anizkoitzen arteko elkarreragindakoa*, subjektu guztiek saiakuntza baldintza guztiak jasotzen dituztenean agertzen da eta bi *nahasketa sorburu mota* nagusi eragiten ditu: *hurrenkera efektuak* eta *hondakin efektuak* edo "*carry-over*" *efektuak*, alegia. Normalki, efektu mota biak *sekuentzia efektuak* bezala ezagutzen dira.

Hurrenkera efektuak, tratamenduak subjektuengan ezartzerakoan jarraitutako ordenak eragiten duen alborapenari dagozkio. Nahasketa sorburu hau, aldagai arrotzek, lehen tokian, bigarrenetan, hirugarrenean eta horrela elkarren segidan ezartzen diren tratamenduetan eragin desberdina dutelako agertzen da. Normalean, hurrenkera efektuak lehen eta azken tratamenduaren artean igarotako denbora tartean zehar bilakatzen diren zenbait faktoreren ondorioz sortzen dira, ikaskuntza eta trebezia (*hurrenkera efektu positiboak*) edo frustrazioa eta nekea (*hurrenkera efektu negatiboak*) bezalako faktoreren ondorioz, hain zuzen ere. Faktore hauek denboraren igarotzearekin zuzenean erlasionatzen dira eta saiakuntzaren barnean une desberdinetan aurkeztutako tratamenduetan modu desberdinean eragiten dute. Demagun, esate baterako, *zereginaren konplexutasunak* O.H.O.-ko 6. mailako haur talde batek *zeregin desberdinak burutzeko behar duen denboran* izan dezakeen eragina aztertu nahi dela. Esperimentatzaileak hitzezko arrazonomendu lan desberdinak, bakunak zein konplexuak, aurkeztu dizkio haur talde berberari eta, ondoren, haur hauek berauek aurkeztu zaizkien lanak egiteko behar duten denbora neurtzen du. Egoera honetan, haurrek zeregin bakun guztiak, esate baterako, lehen tokian burutzen badituzte, nekez jakin genezake emaitzetan

agertutako diferentziak tratamendu esperimentalen ondorio izan ote diren edo aldagai askearen maila bakoitzean eragin desberdina duten zenbait aldagai kontaminatzailek, hala nola ikaskuntza edo nekea bezalako aldagaiek, sortu ote dituzten.

Saiakuntza talde bereko subjektu guztiek tratamendu anizkoitzak erantzutetik eratorritako bigarren nahasketa sorburua *hondakin efektuak* edo "*carry-over*" *efektuak* izenez ezagutzen da. Efektu hauek, tratamendu baten eragina hurrengo tratamendua ezartzeko unean oraindik desagertu ez denean agertzen dira, hots, eragin horrek tratamendu epea amaitu ondoren jarraitzen duenean. Efektuak denbora epe bakar batean zehar jarraitzen badu, *lehen mailako hondakin efektua* eragiten duela esaten da. Hondakin efektuak tratamendu osteko denbora epe bat baino gehiagotan zehar hedatzen badira, berriz, *k ordeneko "carry-over" efektuak* bezala ezagutzen dira. Efektu horiek agertuz gero, menpeko aldagaian behatutako diferentziak lehen tratamenduaren, hurrengoaren edo, lehenaren eragin luzea dela eta, bien arteko interakzioaren ondorio ote diren jakitea ezinezkoa da. Barne balidezia zalantzan jartzeaz gainera, hondakin efektuek kanpo balidezari eragiten diote baita ere, ikerketa zehatz batean erabilitako tratamenduen sekuentzia berbera barnean ematen ez duten beste egoera batzutura emaitzak orokortzeko posibilitatea murrizten baitute.

Tratamendu anizkoitzen arteko elkarreragindakoak mehatxu garrantzitsua suposatzen du *subjektubarneko diseinuekin* lan egiten denean eta, oro har, denbora tarte handiaz aldendutako *testaurre eta testoste neurriak* hartzen diren egoeretan. **Mehatxu hau kontrabalentzeo edo ekiponderazio teknikaren bat erabiliz kontrola daiteke.** Beren egituran bertan kontrol prozedura hori barne ematen duten diseinuen, hots, *cross-over* edo *ordezko diseinuen* eta *errepikapen neurketako karratu latinozko diseinuen* erabilerak tratamendu anizkoitzen arteko elkarreragindakotik eratorritako mehatxua kontrolatzeko aukera ematen du baita ere.

4.2.6.- Erregresio estatistikoa

Erregresio estatistikoa deritzon fenomeno, neurketa sistema berbera bigarren aldiz aplikatzerakoan muturretako puntuazioek banaketaren batazbestekorantz jotzeko aurkezten duten joerari dagokio (Neale eta Liebert, 1973). Mehatxu hau agertzen zaigu subjektu esperimentalak berauek aldagairen batean lortutako muturretako puntuazioen arabera esleitzen direnean taldeetara. Demagun, adibidez, *depresioak aditasunean* izan dezakeen eragina aztertu nahi dela. Horretarako, depresioa neurtzen duen galdesorta batean subjektuek aurkeztutako muturretako balioetatik abiatuz, depresio handiko eta txikiko talde bana hautatzen dugu. Kasu honetan, zera aurrean daiteke: testa berriro aplikatzen bada, puntuazio baxuko taldeak lehen neurketan aurkeztutakoa baino batazbesteko puntuazio handiagoa lortuko duela eta, era berean, depresio maila handiko taldeak aurreko neurketan baino balio txikiagoak eskuratuko dituela, hain zuzen ere. Hots, bigarren aplikazioetan puntuazioek jatorrizko populazioaren batazbestekorantz jotzen dute.

Hori horrela izanik, ikertzaileak benetan adierazi beharko luketena adierazten ez duten talderekin lan egiten du. Gainera, aldagai askearen eragina nahas daiteke eta, beraz, barne balidezia murriz daiteke interbentzioaren efektua puntuazioek batazbestekorantz hurbiltzeko aurkezten duten berezko joeratik ezin bada bereiz.

Cook eta Campbelli (1979) jarraituz, batazbestekoarekiko erregresioaren zergatiak *neurketa tresnen fidagarritasun eza* eta *borondatez aukeratutako azpitalde baten batazbestekoaren eta berari dagokion populazioaren batazbestekoaren arteko diferentzia* izango lirateke. Neurketa tresnaren zehaztasuna eta fidagarritasuna hobetzen diren neurrian eta batazbestekoen arteko desberdintasuna gutxitzen den neurrian, erregresio estatistikoaren eragina ere murriztu egiten da.

Ikertzaile gehienek erregresio estatistikoa neurketa errorearen ondorio hutsa bezala azaltzen duten arren, Pedhazur eta Pedhazur Schmelkinek (1991) zera azpimarratzen dute: *fenomeno hau bi aldagaien artean koerlazio perfekturik ez dagoen orotan agertzen dela*, eta hori neurketa erroreren edo aldagai bakoitzari dagozkion beste zenbait ezaugarriren eraginez gerta daiteke. Ausazko neurketa erroreek erregresio estatistikoan izan dezaketen efektuari dagokionez, ikertzaileek diote errore horiek bi aldagaien artean edo bi une desberdinetan neurtutako aldagai beraren bi puntuazioen artean egon daitekeen koerlazioa ahultzen dutela. Halatan, zorizko erroreak areagotzen diren neurrian, erregresio estatistikoa ere handitzen doa. Fenomenoa oso erraz nabaritzen da muturretako puntuazioak dituzten subjektuen kasuan. Ondorioz, goian aipatu dugun bezala, lehenengo neurketan puntuazio oso handiak edo oso txikiak lortzen dituzten subjektuek batazbestekoarekiko erregresioa aurkeztuko dute bigarren neurketan. Arrazonamendu berberari jarraituz, bi aldagai aldi berean neurtzen direnean, bietako batean muturretako puntuazioak eskuratzen dituzten subjektuek beste aldagaian kopuru eta norabide berdineko ausazko errorerik ez dutela aurkeztuko suposa daiteke, beraz bigarren faktorean beren puntuazioak ez dira lehenengoan bezain muturretakoak izango.

Mehatxu hau kontuan hartu beharrekoa da *subjektubarneko diseinuetan* edo subjektu talde berberan *testaurre eta testoste neurriak* hartzen direnean. **Erregresio estatistikoa taldeak zoriz osatuz murriz daiteke.** Hori posible ez bada, hots, taldeak aurrez finkatutako irizpide baten arabera osatu behar badira, *tratamendua aplikatu aurretik zenbait neurketa egitea* eta *subjektuaren puntuaziorik fidagarriena erabiltzea* gomendatzen da (Arnau, 1990).

4.2.7.- Subjektuen hautapen diferentziala

Subjektuen hautapen diferentziala, *tratamendua* ezarri aurretik saiakuntza talde desberdinen artean egon daitekeen baliokidetasun ezari dagokio eta subjektuak taldeetara esleitzeko erabilitako prozedurak taldeen arteko baliokidetzat ziurtatzen ez

duenean agertzen zaigu. Jakina den legez, interbentzioa aplikatu aurretik talde desberdinen artean diferentzia sistematikorik ez dagoela ezin ziurta badaiteke ere, tratamenduaren efektua ezin bereiz daiteke taldeen arteko berezko desberdintasunak dakarkigunatik, eta horrek serioski eragiten dio barne balideziari. Demagun, esate baterako, *arrazonamendu logikoa garatzeko metodo berri baten eraginkortasuna* aztertu nahi dugula. Gure hipotesia egiaztatzeko asmoz, Instituto jakin batean U.B.I.-ko 3. mailako bi gela hautatzen ditugu. Hauetako batean, ikasleek metodo berria erabiltzen ikasteko zenbait eskola hartzen dituzte. Beste gelan, berriz, ez da inolako interbentziorik egiten. Egoera honetan, bi taldeen artean arrazonamendu logikoan desberdintasun esanguratsuak hautemango balira ere, oso zaila gertatuko litzateke zergati-ondorio izaerako inferentziak burutzea, ezingo baikenuke jakin desberdintasunak metodo berriaren, bi gelatako ikasleen arteko alde aurreko desberdintasunen edo bi faktore hauen arteko interakzioaren ondorioz agertu ote diren.

Subjektuen hautapen diferentziala *ausazkoak ez diren esleipen prozedurak* erabiltzen dituzten *taldearteko diseinuetan*, hots, *sasisaiakuntza ikerketetan* agertzen zaigu batik bat. *Saiakuntza diseinuetan* subjektuak tratamendu baldintzetara zoriz esleitzen direnez gero, hautapen diferentzialak ez du eragin garrantzitsurik. Halaber, mehatxu honek ez gaitu kezkatu behar *subjektubarneko diseinuekin* lan egiten dugunean.

Oraintxe esandakotik ondoriozta daitekeen bezala, **hautapenak eragindako errorea kontrolatu ahal izateko prozedurarik egokiena zorizkapena** da. Teknika honek saiakuntza taldeen arteko baliokidetasuna ziurtatzen du eta manipulazio esperimentalak eragindakoa ezik, beste edozein efektu sistematikorik ez dela egongo segurtatzen du. Dena den, badira mehatxu hau kontrolatzeko erabil daitezkeen beste zenbait teknika, hala nola *blokeo teknika*, *subjektu baliokideetan oinarritutako parekatzea*, *kobariantza analisia*, etab.

4.2.8.- Hilkortasun esperimentala edo subjektuen galera selektiboa

Hilkortasun esperimentala, arrazoi desberdinak direla eta, hala nola irizpide batzuri ez jarraitzea, gaixotasuna, heriotza, nekea, etab. denboran zehar saiakuntzaren barnean jazo daitekeen subjektu galerari dagokio. Kazdinek (1992) dioten bezala, hilkortasuna nahitaezkoa gertatzen da denbora epe luzetan zehar hedatzen diren saiakuntzetan, eta barne balidezia esanguratsuki eragozten du modu selektiboan ematen bada, hots, saiakuntza talde desberdinei maila desberdinean eragiten badie. Kasu horretan, subjektuen banaketa desberdina izango da talde desberdinetan, hau da, baliokideak ez diren taldeen aurrean egongo gara eta, beraz, ezingo dugu bereiztu emaitzak tratamenduaren ondorioz edo taldeen arteko diferentzia sistematikoen ondorioz agertu ote diren.

Adibide gisa, demagun *herri mailan nagusi den hizkuntzak (gehiengo euskalduna vs. gehiengo erdalduna) subjektuek euskarari buruz daukaten jarreran* izan dezakeen eragina aztertu nahi dugula. *Gizabanako beraren hizkuntzaren ezagutzak* ere berorrek hizkuntza horrekiko agertzen duen jarrerari eragiten diola suposatuz, euskaldun eta erdaldun kopuru berbera esleitzen da saiakuntza talde bakoitzera. Erdaldunek ikerketa objektuari buruz daukaten interes eza dela eta, gehiengo erdaldunetz osatutako herriari dagokion taldeko subjektu erdaldunen %35ak ikerketa uzten du. Analisi estatistikoek hurrengo adierazten dute: euskararekiko jarraera esanguratsuki positiboagoa dela biztanle erdaldun gehiago dituen herriko taldean, gehiengo euskalduneko herriari dagokion taldean baino. Baldintza hauetan, honelako inferentzia burutzea desegoki litzateke, gehiengo erdalduneko herriari dagokion taldean, gehiengo euskaldunekoan baino subjektu euskaldun gehiago egongo bailirateke eta, hortaz, probabilitatearen arauak errespetatu nahiko bagenu, ez litzateke zilegi izango taldeak elkarren artean konparatzea. Azken buru, emaitzak aldagai askearen ondorio baino gehiago, hilkortasun esperimentalak eragindako alborapenaren ondorio izan bailitezke.

Subjektuak ikerketatik arrazoi zehatzen ondorioz egozten badira, hilkortasuna mehatxu garrantzitsua da *subjektuarterko diseinuekin* lan egiten denean. Alabaina, *errepikapen neurketako diseinue*tan ere errore sorburu izan daiteke subjektuen galera selektiboa bada, hots, subjektu trebeenek, esate baterako, ikerketan jarraitzen badute eta trebezia gutxiagokoek, berriz, berau uzten badute, edo alderantziz (Pascual, García eta Frías, 1994).

Behin agertu ondoren, **hilkortasun diferentziala oso nekez kontrola daiteke**. Izan ere, galdu direnen ordainean beste subjektu batzuk zoriz lortzea ez da lan erraza eta, gainera, taldeen artean diferentzia sistematiko gehiago eragin ditzake. Taldeak berriro orekatzearren subjektuak iraziteak zenbait arazo ekar diezazkiguke ere, hala nola laginaren tamaina murriztea, testaren indarra gutxitzea, etab. Horregatik, *kontrol prozedurarik gomendagarriena, mehatxu hau saihestearren alde aurretik ahalik eta neurri gehienak hartzean* edo *ikerketa, bertan hilkortasun esperimentalak agertzeko probabilitatea txikia izan dadin antolatze*ko moduan datza: saiakuntza gehiegi ez luzatu, oso aspergarriak edo ezatseginak suerta daitezkeen egoerak saihestu, gehiegizko eskakizun irizpideak ez ezarri, etab.

Aurreko zortzi ataletan azaldu den legez, barne balideziaren aurkako zenbait mehatxu ikerketa diseinu zehatzekin lan egiten dugunean agertzen zaizkigu soilik. Zentzu honetan, *subjektuen hautapen diferentziala* errore sorburu izan daiteke *subjektuarterko diseinuak* erabiltzen ditugunean; *historia, heltzea, testen aplikapena, tratamendu anizkoitzen arteko elkarreragindakoa* eta *erregresio estatistikoa*, berriz, barne balidezia zalantzan jar dezakete *subjektubarneko diseinuekin* lan egiten dugunean edo *tratamendu aurreko neurriak* hartzen ditugunean. Haatik, subjektuak

ez badira subjektuarterko diseinu bateko baldintza desberdinetara zoriz esleitzen eta beren erantzunak denboran zehar neurtzen badira, hautapenak subjektubarneko diseinuei dagozkien mehatxuekin edo diseinu mota bietan eman daitezkeenekin elkarreragin dezake, tratamenduaren eragina nahasten duten interakzioak sortaraziz, hots: *hautapena eta historia*, *hautapena eta heldzea*, *hautapena eta testen aplikapena*, *hautapena eta neurketa baliakizuna*, *hautapena eta erregresioa eta hautapena eta hilkortasunaren arteko elkarreragindakoak*, hain zuzen ere.

4.2.9.- Subjektuen hautapen diferentziala eta beste zenbait mehatxuren arteko elkarreragindakoak

Hautapena eta historiaren arteko elkarreragindakoa

Saiakuntza taldeak zoriz osatu ez direnean eta historia lokal desberdina izan duten testuingurutatik datozenen ematen da. Horrelako kasuetan, *historiak* edo *testaurre-testoste bitartean* jazotzen diren gertaera nahiz gora-beheren multzoak eragin desberdina izan dezake talde desberdinetan. Demagun, esate baterako, *talde terapia berri batek* gaixo psikiatriko kronikoen *zenbait arazo psikologikotan* izan dezakeen eragina aztertzeke asmoz ikerketa bat burutzen ari garela. Horretarako, ospitale psikiatriko batean bi gaixo talde borondatez hautatzen ditugu. Batak baldintza esperimentalak jasotzen du eta besteak, berriz, kontrol talde gisa jotatzen du. Interbentzio garaian, saiakuntza taldea, terapia burutzeko bereziki prestaturik dagoen gela batera eramaten dugu. Kasu honetan, emaitzek talde terapia berria eraginkorra dela adieraziko balute ere, ezingo genuke ziurtatu taldeen arteko desberdintasunak tratamenduaren ondorio hutsa izan direnik, subjektuek testaurre-testoste bitartean jasan dituzten gora-behera desberdinen (gela aldaketaren) ondorio ere izan baitaitezke. Ildo honetatik, ikerketa batzuk agerian utzi dute terapia aurrera eramateko testuinguruaren aldaketa hutsak hobekuntzak eragin ditzakeela zenbait gaixo psikiatrikorengan (Gripp eta Magaro, 1971; Higgs, 1970).

Hautapena eta heldzearen arteko elkarreragindakoa

Mehatxu hau agertzen zaigu, ausazkoa ez den esleipen prozedura baten bidez osatutako talde desberdinek heldze maila edo erritmo desberdina aurkezten dutenean. Honen ondorioz, tratamenduaren eragina desberdina izan daiteke talde bakoitzean. Adibide argigarri bat eskaintzearen Kusche eta Greenbergen (1983) ikerketa

aipatuko dugu, zeintzuk haur entzungorrek "on" eta "txar" kontzeptutaz entzumen arazorik gabeko haurrak baino beranduago jabetzen zirela aurkitu zuten. Diferentzia ebolutibo hauek kontuan hartuko ez balira, *"on" eta "txar" kontzeptuen ikaskuntzarako programa* zehatzen baten eraginkortasuna haur entzungor eta normalengan aztertzean interesatutako edozein ikertzaile, programa, entzumen akatsik gabeko haurrengan haur entzungorrengan baino eraginkorragoa dela oker ondoriozta lezake. Alabaina, emaitzak hautapena eta heltzearen arteko elkarre-ragindakoaren ondorio hutsa izan litezke.

Hautapena eta testen aplikapenaren arteko elkarreragindakoa

Interakzio hau errore sorburu izan daiteke, zoriz ez osatzearen ondorioz haseratik elkarren artean desberdinak diren taldeek eragin desberdina jasaten badute edo esperientzia, ezagutza nahiz praktika maila desberdina eskuratzen badute testaurre bati erantzun ondoren.

Hautapena eta neurketa baliakizunaren arteko elkarreragindakoa

Elkarreragindako hau jazotzen da, subjektu talde desberdinak ez direnean aldi berean ebaluatzen edo talde horietako subjektuek neurketa baliabideekin trebezia maila desberdina aurkezten dutenean. Egoera honetan, zorizkoa ez den esleipen irizpide baten arabera eratu diren talde hauek, tarte desberdinez osatutako test batean oso puntuazio desberdinak lor ditzakete. Horrela, *sabai efektua*, talde baten puntuazioak neurketa eskalaren goi mailatan metatzen badira, edo *zoru efektua*, puntuazioak eskalaren behe mailatan geratzen bazaizkigu, ager daiteke.

Hautapena eta erregresio estatistikoaren arteko elkarreragindakoa

Errore sorburu hau agertzen zaigu, zorizkoa ez den esleipen prozedura erabiltzearen ondorioz, saiakuntza taldeek testaurrean muturretako puntuazio desberdinak aurkezten dituztenean eta, hortaz, batazbestekoarekiko erregresioa ere talde desberdinetan modu desberdinean ematen denean.

Hautapena eta hilkortasun esperimentalaren arteko elkarreragindakoa

Ausazko esleipen prozedura erabili ez delako erkaketa taldeak haseratik elkarren artean baliokideak ez badira, saiakuntza baldintza batean bestean baino subjektu gehiago gal daitezke. Tratamenduen efektuak interakzio hau kontuan hartu gabe estimatzen badira, emaitzen interpretazioa alboratua izango da.

Lehenago aipatu den bezala, oraintxe deskribatu ditugun interakzio efektuek barne balidezia zalantzan jartzen dute zoriz osatu ez diren taldeak elkarren artean erkatzen direnean eta menpeko aldagaia denboran zehar neurtzen denean. **Elkarreragindako hauek ager ez daitezzen prozedurarik egokiena zorizkapena erabiltzea da.**

4.2.10.- Kausazko eraginaren norabideari buruzko zalantzagarritasuna

Kausazko eraginaren norabideari buruzko zalantzagarritasuna, zergatia zein den eta ondorioa zein den zehaztu ahal izateko, hau da, X-ek Y sortarazten ote duen edo Y X-en zergatia ote den jakiteko zailtasunari dagokio. *Esperimentalak ez diren zeharkako ikerketetan* garrantzi berezia hartzen duen mehatxua da. Aitzitik, *ikerketa esperimentalak* egiten direnean, edo *luzerako diseinuak* erabiltzen direnean ez da errore sorburutzat hartzen, kasu bietan denbora lehentasuna argi eta garbi ezagutzen baita. Lehenago aipatu dugunez (4.1. atalean), badira luzerako zenbait diseinu, *panel diseinuak* esate baterako, kausazko eraginaren norabidea aztertzeko analisi teknika bereziak (*gurutzatutako atzerapenen koerlazioa panel moduan*) bere barnean ematen dituztenak ere. *Kausazko eraginaren norabidea guztiz finkaturik daukaten koerlazio ikerketetan* ere ez du inolako arazorik sortzen. Esate baterako, Cook eta Campbell-ek (1979) dioten bezala, inguruneke tenperatura jaisterakoan gasolinaren kontsumoa gehituko dela ondorioztatzea guztiz logikoa da, baina gasolina gehiago kontsumitzeak inguruneke tenperaturaren beherakuntza dakarrela esateak, berriz, ez du inolako zentzurik.

4.2.11.- Egoerak eragindako efektu erreaktiboak

Orain arte azaldutako mehatxu guztiak normalean barne balideziaren aurkako mehatxu gisa hartzen badira ere, badaude beste zenbait errore sorburu, zeintzuen sailkapenari buruz desadostasun handia dagoen. Errore sorburu hauek *egoerak eragindako efektu erreaktiboak* bezala ezagutzen dira eta autore batzuk *barne balideziaren aurkako mehatxutzat* hartzen dituzten bitartean, beste batzuen

ustez *konstruktua balideziari* eragiten diote funtsean. Azken hauen artean Judd eta Kenny (1981) bezalako ikertzaileak aipa ditzakegu haien aburuz egoerak eragindako efektu erreaktiboak tratamendu aldagaiaren izaerari eta beronek interes gune den konstruktua teorikoarekin duen harremanari dagozkien batik bat. Judd eta Kennyk proposatutako irizpidearekin bat etorritik, Vallejo (1991a) honako hau dio: zorizkapenaren bidez ezin kontrola daitekeen barne balideziaren aurkako edozein mehatxu, barne balideziaren aurkako mehatxutzat baino gehiago, konstruktua balideziaren aurkako mehatxu gisa har litekeela, eta alderantziz. Guztiek onartzen duten sailkapen irizpiderik ez dagoenez, guk Cook eta Campbell (1979) ereduari helduko diogu, non aipatutako mehatxuak, emaitzak azaltzerakoan aldagai askearekin norgehiagoka ari diren zergatiak bezala eta, beraz, barne balideziaren aurkako mehatxuak bezala hartzen diren. Dena den, aitortu beharra dago azken urteetako zenbait lanetan autore hauek konstruktua balideziaren aurkako mehatxutzat hartzen dituztela egoerak eragindako efektu erreaktiboak (Cook, Campbell eta Peracchio, 1990).

Egoerak eragindako efektu erreaktiboak zorizko esleipen prozedura erabiliz ezin kontrola daitezkeen mehatxuak dira, halatan *saiakuntzen*, *sasisaiakuntzen* eta *koerlazio ikerketen* emaitzak zalantzan jar ditzakete. Normalki, elkarrekintza sozialeko egoeratan sarritan agertzen diren erreaktibitate arazori dagozkien. Cook eta Shadish (1984) arabera, efektu hauen kontrolak tratamenduen ezarpen kalitatean eta, horretaz, saiakuntza efektuen estimazio egokian esanguratsuki eragiten du. Mehatxu hauen artean *tratamenduen hedapena edo imitazioa*, *tratamenduaren pareko ordaina*, *ordaina lortzeko lehiakidetasuna* eta *subjektuen frustrazioa edo desmoralizapena* azpimarra ditzakegu.

Tratamenduen hedapena edo imitazioa

Tratamenduen hedapena edo imitazioa ematen da, kontrol taldeko edo edozein baldintza esperimentaleko subjektuek beste talde edo tratamendu baldintza batzuri buruzko informazioa jasotzen dutenean (e.b. komunikabideen bidez edo beste zenbait taldetako kideekin zuzenean hitz eginez). Demagun, esate baterako, subjektu gizonei agindutako *jan-neurri berri baten* eraginkortasuna aztertu nahi dugula. Horretarako, argaltzeko klinika pribatu batera jotzen duten subjektuen erdiak saiakuntza taldera esleitzen ditugu eta beste erdiari, kontrol talde bezala jokatzeko dutenei alegia, klinika horretan normalean erabiltzen den *jan-neurria* ematen diegu. Interbentzio epean zehar, *jan-neurri berria* proposatu duen medikuari telebistan elkarriketa egiten diote eta bertan *jan-neurriari* buruzko informazio zehatza eskaintzen du. Hori horrela, kontrol taldeko subjektuek telebistako programa ikusten badute, tratamenduaren eragina zalantzan jar daiteke.

Azken batez, beste tratamendu batzuri buruzko informazioa eskuratzeak, subjektuek beren tratamenduaren (edo tratamendu ezaren) aurrean aurkezten dituzten erantzunak edo portaerak aldaraz ditzake. Izan ere, elkarren artean erkatzen direnak ez dira jadanik saiakuntza taldea eta kontrol taldea edo tratamendu desberdinak jasotzen dituzten bi talde, antzeko baldintzen pean dauden subjektuak baizik. Hortaz, tratamenduen hedapenak interbentzioaren eraginari buruz ondorio desegokiak ateratzera eraman gaitzake.

Tratamenduen hedapena errore sorburu izan daiteke *subjektuarterko diseinuekin* lan egiten denean. **Mehatxu hau kontrolpean egon dadin**, saiakuntza subjektuak ikerketa amaitu arte tratamendu experimentalari buruzko inolako iruzkinik ez egiteak daukan garrantziaz jabetu behar dira. Halaber, informazio sorburua edozein delarik ere (komunikabideak, terapia, esperimendatzailea, etab.), subjektuek beste saiakuntza baldintza batzuri buruzko informaziorik jaso ez dezaten saiatu behar gara.

Tratamenduaren pareko ordaina

Mehatxu hau agertzen zaigu, tratamendurik jasotzen ez duten subjektuek pairatu behar izaten dituzten desabantailak gutxitzeko asmoz, kontrol taldeko subjektuei *ordain tratamenduren* bat (dirua, beste zerbitzu batzuk, edozein motatako faboreak, etab.) eskaintzen zaienean. Horrelako egoeran, tratamendu experimentalak hartzen duen taldea beste interbentzio mota bat hartzen duen taldearekin erkatzen da, eta horrek ikerketa objektu den tratamenduak benetan eragindako efektuen hautematea eragotz dezake.

Tratamenduaren pareko ordaina testuinguru aplikatuetan agertzen da normalki, hala nola ikastetxeetan, ospitaletan, lantegietan, etab., eta garrantzi berezia dauka *subjektuarterko diseinuekin* lan egiten dugunean. **Mehatxu hau kontrolatzeko prozedura** egokienetariko bat, *itsu bikoitza teknika* erabiltzea da.

Ordaina lortzeko lehiakidetasuna

Subjektuek kontrol taldeko edo saiakuntza taldeko partaide direla dakitenean, talde bien artean lehiakidetasun soziala sor daiteke. *Ordaina lortzeko lehiakidetasuna*, baldintza desanbantailtsuenean daudela sumatzen badute,

itxarondako desberdintasunak gutxitzeko edo alderantzikatzeko asmoz kontrol taldeko subjektuek ager dezaketen lehiakidetasun maila garaiari dagokio.

Ordaina lortzeko lehiakidetasuna, sarritan, berezko unitateak (ikastetxeak, lantegiak, etab.) tratamenduetara esleitzean direnean agertzen da edo kontrol taldeko subjektuek tratamendu esperimentalaren eraginkortasunak beren lanpostua, maila sozioekonomikoa, karrera profesionala, etab. zalantzan jar dezakeela hautematen dutenean. Errore sorburu izan daiteke *subjektuarteko diseinuekin* lan egiten denean eta *itsu bakuna teknikaren* bidez **kontrola dezakegu**.

Subjektuen frustrazioa edo desmoralizapena

Baldintza desanbantailtsuenaren partaide direla sumatzeak kontrol taldeko subjektuengan eragiten duen frustrazioa, desmoralizapena edo gaitzondoaren ondorioz, subjektu hauen errendimenduak jasan dezakeen beherakuntzari dagokio. Honela, testuinguru industrialean, esate baterako, tratamendu ezatseginena jasotzen duten langileek beren produkzio maila eta, beraz, lantegiaren etekinak gutxituz erreakziona dezakete, eta horrek saiakuntza tratamenduak benetan eragindako efektuen hautematea eragotziko luke.

Errore sorburu hau ordaina lortzeko lehiakidetasuna sortarazten duten antzeko egoeratan agertzen da. Azken hau bezala, *subjektuarteko diseinuekin* lan egiterakoan hartu behar dugu kontuan eta *itsu bakuna teknika* erabiliz **kontrola daiteke**.

4.1. Taulan barne balideziaren aurkako mehatxu nagusienak laburki deskribatzen dira eta berauek kontrolatzeko erabil daitezkeen zenbait prozedura aipatzen dira. Hurrengo atalean, barne balidezia maila egokia lortu ahal izateko ikerketa eremuan gehienbat erabiltzen diren kontrol metodoak zertan datzaten azaltzen dugu.

4.1. Taula.- Barne balideziaren aurkako mehatxu nagusienak eta berauek kontrolatu ahal izateko prozedurak

MEHATXUA	DEFINIZIOA	KONTROL PROZEDURAK EDO TEKNIKAK
1.- Historia	- <i>Barne historia</i> edo <i>historia proaktiboa</i>: Subjektuak elkarren artean bereizten dituzten ikasitako faktoreei edo faktore genetikoei (ezaugarri fisikoak, buru gaitasunak, jarrerak...) dagokie - <i>Kanpo historia</i> edo <i>historia erretroaktiboa</i>: Tratamenduarekin zerikusirik izan gabe, esperimentuan bertan edo egoera hortatik kanpo jazo daitezkeen eta M.A.-ri eragiten dioten gora-beherak edo gertaerak	- Zorizkapena - Kontrol taldea - Konstantzia teknika
2.- Heltzea	Gizabanakoarengan denbora igarotzearen ondorioz agertzen diren eta bere erantzunetan eragiten duten aldaketa biologiko eta psikologikoak	- Kontrol taldea - Konstantzia teknika
3.- Testen aplikazioa	Tratamenduaren eraginatik kanpo, testaurrearen aplikapenak subjektuek testostean emandako erantzunetan eragin dezake	- Kontrol taldea
4.- Neurketa baliakizuna	Neurketa baliabideetan edo behatzaile edo kalifikatzaileengan denboran zehar jazotako aldaketek, erroreak sor ditzakete subjektuen erantzunak neurtzerakoan	- Estandarizazioa/Automatizazioa - Kontrabalantzeoa edo ekiponderazioa - Itsu bikoitza teknika - Behatzaile edo kalifikatzaileen trebakuntza - Fidagarritasun egiaztapen anizkoitzak
5.- Tratamendu anizkoitzen arteko elkar eragindakoa	- <i>Hurrenkera efektuak</i>: Ikaskuntza edo nekea bezalako faktoreek, saiakuntzaren barnean orden desberdinean aurkeztutako tratamenduen aurrean subjektuek ematen dituzten erantzunetan modu desberdinean eragiten dute - <i>Hondakin efektuak</i> edo "<i>carry-over</i>" efektuak: Tratamenduaren efektuek bizirik diraute tratamendu epea amaitu ondoren eta subjektuek hurrengo tratamenduan emandako erantzunei eragiten diete	- Kontrabalantzeo edo ekiponderazio teknikak - <i>Cross-over</i> edo <i>ordezko diseinuak</i> eta <i>errepikapen neurketako karratu latinozko diseinuak</i> erabili
6.- Erregresio estatistikoa	Neurketa sistema berbera berriz ere aplikatzen denean, muturretako puntuazioek banaketaren batzbestekorantz hurbiltzeko aurkeztzen duten joera	- Taldeak ausaz osatu - Puntuazio egonkorrak lortzearen, interbentzioa aplikatu aurretik zenbait neurketa burutu

4.1. Taula (jarraipena).- Barne balideziaren aurkako mehatxu nagusienak eta berauek kontrolatu ahal izateko prozedurak

MEHATXUA	DEFINIZIOA	KONTROL PROZEDURAK EDO TEKNIKAK
7.- Subjektuen hautapen diferentziala	Hautapen edo esleipen prozedura desgokia erabiltzearen ondorioz, tratamenduak ezarri aurretik talde desberdinen arteko baliokidetasun eza	- Zorizkapena - Beste zenbait kontrol teknika esperimental (e.b. blokeoa, subjektu baliokideetan oinarritutako parekatzea, etab.) edo estatistikoak (e.b. kobariantza analisisa)
8.- Hilkortasun esperimental edo subjektuen galera selektiboa	Saiakuntza talde desberdinetan subjektu kopuru desberdina galtzen da	- Ikerketa, hilkortasun esperimentalaren jazoetako probabilitatea ahalik eta txikiena izan dadin moduan antolatu
9.- Subjektuen hautapen diferentziala eta beste zenbait mehatxuren arteko elkarrengatutakoak	Taldeak osatzeko ez-zorizko esleipen prozedura erabiltzearen ondorioz, zenbait mehatxuk modu desberdinean eragiten dute talde bakoitzean	- Zorizkapena
10.- Kausazko eraginaren norabideari buruzko zalantzaragarratasuna	Zergatia zein den eta ondorioa zein den zehaztasunez finkatzeko zailtasuna	- Luzerako ikerketak egin - Zeharkako ikerketak burutuz gero, estrategia esperimentalari berez dagozkion kontrol prozedurak erabili
11.- Tratamenduen hedapena edo imitazioa	Subjektuen artean informazioa elkarrengatutako ondorioz, tratamendu (edo kontrol) talde desberdinetako subjektuek antzeko tratamendu baldintzak jasotzen dituzte	- Subjektuek tratamendu esperimentalari buruzko informazioa zabaltzea ekidin - Subjektuek beste saiakuntza baldintza batzuri buruzko informazioa sorburu desberdinetatik ez jasotzea
12.- Tratamenduaren pareko ordaina	Kontrol taldeko subjektuek, efektu onuragarriak dakartzkien eta tratamendu gabeziaren ondoriozko desabantailak sahisteko xedea jarraitzen duen ordain tratamendua jasotzen dute	- Itsu bikoitza teknika
13.- Ordaina lortzeko lehiakidetasuna	Baldintza desabantailtsuenean daudela sumatuz gero, saiakuntza taldearekin erkatuz itxaron daitekeen diferentzia gutxitzeko edo horren kontra jotzeko asmoz, kontrol taldeko subjektuek lehiakortasun maila altua agertzen dute	- Itsu bakuna teknika
14.- Subjektuen frustrazioa edo desmoralizazioa	Baldintza desabantailtsuenean daudela sumatuz gero, kontrol taldeko subjektuak frustraturik, desmoralizaturik edo sumindurik sentitzen dira eta errendimendua gutxitzen da	- Itsu bakuna teknika

4.3.1.- Ezabapena

Kontrol teknika hau *aldagai kontaminatzailea* ikerketa egoeratik egoztean datza. Esate baterako, animaliekin lan egiten denean, ohikoena da animaliak zaratatik isolatzen dituzten kaxak erabiliz "*zarata*" *aldagai nahastatzailea* egozte. Halere, psikologian agertzen zaizkigun *aldagai arrotz* gehienak nekez elimina daitezkeela onartu beharrean gaude.

4.3.2.- Konstantzia

Konstantzia, psikologian sarritan erabiltzen den prozedura da eta *aldagai arrotzaren* maila bakarra erabiltzean edo aldagai hori konstante (aldatzeko posibilitaterik gabe) mantentzean datza. Zenbait autorek, *estandarizazio* izena ere ematen diote teknika honi (e.b. Arnau, 1990; Burns eta Dobson, 1981) eta beste batzuk, *uniformetasun* nahiz *konstantzia* hitzak adiera berberaz erabiltzen dituzte prozedura hau aipatzerakoan (e.b. Pedhazur eta Pedhazur Schmelkin, 1991). Subjektu guztiei tratamendua ordu berean ezartzea (*egunaren ordua aldagai nahastatzailearen* kontrola) edo argibideak magnetofoiaren bidez aurkeztea (*argibideak aldagaiaren* kontrola), kontrol teknika honen adibideak dira. Teknika hau *subjektuzko aldagai arrotzak* kontrolatzeko erabiliz gero, Christensenek (1988) *konstantzian oinarritutako parekatzea* deritzona, konstantziak ikerketan bertan parte hartzen duten subjektu motatatik kanpo emaitzak ezin orokortzearen desabantaila aurkezten du. Atok (1991) **itsu bakuna** eta **itsu bikoitza** teknikak konstantzia estrategiaren barnean sartzen ditu. Badago arrazoirik prozedura hauek *barne balideziari* zein *konstruktuko balideziari* lotzeko, subjektuen eta esperimintatzaileen hautemate nahiz espektatibetatik eratorritako erroreak kontrolatzeko aukera ematen baitigute. Alabaina, aurreko atalean teknika biak barne balideziaren aurkako mehatxuei aurre egiteko prozedurak bezala aurkeztu direnez gero, egokiago deritzogu oraingoan Atoren (1991) irizpideari jarraitzea. **Itsu bakuna teknikan**, esperimintatzaileak saiakuntza hipotesia ezagutzen du, baina "itsua" da subjektu bakoitzari esleitutako tratamendu baldintzari dagokionez. Ondorioz, ezin dezake talde desberdinekin modu desberdinean inkontzienteki jokatu. Aurrekotik erator daitekeen bezala, subjektuak ez daki saiakuntza taldearen edo kontrol taldearen partaide ote den. **Itsu bikoitza teknikan**, berriz, ez subjektuek ez eta esperimintatzaileak ere, enpirikoki egiaztatu nahi den hipotesiari buruzko inolako ezagutzarik ez daukate.

4.3.3.- Orekatzea subjektuarteko diseinuetan

Aldagai arrotzen eragina ezin daitekeenean egotzi edo konstante mantendu, bere efektuak tratamendu desberdinen artean modu orekatuan bana ditzakegu *orekatzea* ("*balancing*") izeneko teknika erabiliz. Kontrolatu nahi den aldagai

kontaminatzaile motaren arabera **orekatzea lortzeko metodo desberdinak** daude. Anderson eta Borkowski (1978), Christensen (1985), Conrad eta Maul (1981), McGuigan (1983) eta Myersi (1980) jarraituz, *subjektuarterko diseinuen* kasuan Atok (1991) *zorizkapena*, *parekatzea* eta *doikuntza estatistikoa* azpimarratzen ditu eta *subjektubarneko diseinuetarako*, berriz, *kontrabalantzeoa* edo *ekiponderazioa* gomendatzen du.

4.3.3.1.- Zorizkapena

Zorizkapena kontrol tekniken arteko garrantzitsuena da, errore sorburu ezagunek nahiz ezezagunek emaitzetan alborapen sistematikorik ez dutela eragingo probabilitate mailan ziurtatzeko aukera ematen baitigu.

Aldagai arrotzak erabat kontrolatu nahi badira, jatorrizko populaziotik subjektu lagina *zoriz hautatu* behar da. Ondoren, *subjektuak zoriz esleitu* behar dira saiakuntza talde desberdinetara. Azkenik, *tratamenduak* talde desberdinen artean *banatzerakoan* ere *zoriari jarraitu* behar diogu.

Zorizko hautapenak populazioko subjektu guztiek laginean parte hartzeko probabilitate berbera dutela ziurtatzen du eta, beraz, jatorrizko populazioaren *lagin errepresentatiboak* lortzeko aukera ematen digu. Halere, *aldagai nahastatzaileen* kontrola benetan ziurtatzen duena ez da *zorizko hautapena*, *zorizko esleipena* baizik, zein ikerketa esperimentalaren berezko ezaugarri gisa har daitekeen. *Zorizko esleipena*, probabilitatearen arauari jarraituz saiakuntza taldeen arteko baliokidetasuna lortzeko prozedurarik zehatzena da. Bai subjektuen hautapenak eta baita esleipenak ere zoriaren irizpideari jarraitzen diotenean, aldagai arrotzen banaketa eta eragina saiakuntza talde guztietan berbera izango dela suposa daiteke.

Zorizkapenak kontrol teknika gisa duen balioa eztaba daezina den arren, bere eraginkortasuna zalantzan jar daiteke lagin txikiekin lan egiten dugunean. Horregatik, azken kasu honetan hobe izaten da *parekatzea* deritzon prozedura erabiltzea. Azkenik, zera azpimarratu nahiko genuke: *zorizkapenak* ez duela *aldagai arrotzek eragindako aldakortasuna* egozten, berau saiakuntza baldintza desberdinen artean modu orekatuan banatzen baizik. Gainera, *zorizkapenaren* bidez kontrolatzen den aldagai kontaminatzailearen ondoriozko aldakortasuna, *sistematikoa ez den bariantza* bezala edo *ausazko errore* bezala hartzen da.

4.3.3.2.- Parekatzea

Diseinuaren *zorrozatasuna* errore bariantza gutxituz gehi daiteke eta, ikuspegi hortatik, *parekatzea zorizkapena* baino prozedura egokiagoa da. Prozedura honek *bigarren mailako bariantza sistematikozko* sorburu bat edo gehiago

kontrolatzeko eta *parekatze aldagaia* deritzon *kontrolatutako bariantza sistematikozko* osagai berria sortarazteko aukera ematen digu. Bariantza sorburu berri hau analisi estatistikotik egotzia izan daiteke. Nolanahi ere, Atok (1991) dioen bezala, zorizkapenaren eraginkortasuna ziurtatzen ez duten lagin tamaina txikiekin lan egiten denean eta *parekatze aldagaiaren* eta *menpeko aldagaiaren* arteko koerlazioa $r=0,50$ ekoa baino handiagoa denean soilik erabili behar da parekatzea (Cochran, 1983; Feldt, 1958; Kerlinger, 1986). Teknika honen erabilpena nahikoa konplexua gertatzen da lagin handiak, tratamendu anitz edo kontrolatu beharreko aldagai asko ditugunean. Bestalde, garbi utzi nahi dugu parekatzea ez dela zorizkapenaren ordezkotik gisa hartu behar. Izan ere, ahal den guztietan zorizkapena erabili behar da, ez baita batere erraza aldagai gehienak modu zehatzean parekatzea eta, gainera, ikerketaren emaitzetan eragin dezaketen aldagai garrantzitsu guztiak identifikatzea eta parekatzea ezinezkoa gertatzen da.

Parekatzea modu desberdinetan egin daiteke. Jarraian, maizen erabiltzen diren parekatze motak aurkezten dizkizuegu.

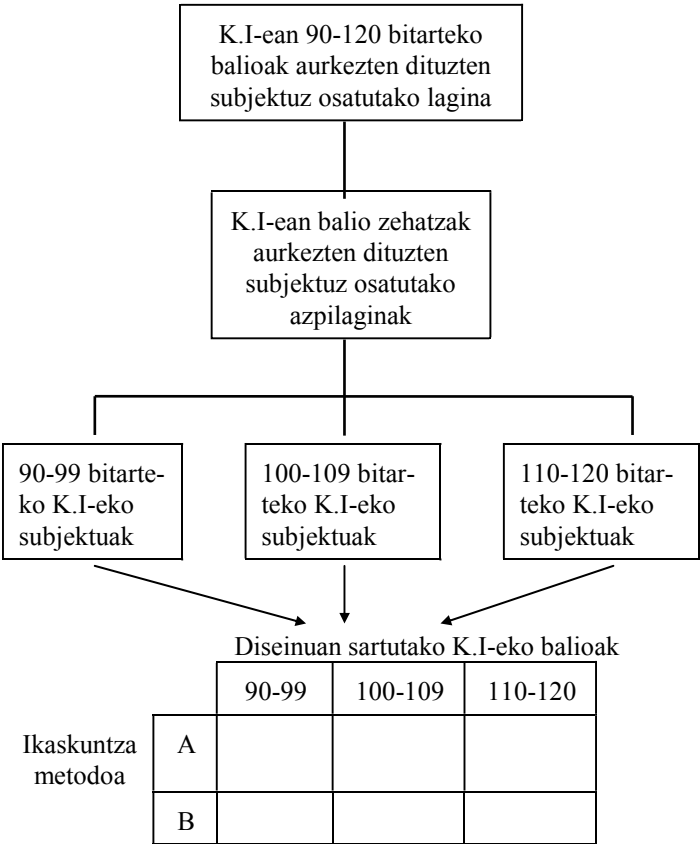
Aldagai arrotza diseinuan sartzean oinarritutako parekatzea

Parekatze mota hau *aldagai arrotza* beste aldagai bat gehiago izango balitz bezala diseinuan sartzean datza, honela berorrek menpeko aldagaiean izan ditzakeen efektuak azter baitaitezke. Azken batean, *lehen mailako bariantza sistematikozko* beste sorburu bat eratuz diseinua *faktorizatzean* datza. Horren ondorioz, *bigarren mailako bariantza sistematikozko* sorburu bat kontrolatzen da eta *errore bariantza* gutxitzen da, azken honi kontrolatutako aldagaiak eragindako bariantza proportzioa kentzen baitiogu. Horrela, diseinuaren *zorrotasuna* areagotzen da. Demagun, esate baterako, *subjektuek matematika ikasgaiean aurkeztutako errendimenduan bi ikaskuntza metodok* izan dezaketen eragina aztertzen ari garela eta *adimena* aldagaiaren efektuak kontrolatu nahi ditugula. Horretarako, koefiziente intelektualeko hainbat balio hauta ditzakegu (e.b. 90-99, 100-109 eta 110-120) eta aldagai aske berri baten balioak bezala har ditzakegu. Estrategia honek adimenaren eragina estimatzeko eta eragin hori *errore bariantzatik* egozteko aukera ematen digu (ikus 4.3. Irudia).

Atok (1991) **geruzatzean oinarritutako parekatzea** bezala aurkezten digu teknika hau. **Aldagai arrotz nahastatzailearen sistematizapena** eta, eskuharkiago, **blokeo teknika** izenez ere aski ezaguna da. Prozedura hau guztiz gomendagarria da aldagai arrotzaren maila desberdinek edo aldagai nahastatzailearen eta diseinuan parte hartzen duten beste zenbait aldagairen arteko elkarreragindakoak menpeko aldagaiean daukaten eragina aztertu nahi bada. Egoki erabilia izan dadin, *aldagai nahastatzailea kategorikoa* izan behar da.

Kontrol akoplatuan oinarritutako parekatzea

Teknika honen helburua gertaeraren eta erantzunaren arteko denbora erlazioak eragin dezakeen errorea kontrolatzean datza. Adibidez, ikerketa bat burutzen ari garelarik, saiakuntza subjektuek shock elektrikoak jasan behar badituzte, *shockak emateko erabilitako denbora sekuentzia aldagai nahastatzailea*, ikerketako partaide guztiek shockak jasaterakoan shock batetik bestera denbora tarte berbera izan dezaten ziurtatuz kontrola daiteke. Kontrol prozedura hau, gertakizunen denbora banaketak eragin dezakeen errorea saihestu ahal izateko oso egokia da.



4.3. Irudia.- Aldagai arrotza diseinuan txertatzean oinarritutako parekatzearen adibidea (Christensenetik moldatua, 1988)

Subjektu baliokideetan oinarritutako parekatzea

Subjektu baliokideetan oinarritutako parekatzea, subjektuak kontrolatu nahi den aldagaiaren (edo aldagaien) arabera parekatzean datza, honela talde baliokideak eratzen direlarik. *Aldagai arrotza diseinuan sartzean oinarritutako parekatzean ez*

bezala, estrategia hau erabiltzean ez dugu *aldagai nahastatzailea* diseinuaren barnean sartzen. Normalki, parekatze mota hau bi modutan lortzen da:

- 1.- **Zehaztasunean oinarritutako kontrola.** Kasu honetan, ikertzaileak subjektu bakoitza *aldagai kontaminatzaile* bakoitzaren maila berberari dagokion (aldagaia kategorikoa bada) edo aldagai horretan oso antzeko balioa aurkezten duen (aldagaia kuantitatiboa bada) bigarren (hirugarren, etab.) subjektu batekin parekatzen du. Ondoren, bikote (hirukote, etab.) bakoitzeko kide bana zoriz esleitzen du tratamendu baldintza bakoitzera. Teknika honen bidez, talde desberdinetako kideak, emaitzetan eragin dezaketen *aldagai nahastatzaileei* dagokienez baliokideak direla ziurta dezakegu; horregatik, parekatze honek esanguratsuki gehitzen du diseinuaren *zorrotasuna*. Dena den, zorrotasuna gehi dadin honako baldintza hau bete behar da: *parekatze aldagaien* eta *menpeko aldagaiaren* arteko koerlazioa handia izatea, alegia.
- 2.- **Maiztasunen banaketan oinarritutako kontrola.** *Zehaztasunean oinarritutako kontrola* oso egokia da zorrotasuna areagotzeko, baina subjektu guztiak parekatzerik ez dagoenez gero, ezinbestean datorkigu zenbaitzuren galera. Maiztasunen banaketan oinarritutako kontrola arazo hori gainditzen saiatzen da eta horretarako, parekatzea subjektuz subjektu egin ordez, subjektu taldeen artean egiten du. Horrela, *aldagai nahastatzaileetan* balio berbereko *banaketa parametroak* (batazbestekoa, desbidazio tipikoa, asimetria, etab.) aurkezten dituzten taldeak parekatzen dira. Dena den, *aldagai arrotz* bat baino gehiago kontrolatzen denean estrategia hau arriskutsua gerta daiteke, talde desberdinetako subjektuak elkarren artean oso desberdinak izanik ere, talde horietako *aldagaien banaketa parametroak* baliokideak izan baitaitezke.

4.3.3.3.- Doikuntza estatistikoa

Aldagai arrotzaren eragina saiakuntza kontrol prozeduren bidez egotzi ezin denean, analisi estatistikoan zenbait doikuntza eginez kontrola daiteke. **Doikuntza estatistikozko tekniken** helbururik nagusia **errore bariantza** gutxitu eta diseinuaren **zorrotasuna** gehitzean datza.

Atori (1991) jarraituz, aldagai nahastatzailea *kategorikoa* baldin bada eta tratamendu efektuak tratamendutik tratamendura egonkorrak badira, *doikuntza* motaren bat aplika daiteke, esate baterako talde batazbestekoren kasurako *batazbestekoren arteko diferentzia doitua*, edo proportzioen kasurako "*logit*" delako eraldakuntza.

Aldagai arrotza *kuantitatiboa* bada, doikuntza prozedurarik egokiena *kobariantza analisisa* (ACOVAR) da. Huitemak (1980) dioenez, *kobariantza analisisa* *erregresio analisisit* abiatuz burutzen da. Lehenik, *menpeko aldagaiaren* erregresio analisisa *tratamendu aldagai(ar)ekiko* eta *aldagai kobariante(ar)ekiko* egiten da.

Honela, tratamendu aldagaiak -edo aldagaiek- eta aldagai kobarianteak -bat edo gehiago izan daitezkeelarik- azaldutako menpeko aldagai(ar)en bariantza adierazten duen *determinazio-koefiziente anizkoitza* kalkula daiteke. Jarraian, kobarianteak azaldutako bariantzari dagokion *determinazio-koefizientea* kalkulatzearren, *menpeko aldagai(ar)en eta kobariantearen arteko* erregresio analisia egiten da. Erregresio analisi horietan kalkulaturako determinazio-koefizienteen arteko kendurak kobariantearen eraginik gabeko menpeko aldagaiaren bariantza azaldua adierazten du. Horrela kontrolatzen da kobariantearen eragina hain zuzen ere. Bide horretatik, azkenik, *kobariante* edo *aldagai nahastatzaile posiblearen* eraginatik kanpo dauden puntuazioak oinarritzat hartuz ohiko bariantza analisia egiten da.

Kobariantza analisiak bariantza analisiak baino indar handiagoa dauka. Izan ere, tratamenduak ezarri aurretik taldeen artean izan daitezkeen diferentzietatik eratorritako alborapena murrizteko aukera baitakar. Alabaina, Huitemak (1980) dioenari lotuz, sasisaiakuntza eremuetan, ACOVAR-a bere balioa zalantzan jar dezaketen zenbait arazoren menpe dago. Zehazki, *kobarianteak eta menpeko aldagaiak* dimentsio berbera neurtzen ez dutenean, edo *kobarianteak* neurketa errore nabarmenak baditu, edota fidagarritasun eskasa duenean alborapena izaten dugu. Hau da, batzbesteko doituaren arteko diferentzien eta, beraz, kobariantza analisiaren F testaren estimazioan alborapena agertzen zaigu.

Arazo hauek direla eta, Huitemak ordezeko analisi teknikak lantzearen beharra aldarrikatzen du. Horrela, behar horri erantzun nahian, Arnauk (1995) sasisaiakuntza diseinuaren eremuan kobariantza teknika ordezkia dezaketen analisirako prozeduren laburpen bikaina egin du. Prozedura horien artean *blokeo edo parekatze teknikan oinarritutako bariantza analisia* (testuliburu honetan *aldagai arrotza diseinuan txertatzean oinarritutako parekatzea* delako atalean aztertu dugu), *diferentzia, aldakuntza edo irabazte puntuazioekin burututako bariantza analisia, estandarizatutako diferentzia puntuazioekin burututako bariantza analisia eta benetako puntuazioetan oinarritutako kobariantza analisia* aipatzen ditu.

Oro har, *blokeo edo parekatze teknikan oinarritutako bariantza analisiak* ACOVARekiko dakarren abantailarik nagusiena hurrengo hau da: teknika horrek menpeko aldagaiaren puntuazioen aldagai kobariantearen puntuazioekiko erregresio zuzenak hartzen duen itxurari buruzko inolako aurreiritzirik ez du eskatzen. Lehenengo estrategia horren ildotik, *diferentzia, aldakuntza edo irabazte puntuazioekin burututako bariantza analisiak* ACOVARek dakarren arazo ugari gainditzen du. Hala ere, jakiniko zenbait baldintzapean desabantaila batzuk bide, bere erabilera ez da gomendagarria. Lehenik, eta aurreko ordezeko analisi teknikak ez bezala, tratamenduaren -edo tratamenduen- eta tratamendu aurreko neurrien artean ager daitezkeen interakzioa kalkulatzeko aukerarik ez dakar. Are gehiago, tratamendu aurreko neurrien populazio batzbestekoak elkarren baliokideak ez direnean, prozedura hau desegokia da. Horrek badu azalpena, ikus dezagun: testaurrean balio handiak dituzten subjektuek balio txikiak dauzkatenean baino *heltze* maila handiagoa agertzen dute; ondorioz, diferentzia puntuazioen batzbestekoak oso desberdinak izan daitezke tratamendu aurreko neurri desberdinetatik abiatzen diren taldeetan.

Gertakizun honekin batera, denbora igarotzeak eta *heltze* edo garapen prozesuek taldearteko eta taldebarneko aldakortasuna areagotzen dute. *Sakabanatze zabaleko gertakizun*¹ hori dela medio, tratamenduen efektuaren estimazioan alborapena sortzen da. Arazo honi aurre egiteko, ziurtatu behar dugu tratamendu aurreko eta osteko neurriek bariantza baliokideak izan ditzaten. Horrela ez balitz, bariantzen baliokidetasuna lortzeko eraldaketaren bat burutzea ezinbestekoa da. Xede hori lortzeko, hain zuzen ere, *estandarizatutako diferentzia puntuazioekin burututako bariantza analisia* erabiltzen da. Azkenik, *benetako puntuazioetan oinarritutako kobariantza analisia* tratamendu aurreko aldagai(ar)en (kobariantearen) neurketa erroreetatik eratorritako arazoak sahiesten ditu. Azken buru, aldagai horren puntuazioak benetako puntuazioetan edo fidagarritasunaren arabera zuzendutako puntuazioetan eraldatzen ditu.

4.3.4.- Kontrabalantzeoa edo ekiponderazioa

Kontrabalantzeoa edo *ekiponderazioa* taldebarneko diseinuei dagozkien nahasketa sorburuak, hots, *hurrenkera efektuak* eta *hondakin* ("carry-over") *efektuak* kontrolatzeko erabiltzen den teknika da. Hiru ekiponderazio mota nagusi bereiz daitezke: *subjektubarneko ekiponderazioa*, *taldebarneko ekiponderazioa* eta *ausazko ekiponderazioa*, hain zuzen ere.

Subjektubarneko ekiponderazioa (ABBA teknika)

Subjektubarneko ekiponderazioan *hurrenkera* eta *hondakin efektuak* kontrolatzen dira subjektu bakoitzari tratamenduak *hurrenkera* zehatz batean, lehenik, eta alderantzizko *hurrenkera*an, ondoren, jasoeraziz. Saiakuntza bi tratamenduz osaturik badago, subjektu bakoitzak *AB* baldintzak jasotzen ditu lehenik eta *BA* baldintzak ondoren, beraz *ABBA* sekuentziari aurre egiten dio. Subjektu bakoitzaren kasuan, bi *A* baldintzen pean eta bi *B* baldintzen pean lortutako emaitzak elkarrekin konbinatzen dira, eta tratamenduen efektuak proba estatistiko egokiaren bidez erkatzen dira.

ABBA metodoa ondoko aurretikoan oinarritzen da: *hurrenkera* eta *hondakin efektuak* sekuentziaren barneko elkarren segidako posizio bakoitzerako *lerro izaeradunak* edo *egonkorrak* dira, hau da, posizioak edozein direlarik, posizio batetik bestera igarotzeak alborapen edo errore kopuru berbera eragiten du beti. Aurretiko hau betetzen ez bada, teknika honek ez ditu *hurrenkera* eta *hondakin efektuak* kontrolatzen. Normalean, bi baino tratamendu baldintza gehiagorekin erabiltzen ez den prozedura da, zeren sekuentziaren luzera tratamendu kopurua handitu ahala areagotzen baita eta, beraz, subjektuarentzat oso neketsua gerta daiteke.

¹ Gazteleraz *patrón de dispersión en abanico* esaten zaio (Arnau, 1995). Euskaratzerakoan nahiago izan dugu kontzeptuaren adierari lotzea erabilitako etiketa itzultzea baino.

Taldebarneko ekiponderazioa

Taldebarneko ekiponderazio teknika tratamendu sekuentzia desberdinak subjektu azpitalde desberdinen artean banatzean datza, honela *hurrenkera* eta *hondakin efektuak* talde mailan, eta ez subjektu mailan, kontrolatzen dira. Zentzu honetan, bi tratamendu baino gehiago ditugunean aurreko teknika baino eraginkorragoa da. Taldebarneko ekiponderazioa bi prozedura desberdinen bidez lor daiteke:

- 1.- **Taldebarneko ekiponderazio osoa.** Estrategia hau aplikatu ahal izateko, tratamendu konbinaketa desberdinetatik eratorritako sekuentzia posible guztiak eratu behar dira eta sekuentzia horietako bakoitza subjektu azpitalde desberdin bati eman behar zaio. *Lerro tankerako edo egonkorak diren hurrenkera* eta *hondakin efektuak* kontrolatzeko teknika bikaina izan arren, ez da gehiegi erabiltzen. Bere erabilpen eskasaren arrazoia honako hau da: sekuentzia kopurua faktorialki areagotzen dela tratamendu kopurua aritmetikoki handitzen den neurrian (lau tratamendurekin sekuentzia posibleen kopurua $4!=24$ da; bost tratamendurekin $5!=120$ ra igotzen da, etab.). Ondorioz, beharrezko subjektu kopurua gehiegizkoa izatera hel daiteke.
- 2.- **Taldebarneko ekiponderazio osagabea.** Teknika honetan ez da beharrezkoa tratamenduen arteko konbinaketatik eratorritako sekuentzia guztiak erabiltzea; aldiz, sekuentzia kopuruak tratamendu kopuruaren berdina izan behar du. Honela, sekuentzia bakoitza subjektu azpitalde desberdin bati eman behar zaio *karratu latino*a deritzon antolakuntza bereziari jarraituz. Teknika hau erabiltzeak bi *baldintza* betetzea eskatzen du: (a) Sekuentzia bakoitzean, tratamendu baldintza bakoitza hurrenkera posizio bakoitzean aldi kopuru berbera agertu behar da, eta (b) tratamendu baldintza bakoitza beste edozein baldintzaren aurretik eta atzetik aldi kopuru berbera agertu behar da. Bi baldintza hauek bete daitezten *karratu latinoaren* lehenengo errenkada (sekuentzia) garatzeko *formula tipiko* bat aplikatu behar da, hots: $1, 2, n, 3, (n-1), 4, (n-2), 5, \dots$, eta honela elkarren segidan, non n tratamendu kopuruari dagokion (esate baterako, saiakuntza 6 tratamendu baldintzaz osaturik balego, lehenengo sekuentzia ondokoa litzateke: $A B F C E D$). Behin lehen errenkada (sekuentzia) osatu ondoren, ondorengoak lortzeko aurreko errenkadako elementu bakoitzetik berorri jarraitzen dion hurrengo elementura pasatu behar gara, n errenkadak osatu arte. Dena den, formula hau ez da baliagarria tratamendu kopurua bakoitia denean, kasu horretan goran aipatutako bigarren baldintza ez baita betetzen. Arazoa gainditzeko, formula tipikotik abiatuz lortutako sekuentzia kopurua bikoiztu behar dugu. Horretarako, beste hainbat sekuentzia eratu beharko ditugu aurreko bakoitza "ispiluan proiektatuz" ("karratu latino bikoitz"eko antolakuntza). Taldebarneko ekiponderazio osoak bezala, estrategia honek *lerro tankerako sekuentzia efektuak* soilik kontrolatzen ditu.

Ausazko ekiponderazioa

Ausazko ekiponderazioa, subjektu bakoitzarentzat sekuentzia bana zoriz hautatzean datza. Adibidez, saiakuntzan 20 subjektu baldin badaude, 20 sekuentzia zoriz hautatzen dira eta sekuentzia bana zoriz esleitzen zaio subjektu bakoitzari. Prozedura hau, zoriak *hurrenkera* eta *hondakin efektuak* egoki kontrolatzen dituenaren ustetik abiatzen da. Alabaina, zoriak errorea tratamendu desberdinen artean modu orekatuan banatuko du baldin eta subjektu kopuru nahikoa erabiltzen bada. Izan ere, teknika honek alborapenaren kontrola ziurta dezan, gutxienez 50 laginketa unitate erabili behar direla estimatzen da (Zimmy, 1961).

5.- Konstruktua balidezia

- 5.1.- Konstruktua balideziaren definizioa eta ezaugarri nagusiak
- 5.2.- Konstruktua balideziaren aurkako mehatxuak
 - 5.2.1.- Azalpen operazional desegokia
 - 5.2.2.- Operazionalizazio bakar baten erabilerak eragindako erroreak
 - 5.2.3.- Operazionalizazio metodo bakar baten erabilerak eragindako erroreak
 - 5.2.4.- Konstruktua eta konstruktua mailen arteko nahasketa
 - 5.2.5.- Tratamendu anizkoitzen arteko elkarreragindakoa
 - 5.2.6.- Tratamenduaren eta aldeez aurretik aplikatutako testen arteko elkarreragindakoa
 - 5.2.7.- Konstruktua batetik bestera orokortzeko aukera eskasa
 - 5.2.8.- Egoerarekin eta elkarrekintza sozialarekin zerikusia duten faktoreetatik eratorritako mehatxuak
 - 5.2.8.1.- Hipotesien sorkuntza testuinguruko ezaugarrietan oinarrituz
 - 5.2.8.2.- Saiakuntza subjektuek hartzen dituzten errol desberdinak
 - 5.2.8.3.- Esperimentatzaileak eragindako erroreak
 - 5.2.8.4.- Subjektuak eta esperimentatzaileak eragindako errorearen kontrola

5.- KONSTRUKTU BALIDEZIA

5.1.- Konstruktua balideziaren definizioa eta ezaugarri nagusiak

Konstruktua balidezia bi aldagaien arteko erlazio kausalaren interpretazio teorikoari dagokio. Cook eta Campbelli (1976, 1979) jarraituz, menpeko aldagaiaren neurriaren eta aldagai askearen manipulazioaren eta neurtzen edo manipulatzeko ari garen konstruktua teorikoaren arteko egokitzazio maila bezala uler daiteke. *Konstruktua balidezia, fenomeno behagarrietatik (indikatzaileretatik) abiatuz behagarriak ez diren fenomenoei (konstruktua teorikoei) buruz egiten diren inferentzien balideziari dagokio*. Honelako inferentziak egitea ez da batere erraza, konstruktua bera indikatzaile anitzen bidez errepresenta baitaiteke eta, era berean, indikatzaile berak konstruktua desberdinak adieraz ditzake. Bestalde, indikatzaileek adiera desberdinak har ditzakete leku, kultura, azpikultura, denbora epe, etab. desberdinen arabera. Ondorioz, balidezia mota honen arazo nagusienetarikoa bat, itxuraz *zergati -eragile- edo ondorio -eragindako- konstruktua* zehatz bat adierazten duen indikatzailea konstruktua bat edo gehiagoren arabera interpretatua izateko posibilitatean datza (Cook eta Campbell, 1979).

Bi konstruktua balidezia mota bereiz daitezke: *ondorioaren edo eragindakoaren konstruktua balidezia* eta *zergatiaren edo eragilearen konstruktua balidezia*, hain zuzen ere. *Ondorioaren konstruktua balideziaz* zera ulertzen da: menpeko aldagai bezala hartutako aldagaiek, hots, *erreflektagailu* (e.b. Costner, 1969) edo *ondorio indikatzaile* (e.b. Blalock, 1971) deritzotenek zein punturaino islatzen duten neurtu nahi den ezaugarri teorikoa. Bestalde, *eragilearen konstruktua balideziaz* honako hau adierazten du: aldagai asketzat hartutako tratamenduek, hau da, *ekoizle* (Costner, 1969) edo *zergati indikatzaile* (Blalock, 1971) izenekoek zein punturaino errepresentatzen duten behatutako ondorioaren eragile bezala hartzen den konstruktua teorikoa.

Konstruktua batek balidezia-maila handia izan dezan, Campbell eta Fiskek (1959) *balidezia bateragarria* eta *balidezia bereizgarria* deritzoten bi balidezia-motak izan behar ditu. *Ondorio -eragindako- konstruktuen* kasuan, neurriak *balidezia bateragarria* handia dauka interes gune den konstruktua teorikoan jazotako aldakortasun gehiena azaltzen badu. Horretarako, konstruktua beraren neurri desberdinek elkarrekiko bateratasuna edo koerlazio handia agertu behar dute. Bestalde, neurriak jasandako aldakuntzak ez dira ikertzailearentzat interes teoriko gutxi edo eza daukaten konstruktuen aldakuntzen ondorio izan behar, hau da, ondorio indikatzaileak ez du azaldu behar interes gune den objektutik desberdinak diren konstruktua teorikoen aldakortasuna. Kasu honetan, konstruktua desberdinen operazionalizazio desberdinek ez dute elkarren arteko koerlazio agertu behar. Azken honek neurriak *balidezia bereizgarria* handia duela adierazten du.

Vallejok (1991a) dioen bezala, *zergati konstruktuen* kasuan, ikertzaileak tratamendua egoki operazionalizatu behar du, berorrek interes gune den tratamendu teorikoa errepresenta dezan (*balidezia bateragarri handia*) eta azterketa gune ez diren beste zenbait tratamendu, hala nola esperimendatzaileak itxaroten duena, ikerketa testuinguruaren ezaugarriak, etab. isla ez ditzan (*balidezia bereizgarri handia*).

Berez *kanpo balideziari* dagozkion arren, *populazioaren* edo *subjektuen konstruktua balideziaz* eta *testuinguruaren konstruktua balideziaz* hitz egin daiteke ere. *Populazioaren konstruktua balideziaz* zera ulertzen da: lagina osatzen duten subjektu zehatzek zein punturaino errepresentatzen duten interes gune den populazio teorikoa eta ez ikasketa objektu ez den beste edozein populazio. Modu berean, *testuinguruaren konstruktua balideziaz* zera adierazten du: ikerketa burutzeko erabilitako testuinguruak zein punturaino islatzen duen interes gune den ingurunea, hots, *testuinguruaren konstruktua teorikoa* deritzona.

Psikologiaren eremuan, konstruktua egoki operazionalizatzea, teoria psikologikoen garapen nahiz zehaztasun mailaren eta ikertzaileak ikerketa objektu zehatzei buruz daukan ezagutza mailaren menpe dago funtsean. Hori horrela, Cook eta Campbell (1979) aburuz, *eragile eta eragindakoaren konstruktua balidezia* handia lortu nahi duen ikertzaile orok ondoko ekintzak burutu beharko lituzke: (1) konstruktua definitzeko moduari buruzko hausnarketa sakona egin, (2) interes gune diren konstruktua aztertu behar ez diren antzeko konstruktuetatik bereiztu, (3) ikasketa objektu diren konstruktuen indikatzaile egokitzat har daitezkeen neurriak edo manipulazioak finkatu, (4) konstruktua hipotetikoen neurri anizkoitzak hartu edo manipulazio anizkoitzak egin eta (5) neurriak hartzerakoan edo manipulazioak egiterakoan metodo anitz erabili.

Balidezia mota honi ematen dioten definizioan esplizituki agertzen ez den arren, Cook eta Campbell (1979) zera azpimarratzen dute: *konstruktua balidezia* ez datzala soilik konstruktua teorikoen eta berauek operazionalizatzeko erabilitako indikatzaileen arteko egokitzen maila aztertzean, izaera askoz dinamikoagoa izatean eta ikerketan lortutako ondorioetara ere egokitu ahal izatean baizik. Zentzu honetan, edozein ikerketatan behatutako datuak oso baliagarriak gerta daitezke ikertzaileak ondorio -eragindako- eta zergati -eragile- konstruktuei buruz dituen ideiak zuzendu eta datuetara haseran planteatutako konstruktua baino hobekiago egoki litezkeen beste konstruktua batzuk proposatu ahal izateko.

Cook eta Campbell (1979) helburu nagusia zergati -eragile- eta ondorio indikatzaileen eta interes gune diren konstruktua teorikoen arteko egokitzen eza ezartzea bada ere, Cronbach eta Meehle (1955) honela ulertzen dute *konstruktua balidezia*: konstruktua teorikoa zehatz batek jokatu balu aurrean litekeen datu multzoaren eta ikerketan benetan lortzen den datu multzoaren arteko egokitzenetik abiatuz, konstruktua hipotetikoei buruzko inferentziak egiteko posibilitatea bezala. Autore hauek "*sare nomologiko*" hitza erabiltzen dute edozein konstruktua definitzeko aukera ematen duen *aurreandako harreman multzoak* hitz egiteko. Egia esan,

kontzeptu hau Cook eta Campbellen ereduan inpliziturik agertzen zaigu, bertan ikasketa objektu den konstruktua konstruktua hori bera errepresentatzen duten indikatzaileekin erlazonatu behar dela eta beste antzeko indikatzaileekin, berriz, inolako harremanik ez duela agertu behar azpimarratzen denean.

5.2.- Konstruktua balideziaren aurkako mehatxuak

Atori (1991) jarraituz, *konstruktua balidezia* bi arrazoi nagusirengatik narria daiteke: (1) Konstruktua teorikoak ez direlako egoki definitzen eta, ondorioz, ezta operazionalizatzen ere eta (2) operazionalizazio desberdinen arteko erlazioak ez direlako enpirikoki aztertzen. Arazo biak gainditzeko modua *konstruktuari buruzko teoria egokia* eraikitzean datza. Teoria honek zera eskatzen du: konstruktua zehatz-mehatz definitzea eta berari dagokion *sare nomologikoa* zehaztea, hots, konstruktua beste zenbait konstruktuekin eta beste zenbait operazionalizazioarekin nola harremanetan den azaltzea. Behin teoria hori eraiki ondoren, konstruktua orok konstruktua balidezia handia izan dezan bete beharreko bi aurretikoak, balidezia bateragarria eta balidezia bereizgarria izatea alegia, betetzen ote diren aztertu behar da.

Campbell eta Fiskek (1959) *ezaugarri anitz-metodo anitz matrizea* deritzonaren analisisa proposatzen dute bai indikatzaileen balidezia bateragarria eta baita balidezia bereizgarria aztertzeko ere. *Ezaugarri anitz-metodo anitz matrizea*, bi edo metodo (indikatzaile) gehiagoren bidez neurtutako bi edo ezaugarri (konstruktua) gehiagoren arteko koerlazioak agertzen dituen matrizea da (Pedhazur eta Pedhazur Schmelkin, 1991). Oro har, konstruktua beraren indikatzaileen arteko koerlazio handiek *balidezia bateragarria* adierazten dute. Konstruktua desberdinen indikatzaileen arteko koerlazio txikiak, berriz, *balidezia bereizgarriaren* adierazle dira. Campbell eta Fiskek (1959) ezaugarri anitz -metodo anitz matrizean oinarrituz balidezia bateragarria eta balidezia bereizgarria aztertu ahal izateko zenbait irizpide eskaintzen dituzte. Beste ikertzaile batzuk matrize mota hauek aztertzeko hurbilpen desberdinak proposatu dituzte ere (berrikuspenetarako ikus Alwin, 1974; Schmitt, Coyle eta Saari, 1977; Schmitt eta Stults, 1986). Autore gehienek ustez (e.b. Marsh eta Byrne, 1993), analisi prozedurarik egokiena ezaugarri anitz-metodo anitz matrizean *faktore analisi baieztatzailea* egitea da.

Aurrekoaz gain, Atok (1991) *tratamendu efektuen tamainan oinarritutako metodo* bat proposatzen du indikatzaileen balidezia bateragarria eta bereizgarria ebaluatzeko. Konstruktua beraren indikatzaile desberdinetan *tratamendu efektuak* aurkituz gero, *balidezia bateragarria* dagoela ondoriozta genezake. Efektu horiek konstruktua desberdinen indikatzaileetan ez aurkitzeak, berriz, indikatzaileek *balidezia bereizgarria* dutela ondorioztatzera eramango liguke.

Oraintxe deskribatu ditugun metodoetatik ondoriozta daitekeen bezala, indikatzaile anitzen erabilera (*operazionalizazio anizkoitza*) beharrezkoa gertatzen da

ondorioaren konstruktua balidezia ebaluatu ahal izateko. Atoren (1991) iritziz, logika berbera, baina ez zorrotasun analitiko bera, aplikatu diezaiekegu zergatien, subjektuen eta testuinguruaren konstruktua balideziari ere.

Azaldu dugunarekin bat, konstruktua balideziaren aurkako mehatxu nagusienak konstruktuen dimentsio guztiak kontutan hartzen ez dituzten operazionalizazioak (*konstruktuen azpierrepresentazioa*) edo erreferentzia konstruktua teorikorako garrantzitsuak ez diren dimentsioak kontutan hartzen dituzten operazionalizazioak (*konstruktuari atxikiriko irrelebantziak*) egitea da. Arazo hauez gainera, ikerketa aurrera doan heinean bai saiakuntza subjektuek eta baita esperimendatzailerak ere izan ditzaketen pertzepzio, motibazio, itxaropen, etab.-en eraginez, berauengan eman daitekeen portaera aldakuntzatik eratorritako errore sorburuak kontutan hartu behar ditugu ere. Jarraian, konstruktua balidezia zalantzan jartzen duten faktore nagusienak eta berauek kontrolatzeko erabili behar diren kontrol teknikak edo prozedurak deskribatzen ditugu. Lehenago aztertu diren balidezia-moten (ondorio estatistikozko balideziaren eta barne balideziaren) kasuan bezala, oraingoan ere Cook eta Campbellen (1979) ereduari helden diogu funtsean.

5.2.1.- Azalpen operazional desegokia

Erreferentzia konstruktua teorikoa definitzeko erabilitako operazionalizazioak *konstruktua horrek berak kontzeptualki adierazten duena kontutan hartu behar du eta berorren ezaugarri nagusiak islatu behar ditu*. Bestela, konstruktua teorikoaren azalpena zehazgabea da eta konstruktua balidezia txikia agertzen du. Esate baterako, *euskarri soziala* konstruktua aztertu nahi bada, bere operazionalizazioa kontzeptu horren analisi sakon batetik abiatu beharko da. Honela, ez ditu euskarri soziala jasotzen duenaren pertzepzio eta sentimenduak soilik bildu behar, baizik eta, beste zenbait alderdiren artean, euskarri hori eskaintzen duten pertsonen ekintzak ere kontutan hartu beharko ditu.

5.2.2.- Operazionalizazio bakar baten erabilerak eragindako erroreak

Ikerketa askotan *zergati -eragile- konstruktua* definitzeko manipulazio bakarra planteatzen da eta, modu berean, *ondorio -eragindako- konstruktua* bakoitza ebaluatzeko ere neurri bakarra proposatzen da. Cook eta Campbellek (1979) dioten bezala, operazionalizazio bakunek erreferentzia konstruktua teorikoak azpierrepresentatzera jotzen dute eta, gainera, irrelebantziak barne ematen dituzte. Horregatik, interes gune den konstruktua teorikoa definitzeko operazionalizazio bakarra erabiltzen duten ikerketen konstruktua balidezia, operazionalizazio anizkoitzak erabiltzen dituztenena baino txikiagoa da.

Zergati -eragile- konstruktuen operazionalizazio anizkoitzak osoaldi gutxitan erabiltzearen arrazoiak ulertzen lagun diezaguket zenbait argudio badira.

Lehenik, tratamendu kopuruaren gehikuntzak subjektu gehiagoren beharra dakar eta, diseinu faktorialen kasuan, tratamendu baldintza bakoitzaren barnean subjektu kopurua txikiegia izatea eragin dezake. Bestalde, tratamenduak aurkeztean unetik- - unera zenbait gora-behera ematen bada, ondorio estatistikozko balidezia (fidagarritasuna tratamenduak ezartzerakoan, hain zuzen ere) zalantzan jar daiteke. Arazo hauek egon arren, Cook eta Campbellek uste dute *indikatzaile anizkoitzen erabilera konstruktua balidezia areagotzeko prozedurarik egokiena* dela eta, beraz, ahal den guztietan erabiltzea gomendatzen dute. Gainera, *ondorio -eragindako- konstruktuen* kasuan, neurri edo indikatzaile desberdinen bidez datu osagarriak biltzek ez du lan gehiegi suposatzen.

Zergati -eragile- konstruktuen operazionalizazio anizkoitzaren adibide gisa, demagun dietetika eta nutrizioaren eremuan programa bat ezartzen duen *profesionalaren esperientziak* programa hartzaileen *elikadur ohituretan* izan dezakeen eragina aztertu nahi dela.

Kasu honetan, zergati -eragile- konstruktuen hiru indikatzaile erabil ditzakegu: Donostiako ospitalean lan egiten duen prestigio handiko gizonezko medikua, Ipar Amerikako osasun zentru bateko emakumezko mediku prestigiosoa eta afrikar ospitale bateko ospe handiko gizonezko mediku beltza. Honela, indikatzaile horien arteko desberdintasunek (sexuak eta nazionalitateak) eragindako aldakortasuna azter daiteke eta, azterketa horren bidez, erreferentzia konstruktua teorikoari lotutako irrelebantzien konbinaketa desberdinek datuei modu desberdinean eragiten ote dieten jakin genezake. Azken buru, beste heterogeneotasun sorburu batzuk baztertuz, programaren erantzule den profesionalaren esperientziak subjektuen elikadur ohituretan eraginik ba ote duen egiaztatu daiteke.

5.2.3.- Operazionalizazio metodo bakar baten erabilerak eragindako erroreak

Edozein operazionalizazio metodoren balidezia subjektuekin, esperimentatzailearekin eta, oro har, ikerketa eremuarekin harremanetarako zenbait faktoreen mehatxuaren menpe dago. Horregatik, manipulazio guztiak modu berberan aurkezten direnean edo neurri guztiak sistema berberaz jasotzen direnean, metodo bera irrelebantzia gisa har daiteke eta bere eragina ezin daiteke konstruktua teorikoaren eraginatik bereiztu. Egoera horretan, ezin genezake jakin subjektuen erantzunak zein punturaino diren esperimentatzailearen espektatiben, ikerketa burutzeko erabilitako testuinguru zehatzaren, etab.-en ondorio. Azken buru, emaitzak aldagai askearen mailak aurkezteko eta subjektuen erantzunak jasotzeko erabili diren metodo zehatzen ondorio izan daitezke hein handi batean. Irrelebantzia hauek interes gune den konstruktua teorikotik ezin badira bereiztu, ikerketaren ondorioak guztiz mugatuta egongo dira (Maher, 1978).

Aurreko ataleko (5.2.2) adibidearekin jarraituz, profesional guztiek dietetika eta nutrizioari buruzko programa bideo zinta baten bidez ezarri izan balute, ezinezkoa gertatuko litzateke, esate baterako, programa, subjektuekin harreman

zuzena mantenduz ezarri izandako kasura emaitzak orokortzea. Egoera honetan, "profesionalaren esperientzia" delako konstruktua teorikoa azpierrepresentaturik dago eta egokiago da "programa bideo zintzen bidez ezartzen duen profesionalaren esperientzia" bezala definitzea.

Metodo bakarra erabiltzetik eratorritako arazoak eta alborapenak gainditzearren, zenbait autorek *konstruktua bera operazionalizatzeko metodo anitz erabiltzea proposatu dute*. Ikertzaile batzuk *operazionalizazio anizkoitza* deitzen diote proposamen honi (e.b. Garner, Hake eta Erikson, 1956). Beste batzuk *triangulapen* izena ematen diote (e.b. Denzin, 1978) eta badago berau izendatzeko *metodo anitzeko hurbilpen* hitza erabiltzen duen hirugarren talde bat ere (e.b. Campbell eta Fiske, 1959). Prozedura honi dagokion izena edozein izanik ere, berorrek operazionalizazio metodo bakarra erabiltzetik eta metodoaren eta beste zenbait faktoreren arteko elkarrengindaketatik eratorritako erroreak atzemateko aukera ematen digula azpimarratu beharra dago.

5.2.4.- Konstruktua eta konstruktua mailen arteko nahasketa

Ikertzea batzutan, izaera jarraia duen aldagai baten zenbait maila diskretu manipulatu dira. Konstruktuen maila guztiak aztertzen ez direnez gero, X-ek (aldagai askeak) Y-ri (menpekoko aldagaiari) ez diola eragiten ondoriozta genezake, aldagai aske horrek ikerketan zuzenean manipulatuta ez den maila edo balioen bat hartu izango balu subjektuen erantzunetan eragina izan zezakeen arren.

Mehatxu hau ager daiteke aldagai askearen mailak ez direnean modu egokian hautatzen eta arazo garrantzitsu bilaka daiteke X-en balioez osatutako segidan zehar X eta Y-ren arteko erlazioa lerro tankerakoa ez denean. Errore sorburu hau agertzeko probabilitatea efektuaren tamaina edo tratamenduaren eraginaren tamaina urritzen den neurrian areagotzen da. **Mehatxu hau kontrolatzeko metodorik eraginkorrena**, ikerketak burutzen direnean *ustezko zergatiaren maila asko manipulatzeko eta ustezko ondorioaren maila asko neurtzeko datza*.

5.2.5.- Tratamendu anizkoitzen arteko elkarrengindakoa

Tratamendu anizkoitzen arteko elkarrengindakoa subjektubarneko diseinuekin (errepikapen neurketa osoa eta partzialeko diseinuekin) lan egiten denean, hots, subjektu berberak bi edo tratamendu gehiago elkarren segidan jasotzen

dituztenean agertzen zaigu. *Barne balideziaren* aurkako mehatxuak aztertzerakoan aipatu den bezala, tratamendu anizkoitzen arteko interakzioak *hurrenkera efektuak* eta *hondakin efektuak* deritzotenak sortarazten ditu, zeintzuk barne balidezia zalantzan jartzen duten (ikus 4.2.5. azpiatala). Alabaina, elkarreragindako honek konstruktua balideziari ere eragiten dio, subjektuek tratamendu bakarra jasotzen duten egoeretara emaitzak orokortzeko posibilitatea zalantzan jartzen baitu eta, are garrantzitsuagoa dena, tratamenduaren efektuak tratamendu bat baino gehiago barne ematen dituen testuinguruaren efektuetatik bereizteko aukerarik ez du ematen. Guzti horrek irrelebantziak sortarazten ditu, hauen eragina interes gune den konstruktua eraginarekin nahasten delarik.

Mehatxu hau kontrolatzeko prozedurarik egokiena, *subjektuei tratamendu bakarra aplikatzean edo tratamendu bakoitzarekin analisi bereiztuak egitean datza. Errepikapen neurketako karratu latinozko diseinuak eta cross-over edo ordezkio diseinuak* erabil daitezke baita ere errore sorburu honi aurre egiteko.

5.2.6.- Tratamenduaren eta aldez aurretik aplikatutako testen arteko elkarreragindakoa

Mehatxu hau agertzen zaigu subjektuei *testaurre* bat ezartzen diegunean edo subjektu talde bereberek *errepikapen neurriak* hartzen ditugunean. Egoera hauetan, ondoren jasoko den tratamenduari (edo tratamenduei) dagokionez testaurreak eragin dezakeen sentsibilizazioak, *barne balideziari* eragiteaz gain (ikus 4.2.3. atala), *konstruktua balidezia* ere zalantzan jartzen du. Izan ere, testaurrerik gabeko egoeretara edo menpeko aldagaiaren neurketa bakarra barne ematen duten egoeretara ustezko zergati -eragile-ondorio -eragindako- erlazioa edo tratamenduaren eragina orokortzeko posibilitatea murrizten du. *Tratamenduaren eta aldez aurretik aplikatutako testen arteko elkarreragindakoak* errorerik sortu ote duen jakiteko, testoste neurri bakarra jasotzen duen kontrol talderen bat erabili behar da. Honela, efektuen agerpenak testaurrea aplikatzearekin edo testoste neurri bat baino gehiago jasotzearekin zerikusirik ba ote duen jakin dezakegu. Hori horrela izango balitz, konstruktua bere izaera eraldatuko luketen zorizko irrelebantziak aurkeztuko lituzke.

Mehatxu hau kontrolatzeko prozedurarik egokiena, *testak aplikatzen diren saio bakoitzean elkarrekiko askeak diren saiakuntza talderekin lan egitea da.*

5.2.7.- Konstruktua batetik bestera orokortzeko aukera eskasa

Zenbait kasutan, tratamenduek modu desberdinean eragiten dute menpeko aldagai desberdinetan. Honela, aldagai batzutan eragin positiboa izan dezakete eta

beste batzutan, berriz, eragin eza edo negatiboa izan daiteke. Ondorioz, tratamendu zehatz baten benetako eragina zein den egoki estimatzea zaila gertatzen da. Adibidez, *funtzionarioei zuzendutako Eusko Jaurlaritzako euskalduntze plangintzak* subjektuen *euskara gaitasuna* hobe dezake, baina agian ez dio eragiten funtzionarioek *euskararekiko agertzen duten jarrerari*.

Tratamendu zehatz baten eragina jasan dezaketen ondorio -eragindako- konstruktuko guztiak neurtzeak suposatzen duen zailtasunaz kontziente izanik, Cook eta Campbellek (1979) tratamenduek ikerketan jatorriz planteatutako konstruktuetatik desberdinak diren beste konstruktuko batzuri eragiten ote dieten jakitearren, subjektu lagin desberdinak erabiltzea gomendatzen dute. Aipatutako planteamenduari jarraituz, **mehatxu honen efektuak murrizteko prozedurarik egokiena zergati -eragile- eta ondorio -eragindako- konstruktuen indikatzaile anitz erabiltzean datza.**

5.2.8.- Egoerarekin eta elkarrekintza sozialarekin zerikusia duten faktoreetatik eratorritako mehatxuak

Atal honetan aurkezten diren errore sorburuek harreman herstua mantentzen dute *egoerak eragindako efektu erreaktiboak* izenekoekin, zeintzuk barne balideziaren aurkako mehatxuak aztertzerakoan azaldu genituen (ikus 4.2.11. azpiatala). Gogora dezagun ez dagoela batere argi efektu hauek *barne balideziaren* aurkako edo *konstruktuko balideziaren* aurkako mehatxu bezala sailkatu behar ote diren, tratamendu esperimentalaren eragina nahasteaz gain, erreferentzia konstruktuko teorikoen interpretazio egokia ere eragotz baitezakete. Dena den, bai efektu horiek eta baita atal honetan aurkezten ditugunak ere, saiakuntza subjektuek eta esperimendatzaileak ikerketan zehar garatzen dituzten pertzepzioen ondorioz agertzen dira. Jarraian, pertzepzio horiek sortaraz ditzaketen errore sorburu nagusienak deskribatzen ditugu. Gai honetan sakondu nahi duen ikasleak badu izan alderdi desberdinak aztertzen dituen literatura zabala, hala nola ikerketaren izaera soziala (e.b. Brenner, 1981; Brenner, Marsh eta Brenner, 1978; Friedman, 1967; Wuebben, Straits eta Schulman, 1974), gizabanakorekin burutzen diren ikerketak (e.b. Adair, 1973; Barber, 1976; Silverman, 1977), esperimendatzailearen efektuak (e.b. Rosenthal, 1966) edo artefaktu mota hauek oro har (e.b. Barber, 1976; Rosenthal eta Rosnow, 1969) aztertzen dituzten lanak, hain zuzen ere.

5.2.8.1.- Hipotesien sorkuntza testuinguruko ezaugarrietan oinarrituz

Subjektuek esperimendatzaileak ikerketan aurkitzea espero dituen emaitzei buruz edo berauengan espero duen portaerari buruz interpretazioak edo aieruzioak egiten dituztenean agertzen zaigu mehatxu hau. Aieruzio mota hauek egiteko Ornek (1962, 1969) *saiakuntza egoerako eskakizunaren ezaugarriak* deritzonaz baliatzen

dira. Kontzeptu honek zera adierazten du: subjektuei saiakuntza hipotesia zertan datzan asmatzeko aukera ematen dien eta berauen portaera mugatzen duen testuinguruko ezaugarrien multzoa, alegia. *Saiakuntza egoerako eskakizunaren ezaugarriek* honelako alderdiak barnean ematen dituzte: esperimentatzaileak subjektuei aurkezten dizkien argibideak, subjektuek burutu behar duten zeregin-mota, ingurunearen ezaugarri fisikoak, subjektuek saiakuntzari buruz entzuten dituzten zurrumurruak, ikerketan parte har dezaten eskaintzen zaien informazioa, etab. Azken batez, ikerketaren helburuei eta hipotesiei buruz zenbait aieruzio garatzeko aukera ematen dieten eta berauen portaerarengan eragin handia izan dezaketen elementu anitz barne ematen dituzte.

Azaldu dugunatik ondoriozta daitekeen bezala, *saiakuntza egoerak* berak efektu erreaktiboak eragiten ditu. Hortaz, esperimentuan parte hartzen ari direnaren jakite hutsak subjektuek tratamenduen aurrean emandako erantzunak eralda ditzake. Zentzu honetan, behin eta berriz ikusi izan da subjektuek ez dutela berdin jokatzeko egoera naturaletan eta beren portaera behatua eta ebaluatua izaten ari dela dakiten egoeretan. Hori horrela bada, emaitzak ezin daitezke tratamenduaren eragin hutsa bezala azaldu; aldiz, teorikoki garrantzitsuak diren beste aldagai batzuren eragina ere kontutan hartu behar dugu. Horregatik, zergati -eragile- konstruktua subjektuek saiakuntza egoerako zenbait elementuri emandako erantzunekin nahasten da. Saiakuntza egoerak eragin ditzakeen efektu erreaktiboen adibide argigarria *Hawthorne efektua* bezala ezagutzen dena da (Roethlisberger eta Dickson, 1939). Efektu hau aurkitu zen Western Electric Company delakoaren Hawthorneko lantegian *ingurune fisikoaren aldaketek produktibitatean* izan zezaketen eragina aztertzeko asmoz zenbait ikerketa burutzen ari zirelarik. Emaitzek agerian utzi zuten inguruneke edozein aldaketak, bai positiboa (argi gehiago, atsedenaldiak, etab.) eta baita negatiboa (argi gutxiago, soldaten igoera eza, etab.) izanik ere, produktibitatea gehitzen zuela.

Saiakuntza egoerako eskakizunaren ezaugarrien eragina kontrolatzeko aukera ematen digun prozedura bat *kontrol talderen erabilera* da. **Hipotesien asmatzeko** arazoa, subjektuentzat ikertzailearen hipotesiak zeintzuk diren asmatzea zaila gerta dadin ahaleginduz saihas daiteke. Hori lortzeko, *saiakuntzaren erreaktibitate maila orokorra murriz dezakegu* edo *subjektu desberdinei hipotesi desberdinak nahita eman diezazkiekegu*. *Itsu bikoitza teknika* ere mehatxu honi aurre egiteko sarritan erabili ohi den prozedura egokia da. Alabaina, soluzio hauek guztiak osagabeak dira, saiakuntza subjektuen jarrera ez baita pasiboa eta, beraz, esperimentatzailearen hipotesiekin adostasunean edo desakordioan egonik ere, subjektuek beren hipotesi bereziak planteatzen baitituzkete.

5.2.8.2.- Saiakuntza subjektuek hartzen dituzten errol desberdinak

Atok (1991) dioen bezala, subjektu esperimentalak modu aktibo eta kontzientean parte hartzen du saiakuntzan, baina ikerketa horretan bertan dauden

gainontzeko subjektuen desira, espektatiba eta portaeren menpe eta, azken buru, *saiakuntza egoerako eskakizunaren ezaugarrien* menpe dago. Ondorioz, egoera esperimentalak modu berezian hautematen du eta hautemate horrek ikerketaren helburu nagusienarekin zerikusirik ez duten aldagaien arabera bere portaera aldatzera eramanez diezaioke. Weber eta Cooki (1972) jarraituz, subjektuak saiakuntza egoeran har ditzakeen lau errol nagusiak deskribatzen ditugu (beste berrikuspen batzuk irakurri nahi izanez gero, ikus Adair, 1973; Silverman, 1977). Egoera naturaletan aurkeztu diren errolak diren errolak hartzeak, emaitzak ikerketako testuinguru zehatzetik kanpo orokortzeko aukera murrizten du.

Subjektu zintzoa

Ornek (1962, 1969) *subjektu zintzoaren* deskribapen bikainak eskaintzen ditu. Errol hau hartzen duen subjektuak oso iritzi ona dauka zientziari eta saiakuntzari buruz eta, berauek baieztatze asmoz, ikertzailearen hipotesiak asmatzen saiatzen da. Ikertzailearen eskakizunak benetan bitxiak edo arriskutsuak izate arren, subjektu zintzoak eskakizun horiei guztiei erantzun nahi izaten die, saiakuntzan parte hartze hutsa horrela jokatzeko arrazoi nahikoa dela pentsatzen baitu eta esperimentatzaileak besteentzat nahiz norberarentzat kaltegarria izan litekeen zerbaitegitera ez liokeela inolaz ere bultzatuko uste baitu. Subjektu hauek lankidetasuna eta pazientzia erakusten dute eta beren burua gaiztakeriarik gabeko pertsonak bezala aurkeztu dute, beraien erantzunak ikerketaren emaitzetan kontutan hartuko ez direnaren beldur baitira.

Subjektu irudikorra

Subjektu irudikorra (Rosenberg, 1965, 1969) esperimentatzaileari inpresio ona egiteaz oso kezkatutik dagoen (Silvermanek, 1977, *subjektu harroa* deitzen dio) eta Rosenbergekin *ebaluatua izateari aurrejarrera* deritzona, hots, epaitua edo ebaluatua izateari beldurra azaltzen duen gizabanakoa da. Beren burua oso pertsona gai eta psikologikoki helduak bezala aurkezteko ahaleginak egiten dituzten subjektuak dira. Subjektuek errol hau hartzeko probabilitatea oso garaia da berauek ikertzailea goi mailako estatusa (e.b. unibertsitateko irakaslea) nahiz pertsonak epaitzeko gaitasun berezia (e.b. psikologoa) duela hautematen dutenean edo subjektuek burutu behar dituzten zereginak beren trebetasuna edo nortasun ezaugarri bereziak agerian uzten dituztenean.

Subjektu leiala

Subjektu leialak (Fillenbaum, 1966; Fillenbaum eta Frey, 1970) hipotesi esperimentalarik buruz inolako ezaugarri izango ez balu bezala jokatu du. Eskaintzen zaizkion argibideak modu serio eta kontzientean jarraitzen ditu eta

esperimentatzailearen eskakizunei ahalik eta hobekien erantzuten saiatzen da. Atoren (1991) iritziz, jarrera hau subjektu zintzoarenaren hedapen gisa har daiteke, zein, esperimentatzaileari datu erabilgarriak lortzen laguntzeko asmoz, zeregin oro egoki burutzen ahalegintzen den. Bai subjektu leialak eta baita subjektu zintzoak eta subjektu irudikorrak ere, esperimentatzailearekiko esanekotasuna nahiz menpekotasuna eta taldearen portaerarekiko hurbiltze maila garaia agertzen du.

Subjektu ezeztagarria

Aurreko hiru errolak hartzen dituzten subjektuek lankidetasuna eta jarrera positiboa aurkezten duten bitartean, *subjektu ezeztagarriak* (Silvermanek, 1977, *subjektu gaiztoa* deitzen dio) ez du lankidetzarako jarrerarik erakusten. Aitzitik, esperimentatzaileari ez dio jaramonik egiten edo bere kontra jartzen da. Millerrek (1976) honela definitzen ditu subjektu hauek: hipotesia onartzeko aurkeztu beharko luketeenaren aurkako portaera agertzen duten gizabanakoak bezala, hain zuzen ere. Errol honen aurkikuntza Maslingi (1966) dagokio eta *Brehmen erreaktantzia teoriaren* (1966) bidez azaldu ohi da, hots: beren burua askatasunik gabe eta esperimentatzailearen borondatearen menpe ikusteak, subjektuak saiakuntza hipotesia bete ez dadin ahaleginak egitera eramaten ditu. Alabaina, autore gehienek aburuz (e.b. Kruglanski; 1976; Weber eta Cook, 1972), erreaktantzia oso gutxitan agertzen da saiakuntza egoeratan.

5.2.8.3.- Esperimentatzaileak eragindako erroreak

Autore gehienekin ados, azalpen honetan *esperimentatzaile* hitza oso modu zabalean erabiltzen da, hau da, ikertzaile, behatzaile, teoriko edo portaeraren zergatiak aztertzeaz kezkatutako zientzilariaren errola har dezakeen gizabanakoa bezala. Saiakuntza subjektua bezala, esperimentatzailea ez da behatzaile pasiboa, saiakuntzaren emaitzetan eragin dezakeen elementu aktiboa baizik. Rosenthal eta bere laguntzaileek esperimentatzaileak saiakuntza subjektuengan izan dezakeen eraginari buruzko ikerketa anitz egin dute (Rosenthal, 1976; Rosenthal eta Rosnow, 1984; Rosenthal eta Rubin, 1978). Barberrek (1976) ere, esperimentatzaileak har ditzakeen errol desberdinetatik eratorritako mehatxuak sakonki aztertu ditu. Oro har, *esperimentatzaileak* bi motako *errore sistematikoak* eragin ditzakeela onartzen da: *elkarrekintzan oinarritzen ez diren efektuak* eta *elkarrekintzan oinarritutako efektuak*, alegia.

Elkarrekintzan oinarritzen ez diren efektuak

Elkarrekintzan oinarritzen ez diren efektuak esperimentatzaileak datuak biltzerakoan, analisi estatistikoak egiterakoan edo emaitzak interpretatzerakoan eragin ditzakeen erroreari dagozkie.

Atok (1991) *behatzailearen efektua* deitzen dio esperimentatzaileak gertakizunen behaketan eta neurketan egindako erroreari. Efektu honen barnean sartzen dira neurketa prozesuaren inguruan, hots, neurketa sistemen eraikuntza, hautapena, aplikazioa eta ebaluazioaren inguruan esperimentatzaileak hartutako erabakietatik eta burututako ekintzetatik eratortzen diren mehatxuak. Pedhazur eta Pedhazur Schmelkinek (1991) dioten bezala, nahasketa sorburu hau saihesteko kontutan hartu behar ditugun alderdirik garrantzitsuenak neurketen balidezia eta fidagarritasuna dira.

Datuak aztertzerakoan egindako erroreei dagokienez, sarritan agertzen direla eta arazo hau sakonki aztertzen duten lanak ere badirela esan beharra dago (e.b. Hooke, 1983; Huff, 1954; Kimble, 1978). Besteak beste, errore horien artean ondorengoak aurki ditzakegu: formula desegokien erabilera, askatasun gradu kopuruaren estimazio desegokia, analisiaren aurretikoak betetzen ez dituzten analisi tekniken erabilera, etab. Nahasketa sorburu honi buruz hitz egiterakoan Atok (1991) zera dio: bere hipotesia baieztatzen duten emaitzak lortzeko asmoz, zenbait kasutan ikertzaileak benetako datuak faltsutzen dituela, fenomeno honi *berariazko efektua* deritzolarik.

Hirugarrenik, emaitzak interpretatzerakoan ere, esperimentatzailea sarritan tronpatzen dela ikusi ohi da (*interpretatzailearen efektua*).

Elkarrekintzan oinarritutako efektuak

Elkarrekintzan oinarritutako efektuak esperimentatzailearen jokabideak subjektuen portaeran eta, ondorioz, saiakuntzaren emaitzetan izan dezakeen eraginari dagozkio. Elkarrekintzan oinarritutako efektuen artean, bi efektu mota nagusi bereiz daitezke: *esperimentatzailearen ezaugarriak* eta *esperimentatzailearen espektatibak*, hain zuzen ere.

- 1.- **Esperimentatzailearen ezaugarriak.** Efektu hau, esperimentatzailearen ezaugarri biologiko, psikologiko eta psikosozialetatik erator daitezkeen errore sistematikoen multzoari dagokio. Rosenthalek (1966) ezaugarri horiek hiru kategoriatan sailkatzea proposatzen du. Lehenik *ezaugarri biosozialak* ditugu, zeintzuen artean esperimentatzailearen adina, sexua, arraza eta erlijioa bezalako ezaugarriak sar ditzakegun. Ezaugarri hauei dagokienez, ez dakigu saiakuntza subjektuek ezaugarri desberdineko esperimentatzaileen aurrean erantzun desberdinak ematen ote dituzten edo subjektuen aurrean jokabide desberdinak aurkezten dituztenak ezaugarri biosozial desberdineko esperimentatzaileak ote diren. Bigarren kategoriak *ezaugarri psikosozialak* deritzotenak biltzen ditu. Ezaugarri hauen artean antsietate maila, onespenezko sozialaren beharra, etsakeria, autoritarismoa, adimen maila, estatus sozioekonomikoa, menperatze gaitasuna, etab. aipa daitezke. Azkenik,

hirugarren kategoriak *egoerari dagozkion ezaugarriak* barneratzen ditu, non esperimentatzaileak eta saiakuntza subjektuek tratamendu aurretik elkarren artean izandako harremanak eta esperimentatzailearen esperientzia sartuko liratekeen. Esperimentatzailearen ezaugarriak aztertzen dituzten lan anitzen berrikuspenak eta laburpenak Adair (1973), Barber (1976), Friedman (1967), Mahoney (1976), Rosenthal (1966) eta Rosenthal eta Rosnowen (1969) testuetan irakur daitezke.

- 2.- **Esperimentatzailearen espektatibak.** Efektu hau, esperimentatzaileak saiakuntzan emaitza bezala lortzea espero duenatik eta espektatiba hauek subjektuei heleraztetik erator daitekeen erroreari dagokio. Errore sorburu honen inguruko ekarpenik nagusienak Rosenthalek (1963, 1966, 1969, 1976) egin ditu. Autore honek dioenez, ikertzailearen eta saiakuntza hipotesiaren artean sortzen den harreman afektiboak, esperimentatzaile askok hipotesi hori baieztatzen duten emaitzak lor ditzan eragiten du. Rosenthalen ustez, fenomeno hau esperimentatzaileak bere hipotesiaren aldeko erantzunei kontrakoei baino jaramon gehiago egiten dielako edo, beren portaera saiakuntza hipotesira egoki dezaten, subjektuei testuinguruko seinale zorrotzak inkontzienteki igortzen dizkielako gerta daiteke. Honela, saiakuntza egoeraren eta esperimentatzaileak subjektuekin mantentzen duen harremanaren arabera, espektatibek esperimentatzailearen ahotsean, keinuetan, aurpegi adierazpenetan, argibideak aurkezteko moduan, etab. zenbait aldaketa eragin dezakete, eta aldaketa horiek subjektuen erantzunei eragin diezaiekete. Dena den, Vallejok (1991a) dioen bezala, esperimentatzailearen iradokizun ezkutuek eta intentziogabeek ez dute Rosenthalek uste duen bezainbateko mehatxua suposatzen. Izan ere, Rosenthal eta bere languntzaileek 1962-67 bitartean argitaratutako artikuluei buruz Barber eta Silverrek (1968) egindako berrikuspenean zera agertu zen: bi ikerketek soilik egiaztatzen zutela esperimentatzaileak subjektuei inkontzienteki helerazten dien seinaleen ondorioz, esperimentatzailearen espektatibek saiakuntzen balidezia zalantzan jar dezaketela. Planteamendu honekin ados, Adairrek (1978) zera baieztatzen du: lan eremu honetako arazorik nagusiena, esperimentatzailearen espektatiben efektuak azal ditzaketen mekanismoak zehazteko zailtasunean datzala.

Efektu hau kontrolatu ahal izateko, Rosenthalek (1976) *espektatiben kontrolerako diseinua deritzona erabiltzea* gomendatzen du. Berau diseinu faktorial bat da, non *espektatibak* aldagaiak diseinuan parte hartzen duten gainontzeko faktoreetatik aske jokatzeko duen. Honela, espektatiben eraginaren probabilitatea eta tamaina finkatzeko aukera daukagu. Rosenthal eta Rosnowek (1984) esperimentatzailearen espektatiben eragina neurri handian murrizteko aukera ematen duten beste zenbait prozedura ere eskema baten bitartez laburtzen dizkigute. Prozedura horiek ondorengoak dira: *esperimentatzaile kopurua handitu, esperimentatzaileen jokabidea behatu, hurrenkera efektuak aztertu, esperimentatzailearen eta subjektuen arteko harremana gutxitu,itsu bikoitza teknika*

aplikatu eta espektatiben kontrolerako taldeak erabili, hain zuzen ere. Rosenthal eta Rosnowen (1984) testua, metodo horien bidez *espektatibak* aldagaia nola kontrola daitekeen sakonkiago aztertu ahal izateko erreferentzia bezala har daiteke.

5.2.8.4.- Subjektuak eta esperimentatzaileak eragindako errorearen kontrola

Subjektuak eragindako errorearen kontrola

Saiakuntza egoerako eskakizunaren ezaugarrien ondorioz subjektuek garatzen dituzten pertzepzioek eragindako errorea kontrolatzeko prozedura egokien arteko bat, *itsu bikoitza teknika* da (ikus 4.3.2. atala). Kontrol teknika hau aplikatzen denean esperimentatzaileak subjektuak zein tratamendu taldetan dauden ez dakienez gero, ezin die informazio hori igorri eta, beraz, ez du beraien erantzunetan eragiten.

Subjektuek saiakuntza hipotesiaren inguruan gara ditzaketen pertzepzio desberdinek eragindako alborapena kontrolatzeko aukera ematen duen beste teknika bat, *iruzurra* deritzona da. Prozedura hau, subjektuei benetan kontrastatu nahi diren estatistik desberdinak diren hipotesiak igortzean datza.

Errore mota hauek kontrolatzeko hirugarren metodo bat, subjektuek ikerketan parte hartzen ari direla jakin ez dezaten ahaleginak egitean datza. Prozedura honi *saiakuntza ezkutuariaren teknika* deritzo. Alabaina, *saiakuntza ezkutuak* lekuan lekuko ikerlanak izaten direla eta, beraz, aldagai arrotzak egoki kontrolatzeko aukera eskasa eskaintzen dutela azpimarratu beharra dago.

Oraintxe deskribatu ditugun hiru prozedurez gain, Christensenek (1981) eta Adair eta Spinnerrek (1981) beste zenbait metodo proposatzen dituzte. Metodo horiek bi kategoria handitan sailka daitezke: *atzeranzko hitzezko txostenak* eta *hitzezko txosten lehiakideak* izenekoetan, alegia.

Atzeranzko hitzezko txostenek barnean ematen dituzten teknikan subjektuari saiakuntzaren alderdi zehatzak atzebegiraz oroitzeko eskatzen zaio: esperimentuaren helburuei buruz zer pentsatzen zuen, esperimentatzaileak saiakuntzan lortzea espero zezakeenari buruz zein motatako ideiak zituen, etab.

Azkenik, *hitzezko txosten lehiakideek* prozedura multzo bat barnean ematen dute, zeintzuetan subjektuak saiakuntzari buruz garatzen dituen pertzepzioak eta pentsamenduak saiakuntza burutu bitartean adierazten dituen.

Esperimentatzaileak eragindako erroreen kontrola

Esperimentatzaileak datuak biltzerakoan egin ditzakeen erroreak kontrolatzeko prozeduren arteko bat *behatzaile edo kalifikatzaile anizkoitzak erabiltzean* datza. Gainera, kontrol gehiago lortzen da *behatzaileek edo kalifikatzaileek subjektuei dagozkien saiakuntza baldintzak zeintzuk diren ez dakitenean*, hots, baldintza horiekiko "itsuak" direnean (Rosenthal, 1978). Dena den, neurketa erroreak kontrolatzeko metodorik egokiena *neurketa tresna automatikoak edo estandarrek erabiltzean* datza.

Esperimentatzailearen ezaugarriei dagokienez, hauek errore sorburu bihur ez daitezen, *menpeko aldagaiarekin koerlazonaturik daudenak egonkor mantendu behar dira tratamendu baldintza desberdinetan zehar*.

Azkenik, lehenago aipatu den bezala (5.2.8.3. atala) **esperimen-tatzailearen espektatibak kontrolatu ahal izateko** *espektatiben kontrolerako diseinua, itsu bikoitza teknika* eta *espektatiben kontrolerako taldeak* erabil daitezke, beste zenbait prozeduren artean.

5.1. Taulan konstruktua balideziaren aurkako mehatxu nagusien eta berauek kontrolatzeko erabil daitezkeen prozeduren laburpena aurkezten dizuegu.

5.1. Taula.- Konstruktua balideziaren aurkako mehatxu nagusienak eta berauek kontrolatzeko prozedurak

MEHATXUA	DEFINIZIOA	KONTROL PROZEDURAK EDO TEKNIKAK
1.- Azalpen operazional desegokia	Erreferentzia konstruktua teorikoaren definizioak ez ditu bere ezaugarri nagusiak kontutan hartzen eta, beraz, ezin daiteke konstruktua horren operazionalizazio egokitza hartu	- Konstruktua teorikoaren analisi kontzeptual egokia burutu, bere ezaugarri garrantzitsu guztiak adierazi (irudikatu) ahal izateko
2.- Operazionalizazio bakar baten erabilerak eragindako erroreak	Zergati -eragile- eta ondorio -eragindako- konstruktuen operazionalizazio bakunak erreferentzia konstruktua teorikoak gutxieraz jotzen dute eta, gainera, irrelebantziak barne ematen dituzte. Ondorioz, konstruktua balidezia murrizten dute	- Konstruktuen indikatzaile anizkoitzak erabili
3.- Operazionalizazio metodo bakar baten erabilerak eragindako erroreak	Erreferentzia konstruktua teorikoa operazionalizatzeko metodo bakarra erabiltzen denean, metodo bera irrelebantzia gisa har daiteke eta bere eragina ezin daiteke konstruktuen eraginatik bereiztu	- Konstruktua bera operazionalizatzeko anitz metodo erabili
4.- Konstruktua eta konstruktua mailen arteko nahasketa	Konstruktuen maila guztiak manipulatu ez direnean, tratamenduek subjektuen erantzunetan eraginik ez dutela oker ondoriozta genezake	- Ustezko zergatiaren maila asko manipulatu eta ustezko ondorioaren maila asko neur-tu

5.1. Taula (jarraipena).- Konstruktu balideziaren aurkako mehatxu nagusienak eta berauek kontrolatzeko prozedurak

MEHATXUA	DEFINIZIOA	KONTROL PROZEDURAK EDO TEKNIKAK
5.-Tratamendu anizkoitzen arteko el-karrera-gindakoa	Subjektu berdinei bi edo tratamendu gehiago elkarren segidan aplikatzeak, subjektuek tratamendu bakarra jasotzen duten egoeretara emaitzak orokortzeko posibilitatea zailtzen du eta ez du tratamenduaren efektuak tratamendu bat baino gehiago barnean ematen duen testuinguruaren efektuetatik bereizteko aukerarik ematen	- Subjektuei tratamendu bakarra aplikatu - Tratamendu bakoitzarekin analisi bereiztuak egin - Errepikapen neurketako karratu latinozko diseinuak eta cross-over edo ordezeko diseinuak erabili
6.-Tratamenduen eta alde zuretik aplikatutako testen arteko elkarre- ragindakoa	Testaurrea ezartzeak edo M.A-ren neurri desberdinak hartzeak konstruktuari lotutako irrelebantziak sortarazten ditu eta testaurerik gabeko egoeretara edo M.A-ren neurketa bakarra barne ematen duten egoeretara emaitzak orokortzeko posibilitatea zailtzen du	- Probak aplikatzen diren saio bakoitzean elkarrekiko askeak diren saiakuntza taldeekin lan egin
7.- Konstruktua batetik bestera orokortzeko aukera eskasa	Zenbait kasutan, tratamenduek modu desberdinean eragiten dute konstruktua desberdinetan. Horregatik, tratamendu zehatz baten benetako eragina zein den egoki estimatzea zaila gertatzen da	- Zergati -eragile- eta ondorio -eragindako- konstruktuen indikatzaile anitz erabili
8.- Hipotesien sorkuntza testuinguruko ezaugarrietan oinarrituz	Subjektuak saiakuntza egoerako eskakizunaren ezaugarrietaz baliatzen dira saiakuntza hipotesia asmatzeko eta esperimintatzaileak berauen portaerari buruz zer espero duen atzeman ahal izateko	- Kontrol taldeak erabili - Saiakuntzak sor dezakeen erreaktivitate maila gutxitu - Subjektu desberdinei hipotesi desberdinak aurkeztu - Itsu bikoitza teknika
9.- Saiakuntza subjektuek hartzen dituzten errol desberdinak	Saiakuntza egoerako eskakizunaren ezaugarrien ondorioz subjektuek saiakuntzari buruz garatzen dituzten pertzepzioek, subjektuek errol desberdinak har ditzaten eragiten dute. Errol horien artean, subjektu zintzoarena, subjektu irudikorrena, subjektu leialarena eta subjektu ezeztagarriarena azpimarra daitezke	- Itsu bikoitza teknika - Iruzurra - Saiakuntza ezkutua - Atzerantzko hitzeko txostenak - Hitzeko txosten lehiakideak
10.- Esperimintatzaileak eragindako erroreak	<i>Elkarrekintzan oinarritzen ez diren efektuak:</i> Esperimintatzaileak datuak modu desegokian biltzean, analisi estatistikoak oker egitean edo datuak gaizki interpretatzean eragiten dituen erroreak	- Behatzaile edo kalifikatzaile anizkoitzak erabili (kontrola areagotzen da behatzaileak edo kalifikatzaileak subjektuei dagozkien saiakuntza baldintzekiko "itsuak" direnean) - Neurketa tresna automatikoak (estandarrek) erabili
	<i>Elkarrekintzan oinarritutako efektuak:</i> Esperimintatzailearen jokabideak subjektuen erantzunetan izan dezakeen eraginari dagozkio	- M.A-rekin koerlazio naturik dauden esperimintatzailearen ezaugarriei egonkor eutsi tratamendu baldintza desberdinetan zehar - Espektatiben kontrolerako diseinua, itsu bikoitza teknika, espektatiben kontrolerako taldeak, etab. erabili

**7.- Balidezia-mota desberdinak ikerketaren eremuan:
elkarren arteko harremanak eta lehentasunak**

7.- BALIDEZIA MOTA DESBERDINAK IKERKETAREN EREMUAN: ELKARREN ARTEKO HARREMANAK ETA LEHENTASUNAK.

Azalpenean zehar ikusi den bezala, *ondorio estatistikozko balidezia* aldagai aske(ar)en eta menpeko aldagai(ar)en artean egon daitekeen erlazio funtzionala hautemateko gaitasunari dagokio. *Beste hiru balidezia motak*, berriz, harreman horren interpretazio egokiarekin lotzen dira funtsean. Honela, *ondorio estatistikozko balideziak* tratamenduak azaldutako aldakortasuna hauteman ahal izateko adina zorrotzak diren proba estatistikoak erabiltzea eskatzen du. Bestalde, *gainerako balidezia-motek* alborapen sistematikoa eragiten duten sorburuak kontrolatzera eta sistematikoak ez diren faktoreek sortutako aldakortasuna (errore bariantza) murriztera jotzen dute. *Barne balidezia*, hain zuzen ere, subjektuak saiakuntza talde desberdinetara zoriz esleitzearen eta, tratamendu aldagaian ezezik, gainerako alderditan guztiz egonkorrak diren baldintza esperimentalak ezartzearen menpe dago funtsean. *Konstruktuko balideziak*, erreferentzia konstruktuko teorikoak xehetasunez errepresentatu ahal izateko, tratamendu baldintzen eta hauek eragindako ondorioen operazionalizazio egokia eskatzen du. Eta *kanpo balideziak*, azkenik, emaitzak zein zirkunstantzietara orokortu nahi diren argi eta garbi finkatzera eta subjektu, testuinguru, denboraldi nahiz tratamendu baldintzei loturiko beste zenbait ezaugarriren laginketa egokia egitera behartzen gaitu. Azken batez, ondorio zehatz, fidagarri eta baliagarriak lortu nahi badira, ikerketak ondorio estatistikozko balidezia, barne balidezia, konstruktuko balidezia eta kanpo balidezia maila garaia aurkeztu behar du. Alabaina, balidezia mota horiek guztiak ikerketaren barnean funtsezko alderditzat har daitezkeenez gero, balidezia mota baten edo bestearen lehentasunari buruzko erabakiek faktore anitz kontutan hartu beharko dituzte, hala nola ikerketaren helburuak, kostu ekonomikoak, ikerketatik erator daitezkeen ondorioak, etab. Hori dela eta, garrantzitsutzat jotzen dugu gai honen inguruan zenbait hausnarketa egitea.

Pedhazur eta Pedhazur Schmelkinek (1991) dioten bezala, bi pentsamendu korrante bereiz ditzakegu. Lehenengoak, burutzen den ikerketaren izaeraren arabera balidezia-moten artean lehentasunak finkatzen ditu. Bigarren korranteak, berriz, balidezia-moten artean lehentasun kontzeptua baztertzen du. Zenbait autorek, balideziaren inguruan garatutako sailkapenen erabilgarritasuna ere zalantzan jartzen dute (e.b. Olson eta Peter, 1984). Dunnek, gainera, ez barne balideziak ez eta kanpo balideziak ere, ikerketa batean benetan garrantzitsuak diren arazoei ez dagokiela erdiesten du. Gure ustez, bai balideziaren sailkapenak eta baita balidezia mota baten edo bestearen lehentasunari buruzko hausnarketek ere helburu guztiz praktikoa dute eta erabilgarriak gertatzen dira ikerketaren eremuan. Horregatik, azalpen honek lehen pentsamendu korronteari helduko dio.

Cook eta Campbellen iritziz (1979; ikus baita ere, Cook, Campbell eta Peracchio, 1990), balidezia mota baten gehikuntzak beste balidezia mota baten

gutxitzea dakar normalki. Esate baterako, *ondorio estatikozko balidezia* gehiago eta *barne balidezia* gehiago lortzearren, ikertzaileak subjektu, testuinguru, denboraldi, etab. ahalik eta homogeenak erabiltzen baditu, emaitzen orokorpena edo *kanpo balidezia* guztiz mugatua geratzen da. Halaber, esperimenteraileak subjektuen erantzunei eragiten dizkieten estimulak zorrozki kontrolatzen baditu, *ondorio estatistikozko balidezia* eta *barne balidezia* areagotzen dira, baina bai *kanpo balideziak* eta baita *konstruktuko balideziak* ere aurkako efektua jasaten dute. Bestalde, *ondorioaren konstruktuko balidezia* operazionalizazio anizkoitzaren bidez gehituz gero, neurketa prozesua luzea eta aspergarria gerta daiteke eta horrek, *hilkortasun esperimenterailean* eragiteaz gain, *gizabanakoen erantzunen fidagarritasun maila murriz* dezake. Dena den, Cook eta Campbellek bereziki azpimarratzen dutena *barne balideziaren* eta *erreprezentatibitate* edo *kanpo balideziaren* arteko gatazka da. Autore hauen aburuz, bi balidezia-moten artean alderantzizko harremana dago. Honela, *barne balideziaren* gehikuntzak *kanpo balideziaren* murrizketa dakar, eta alderantziz. Egungo zenbait ikertzailek ere (e.b. Arnau, 1990; Christensen, 1988; Kazdin, 1992) iritzi berbera mantentzen dute, hots, *barne balidezia* areagotzeko asmoz zenbat eta kontrol gehiago aplikatu, talde guztiz homogeen baten erantzunak laboratoriorako testuinguru artifizialean denboraldi zehatzean jasoaz, adibidez, orduan eta mugatuagoa egongo da emaitzak beste pertsona, testuinguru edo denboraldi batzuetara orokortzeko posibilitatea. Aitzitik, *kanpo balidezia* areagotzearren, saiakuntza subjektu talde heterogeenak erabiliz, testuinguru eta denboraldi desberdinetan zehar egiten bada, nahasketa aldagaien kontrola zailagoa gertatzen da eta, hortaz, *barne balidezia* murrizten da.

Gatazka honen aurrean, Cook eta Campbellek (1979) Campbell eta Stanleyren (1966) planteamenduari heltzen diote, non *barne balidezia* ikerketaren "*sine qua non*"-a bezala hartzen den. Autore hauen iritziz, lehentasuna *barne balideziari* eman behar zaio, *barne balideziarik* gabeko saiakuntzak ez baitu ondorio fidagarriak lortzeko aukerarik ematen. Cook eta Campbelli jarraituz, ikertzaile ororen helburua saiakuntza eremuan, kausazko erlazio funtzionalak aurkitzean datza, hau da, efektu zehatz bat(zuk) zergati -eragile- zehatz bat(zur)i dago(z)kiola zalantzarik gabe ondorioztatu ahal izatean. Horregatik, *barne balideziari* garrantzi berezia eman behar zaio esparru esperimenterailean. Behin erlazio funtzionala ezarri ondoren, berau beste subjektu, testuinguru edo denboraldi batzuetara orokor ote daitekeen azter dezakegu. Alabaina, badira *barne balideziari* emandako lehentasuna baztertu eta balidezia mota hori *kanpo balideziaren* pean egon beharko litzatekeela defendatzen duten zenbait ikertzaile. Ideia honen aldeko autoreen artean Cronbach (1979, 1982) dugu nagusi. Kruglanskik (1975), bestalde, bi balidezia-moten arteko gatazka ukitzerakoan, psikologian ikerketa egiteko bi modu desberdin bereiz daitezkeela aipatu ohi du, *unibertsalista* bata eta *berezilea* bestea, bakoitzak helburu nahiz ezaugarri desberdinak dituelarik. Campbell eta Stanleyren eta Cook eta Campbellen posizioa lehenengo moduarekin bat etorriko litzateke eta Cronbachenak, berriz, ikuspegi bereizleari jarraituko lioke.

Barne balideziaren nagusitasuna inolaz ere zalantzan jartzen ez den arren, Cook eta Campbell (1979) **burutzen ari den ikerketaren izaera teoriko edo aplikatuaren arabera balidezia-moten artean lehentasunak** ezartzen dituzte. Honela, *ikerketa teorikoaren* eremuan murgilduriko ikertzaileak lehenik *barne balidezia* lortzen saiatu beharko liratekeela planteatzen dute. Ondoren, *konstruktibaleziaren* eskuratu beharko litzateke (oinarrizko psikologian *zergatiaren -eragile-konstruktibaleziaren* ondorioarenak *-eragindako-* baino garrantzi gehiago hartu beharko luke). *Ondorio estatistikozko balidezia* konstruktibaleziari jarraituko lioke eta azken lekuan *kanpo balidezia* izango genuke. Bestalde, *ikerketa aplikatuaren* esparruan ondoko lehentasun ordena proposatzen dute: *barne balidezia*, *kanpo balidezia*, *ondorioaren konstruktibaleziaren* *balidezia*, *ondorio estatistikozko balidezia* eta *zergatiaren -eragile-konstruktibaleziaren*, hain zuzen ere. Ere eremu aplikatuan *kanpo balidezia* ematen zaion garrantziarekin bat etorritik, Kazdinek (1992) honako hau azpimarratzen du: ekarpen aplikatuak dituen edozein aurkikuntzatan, psikologia klinikotik eta hezkuntza psikologiatik datozenetan sarritan gertatzen den bezala, emaitzak orokortzeko posibilitateak garrantzi berezia hartu beharko lukeela, alegia. Judd eta Kennyren (1981) iritziz, berriz, esparru teorikoan ezarri den lehentasun ordena ikerketa tradizio luzean oinarritzen den bitartean, testuinguru aplikatuan balidezia mota guztien garrantzia berbera izan dela dirudi. Ondorioz, autore hauek ez dira ausartzen ikerketa aplikatuan gehienbat bultzatu beharko litzatekeen balidezia mota zein den zalantza izpirik gabe planteatzera.

Egungo ikertzaile askok Cook eta Campbell (1979) jarrera defendatzen jarraitzen duten arren, balidezia mota desberdinen arteko harremanak gatazkaren arabera definitzen ez dituzten autoreak ere badira. Zentzu honetan, Vallejo (1991a) dio maila molarrean garrantzi sozial handia duten teknika anitz, zorrotasun esperimental garaieko lanetan oinarriturik daudela eta balidezia mota zehatz baten aurkako mehatxuak kontrolatu ahal izateko estrategiaren bat erabiltzeak ez dakarrela ezinbestean kontrako balidezia motaren bat gutxitzea. Vallejoen iritziz, lau balidezia-moten artean sor daitezkeen gatazkak batera konpontzeko aukerarik ez balego, esperimentatzaileak berak ikerketaren helburuak kontutan hartuz bultzatu beharreko balidezia mota zein den erabaki beharko luke.

Psikologian erabiltzen diren diseinuen balidezia *buruzko planteamendu berriak* aztertzen ikusiko dugun bezala, Chen eta Rossik (1987) proposatutako *teorian oinarritutako ereduak ("theory-driven approach")*, balidezia mota bakarrak gehiegi arduratzeari eta gainerakoak baztertzeari uko egiten dio. Egileek, diseinuan baino gehiago, ikertzailearen ezagutza teorikoan oinarritzen den eta aldagai aske(ar)en eta menpeko aldagai(ar)en arteko erlazioa azal dezaketen zergatiak sakonki aztertzen ahalegintzen den ereduak proposatzen dute. Cook eta Campbell

(1979) ereduak ez bezala, Chen eta Rossiren (1987) planteamenduak barne balidezia, kanpo balidezia, konstruktua balidezia eta ondorio estatistikozko balidezia aldi berean gehitzeko zenbait proposamen egiten ditu. Alabaina, eta autore hauen ekarpena gutiesteko asmorik gabe, beren ereduak, beste arrazoi batzuren artean, lau balidezia motak batera areagotzeko proposamen zehatzak eta bideragarriak ez aurkezteagatik kritikatu izan dela esan beharra dago. Chen eta Rossiri kontrajarriz, Cordrayk (1989) uste du ikerketa programen ebaluazioaren egungo esparruan, esate baterako, balidezia mota desberdinen artean lehentasunak finkatu behar direla eta lehentasun hauek ikertzaileak lortu nahi dituen helburuen pean (teorikoak, praktikoak, etab.) egon behar direla.

Lau balidezia motak guztiz komenigarriak direla onartuz ere, kasu gehienetan lehentasunak ezartzea behar-beharrezkoa dela uste dugu. Horregatik, edozein balidezia motaren lehentasunaren inguruko erabakia saiakuntza eta sasisaiakuntza metodologiaren arteko edo oinarritzko ikerketa eta ikerketa aplikatuaren arteko bereizketatik abiatu beharko litzatekeela deritzogu. Oinarritzko ikerketaren helburu nagusia teorien ebaluapena denez gero, testuinguru honetan barne balideziaren lehentasuna zalantzaezina da. Ikerketa aplikatuaren eremuan, berriz, non interbentzio zehatz baten eraginkortasuna eta berau aplikatu ahal izateko aukera ezinbestekotzat hartzen diren, bai barne balidezia eta baita kanpo balidezia ere garrantzi maila berean ezarri beharko genituzke.

9.- Psikologian erabiltzen diren diseinuen balideziari

buruzko planteamendu berriak

9.1.- Marken eredua

9.2.- Chen eta Rossiren eredua

9.3.- Campbellen iradokizun berriak barne balidezia eta
kanpo balidezia kontzeptuen inguruan eta Cook eta
Shadishen hausnarketak diseinuaren balideziaz
testuinguru sozialean

9.- PSIKOLOGIAN ERABILTZEN DIREN DISEINUEN BALIDEZIARI BURUZKO PLANTEAMENDU BERRIAK

Liburu honi amaiera emateko, egokia deritzogu balideziaren zenbait alderdi ikuspegi berri batetik lantzen dituzten bi planteamendu azaltzea. Gure iritziz, proposamen hauek ezin daitezke Cook eta Campbellen ereduaren ordezkotzat hartu, zein egungo ikertzaile gehienek artean bizirik dirauen, baina ikerketaren kalitatea hobetzeko erabilgarri gerta daitezkeen ideia batzuk ekartzen dizkigute. Marken eredua aurkeztuz hasten gara eta, ondoren, Chen eta Rossiren eredua eta beroni egin zaizkion zenbait kritika azaltzen ditugu (eredu biak 2.2.5. atalean laburki deskribatuak izan dira). Azkenik, Campbellek berak barne eta kanpo balidezia kontzeptuei buruz berriki egin dituen iradokizun batzuk, bi balidezia-mota horiek izendatzeko proposatzen duen terminologia berria eta duela gutxi Cook eta Shadishek gizarte zientzien alorrean diseinuaren balideziaren inguruan azken hamarkadan jazotako garapenari buruz egindako hausnarketa batzuk aurkezten dizkizuegu.

9.1.- Marken eredua

Markek (1986), bere ustetan autore klasikoek behar adina landu ez dituzten eta ikerketaren balidezia adierazgarriki areagotzeko aukera ematen duten zenbait alderditan sakontzearen beharra planteatzen du. Marken ereduaren puntu edo proposamen nagusienak honako hauek dira:

- 1.- **Goi mailako orokorpenak egiten saiatu.** Marken ustez, barne balidezia indartzeaz gainera, behe mailako indikatzailetatik abiatuz goi mailako estrapolazioak egin ahal izateko beharrezko urratsak eman behar dira. Autorearen iritziz, bai Cook eta Campbellek (1979) eta baita Cronbachek (1982) ere helburu hori lortzeko planteatzen dituzten iradokizunetan, hiru printzipio onartzen dira inplizituki: (a) *Antzekotasun printzipioa*, hots, ordezeko laginetatik abiatuz gero, antzeko ezaugarriak dituzten populazioei buruzko auresanak egin daitezkeenaren baieztapena; (b) *Egonkortasun printzipioa*, zeinen arabera, orokorpenaren posibilitatea areagotzen den efektua zirkunstantzia desberdinetan zehar behin eta berriz agertzen bada, eta (c) *Azalpen printzipioa*, fenomenoaren zergatik ematen den edo zergatik ez den ematen jakiteak, etorkizunari begira auresanak egiteko aukera ematen duela dioena, hain zuzen ere.

- 2.- **Kausazko prozesuak aztertu.** Marken iritziz, balideziari buruz orain arte proposatu diren sailkapenek ez dituzte kausazko prozesuak esplizituki aztertzen, ez eta kausazko inferentzien balideziaren aurkako mehatxu posibleak kontuan hartzen ere. Gabezia honen aurrean, autoreak zera planteatzen du: diseinuaren balidezia neurri handian gehi daitekeela tratamendu zehatz baten eragina azal dezakeen sekuentzia kausala aztertzen bada. Markek kausazko prozesua aztertzeko bost hurbilketa posible proposatzen ditu: (a) kausazko eredutzea, behaketa kualitatiboa, etab. bezalako teknika desberdinak erabiliz, *prozesua zuzenean behatu*; (b) *Blokeo eredua*, ustez zergatia denaren eragina inhibitzen duen *blokeo eragile* bat aurkitzean eta eragile horren eraginkortasunaren arabera kausazko prozesuari buruzko inferentziak egitean datzana; (c) *Areagotze eredua*, aurreko ereduaren logika berberari jarraitzen diona, baina kausazko mekanismo hipotetiko bat inhibititu ordez, mekanismo hori aktibatzean oinarritzen dena; (d) *Arazketa eredua*, prozesu kausala, tratamenduan eta neurketan parte hartzen duten kausazko elementuetatik abiatuz aztertzean datzana eta (e) *Behaketen eta aurrean arteko egokitzapen eredua*, non behaketa multzo baten eta ikasketa objektu diren kausazko prozesuetan oinarritutako auresate multzo baten arteko adostasuna aztertzen den.

Markek bezala, egungo beste ikertzaile batzuk ere, *prozesu kausalen azterketak* ikerketaren balidezia areagotzen duela defendatzen dute. Bai Trochimek (1985) eta baita Reichardtek (1986) ere, konstruktua balidezian eta barne balidezian oso eragin positiboa duela baieztatzen dute. Halaber, goi mailako orokorpenak errazten ditu. Eta azkenik, kasu askotan tratamendu indartsuagoak edo eraginkorragoak aplikatzeko aukera eskaintzen du.

- 3.- **Balideziaren sailkapenak dogmatikoki ez onartu.** Gomendio honen bitartez, Markek, balideziaren tipologiek ikertzailearen beraren arrazonomendu kritikoaren lekua hartzen badute, eragin negatiboa izan dezaketela ohartarazi nahi du. Planteamendu honi jarraituz, Kahnemann eta Tverskyk (1973) eta Reichardtek (1985) zera diote: zenbait kasutan, sailkapen sistema horiek mehatxu posibleak hautematea eragoz dezaketela eta, beraz, ikertzaileari ondorio desegokiak lortzera eraman diezaioketela. Zentzu honetan, Reichardtek, Seaverrek (1973) burututako ikerketa baten adibidea eskaintzen du, non autoreak ez zuen *erregresio estatistikoa* deritzon mehatxua atzeman, bere ikerketa, normalki nahasketa sorburu hori barne ematen duten ikerketetatik oso desberdina zelako, hain zuzen ere. Bestalde, balideziaren tipologietan konfidantza gehiegi izateak, kontrol tekniarik behar ez denean ere, berauek aplikatzera eraman diezaguke.

Arazo hauei loturik, Markek dio oraindik ez daukagula balideziaren aurkako mehatxuak agertzeko probabilitatea eta beren garrantzia modu zehatzean estimatzeko aukera ematen digun estrategiarik. Hori horrela da, aipatutako estimazioa ikerketaren helburuen eta berorri dagokion testuinguruaren menpe dagoelako batik bat. Horregatik, ikertzailearen beraren analisi kritikoa eta ezagutza teorikoak estrategia hori garatzeko ezinbesteko elementu gisa hartzen dira.

- 4.- **Balidezia garrantziarekin ez nahastu.** Beste zenbait autorek bezala (e.b. Cronbach, Ambron, Dornbush, Hess, Hornik, Phillips, Walker eta Weiner, 1980; Krathwohl, 1985), Markek uste du balidezia egokia izateak ez duela esan nahi ikerketan lortutako ondorioak aplikazio eremu zehatz baten barnean garrantzitsuak direnik.
- 5.- **Ez ulertu balidezia "guztia-deus ez" arazoa izango bailitzan.** Mark, inferentziak kausazko harremanak egoteari edo ez egoteari buruzko erabaki dikotomikotzat hartzen dituzten eredu klasikoetatik aldentzen da. Honela, gehiago hurbiltzen zaio Reichardten (1986) planteamenduari, zeina balidezia efektuaren tamaina zehatza daukaten inferentzien multzoa bezala ulertzen duen.

9.2.- Chen eta Rossiren eredua

Zenbait alditan adierazi dugun bezala, eredu hau psikologian burutzen diren interbentzio programen ebaluazio eremuan sortzen da eta *teorian oinarritutako eredua* ("*theory-driven approach*") bezala ezagutzen da. Bere argudio nagusia, edozein interbentzio programaren ebaluazio egokiak eredu edo teoria bat formulatzea eskatzen duela eta eredutze horrek ikerketaren balidezia zalantzan jar dezaketen alborapen sorburuak kontrolatzeko beharrezko estrategiak barnean eman beharko lituzkeela esatean datza. Tratamendu baldintzen eta efektuen arteko erlazioak azal ditzaketen *kausazko mekanismoen* ulerkuntzan sakontzearen, Chen eta Rossik *aldagai garrantzitsuen* eta *bitarteko prozesuen* arteko harremanak eredutzea proposatzen dute. Balidezia-mota desberdinen artean lehentasunak ezarri ordez, *teorian oinarritutako ereduak* barne balidezia, kanpo balidezia, konstruktua balidezia eta ondorio estatistikozko balidezia areagotzeko esparru kontzeptuala eskaintzen digu (ikus 2.1. Irudia eta berari dagokion azalpena 2.2.5. atalean).

Barne balideziari dagokionez, autoreen ekarpenik nagusia zera esatean datza: ikertzaileak ikasketa objektuari buruzko ezagutza sakonik ez duenean, zorizkapena dela barne balidezia lortzeko estrategiarik egokiena. Aitzitik, interes gune den fenomenoari buruzko ezagutza zabala duenean, barne balideziaren aurkako mehatxuak kontrolatzeko modurik hoberena, aldagai mugatzaileak eta kausazko erlazioan eragin dezaketen elementu guztien arteko elkarreragindakoak barne ematen dituen eredu teorikoa zehaztean datza. Gainera, ereduaren zehaztapenak gainerako balidezia-motak ere hobetzen dituela baieztatzen dute.

Ondorio estatistikozko balideziari dagokionez, Chen eta Rossik diotenez, beren planteamenduak zorizkapenarekiko aurkezten duen abantailarik nagusia, aldagai arrotza(k) ereduaren barnean sartzean datza. Horrek errore bariantza eta konfidantza tarteak murrizten ditu eta, ondorioz, efektua hautemateko probabilitatea handitzen du. Bestalde, programen ebaluazio eremuan murgilduriko ikertzaileek tratamendu baldintzen hautapen egokiari eta hauek ezartzeko moduari eman dioten garrantzi eskasa kritikatzten dute.

Konstruktua balideziari dagokionez, erreferentzia konstruktuen indikatzaile anitzen erabilera eta eraikitzen diren kausazko ereduaren barnean neurketa erroreak kontuan hartzearen garrantzia azpimarratzen dute.

Azkenik, *kanpo balidezia* kontzeptualizatzeko modu berri bat, hots, *orokorpen esplizitu* eta *orokorpen inplizituaren* arteko bereizketa proposatzen dute. *Orokorpen esplizitua*, etorkizunean interbentzioa jasoko duten subjektuak zehatz-mehatz definiturik dituzten eta subjektu horien azpimultzoekin lan egiten duten ebaluaizoei dagokie. Programen ebaluazio eremuan, programa populazioko zein taldeetara zuzenduko den, zein erakundek ezarriko duen eta ezarpenerako erabiliko diren prozedura-moten ezagutza zehatza eskatzen du. Orokorpen-mota hau guztiz desiragarria bada ere, ezin daiteke beti lortu, eta horregatik zenbait kasutan *orokorpen inplizitua* jo behar da. Orokorpen hau egitean, ikertzaileak aztertutako sistemaren eta etorkizunean ebaluazioaren emaitzak orokortzeko interes gune bezala hartzen den sistemaren arteko berdintasunak eta desberdintasunak finkatzen ditu eta informazio hori erreferentzia gisa hartzen du programa etorkizunean berriz ere ezartzeko.

Teorian oinarritutako ereduak ikerketen kalitatea hobetzeko oso erabilgarria dela onartzen bada ere, zenbait kritika jaso ditu. Jarraian, Chen eta Rossiren ereduan egokiro landu ez diren alderdi batzuk agerian uztearren, autore desberdinek burutu dituzten zenbait **analisi kritiko** laburtzen ditugu. Alderdi horiek sakonkiago aztertu nahi dituen irakurleak jatorrizko artikuluetara jo dezake.

Azken urteotan, Cordrayk (1989) planteatu du, Chen eta Rossiren ereduak modu egokian erabili ahal izateko hiru alderdi nagusi kontuan hartu behar direla. Ikus ditzagun alderdi horiek: (1) lehenik, *teorian oinarritutako ereduak* suposatzen du ikertzaileak bere interbentzioa aurrera eramateko inolako mugarik ez duela, eta hori ez da egia. Izan ere, programaren helburuek eta izaerak, bertan parte hartzen duten subjektuen (e.b politikoen edo laguntzaile sozialen) jarrerak, etab. interbentzioa antolatzerakoan aintzakotzat hartu behar diren muga anitz ezartzen dituzte. (2) Bigarrenik, Chen eta Rossiren planteamendua, balidezia-mota guztiak hobetzearen metodologia anitzen erabilera proposatzen dutenenarekin bat dator, baina autoreek ez dute zehazten nola aplikatu daitekeen ikuspegi anizkoitz hori. (3) Hirugarrenik, lau balidezia-motak aldi berean handitzearen Chen eta Rossiren helburua oso nekez lor daiteke, xede horrek baliabide asko eskatzen baititu. Ondorioz, Cordrayk zera proposatzen du: balidezia-moten artean lehentasunak finkatu behar ditugula, horretarako, irizpide teorikoei, diseinuaren indarrarekin zerikusia duten irizpideei edo irizpide praktikoei jarraitu behar zaielarik funtsean.

Bestalde, Palumbo eta Oliverioren (1989) ustez, Chen eta Rossiren planteamenduak interbentzio programak ezarri beharko liratekeen moduari dagokionez elaborazio gehiago behar du. Hori dela eta, autoreek ezarpenerako eredu teoriko desberdinak proposatzen dituzte eta eredu bakoitza balidezia-aren arabera eztabaidatzen dute. Ildo berberan kokatzen da Pattonen (1989) analisi kritikoa, zeinen iritziz, *teorian oinarritutako ereduak*, edozein interbentzio programaren ebaluazioan aplikatu daitekeen eredu bezala aurkezten den arren, eraginkorra izateko, baldintza oso bereziak eskatzen dituen. Zehazki, Pattonek dio Chen eta Rossiren eredu egokia izan daitekeela ikertzaileak esperientzia zabala duenean eta kausazko erlazioen ezarpenari eta orokorpenari garrantzi handia ematen dien kalitate goreneko ebaluazioa burutu nahi duenean, hain zuzen ere. Alabaina, psikologiaren esparruan egiten diren ebaluazio gehienetan, baliabideen, denboraren, esperientzia nahikoa duten subjektuen, etab.-en urritasunak kausalitate edo orokortze arazoei zehaztasunez erantzutea galeraetzen du. Halako zirkunstantzietan lortzen diren ondorioak programaren eta zenbait efektuen arteko erlazioari besterik ez dagozkio eta, hortaz, Chen eta Rossik proposatutako esparru kontzeptuala ezin daiteke erabili.

9.3.- Campbellen iradokizun berriak barne balidezia eta kanpo balidezia kontzeptuen inguruan eta Cook eta Shadishen hausnarketak diseinuaren balideziaz testuinguru sozialean

Barne balideziari dagokionez, Campbellek (1986) uste du, bere adiera klasikoan, balidezia-mota hau baldintza guztiz kontrolatuetan laboratoriora burutzen den saiakuntzari hertsiki lotzen zaiola. Autorea adiera honekin ados ez dagoenez gero, hitz berri bat, behin-behinekoa noski, proposatzen du balidezia-mota hau izendatzeko, hots, "*Kausazko Balidezia Lokala Molarra (Pragmatikoa, Teoriaz kanpoko)*". Lehenik, *Molar* hitza, barne balideziari berez lotzen zaion saiakuntza tratamendu huts eta bakunaren konnotazioa baztertzeko erabiltzen da. Campbellek, hain zuzen, zera adierazi nahi du: sarritan tratamendua elementu desberdinen arteko nahasketa konplexua dela, elementuen elkarrekiko lotura aditu baten iritziak mugatzen duelarik, eta ezin daitekeela guztiz teorikoak diren osagai desberdinen arteko batura hutsa bezala ulertu. *Lokal* hitzak, tratamenduaren efektua testuinguru eta denboraldi zehatzetan kokatu behar dela adierazten du. Bi hitz hauek, *Molar* eta *Lokal* hitzek alegia, barne balidezia kontzeptua, orokortze ahalmenik gabeko zergati-ondorio motako harremanen ezarpenari dagokiola islatu nahi dute. Bestalde, barne balidezia *Kausala* da, menpeko aldagai(et)an behatutako aldaketaren eragileetaz arduratzen delako. Azkenik, *Pragmatiko* eta *Teoriazkanpoko* hitzen bidez, balidezia-mota honen benetako helburua tratamenduak interes gune den fenomenoari eragiten diola ziurtatzean besterik ez datzala adierazi nahi da. Campbellek sortutako hitz berria, guztiz zehatza izan arren, Cook eta Shadishen (1994) esaten dute gizarte zientzien alorreko ikertzaile gehienek jatorrizko hitza erabiltzen jarraitzen dutela.

Kanpo balideziari dagokionez, Campbellek jatorrizko hitza mantentzen du eta, Cookekin batera idatzitako liburuan (Cook eta Campbell, 1979) behin eta berriz egiten duen bezala, zera azpimarratzen du: oso zaila gertatzen dela ausazko prozeduraz hautaturiko subjektu, testuinguru eta denbora aldietara orokortzea. Horren ordez, berriro esaten digu, ikerketaren ohiko praktikan askoz erabilgarriagoa eta errazagoa dela populazio, testuinguru eta denboraldi desberdinen artean orokortzea. Hala ere, orokorpen hau aurrera eramateko "*antzekotasun eta hurbiltasun printzipioa*" deritzon mekanismo berri bat proposatzen du. Printzipio honen arabera, interbentzioa berez ezartzen den jatorrizko eremuko eta etorkizunean interbentzioa jasoko duen eremuko tratamendu, testuinguru, populazio, denbora aldi, etab.-en arteko antzekotasuna handitzen den heinean, orokorpenak hainbat eta fidagarriagoak izango dira. Azkenik, Campbellek honako hau esaten digu: *antzekotasun eta hurbiltasun printzipioaren* aplikazioa, ikertzailearen intuizioan, esperientzian eta ezagutza teorikoetan oinarritu behar dela, alegia.

Amaitzeko, interesgarria iruditzen zaigu Cook eta Shadish (1994) gizarte zientzien alorrean diseinuaren balideziaren inguruan azken hamarkadan izandako garapenari buruzko hausnarketa bikainaren alderdi nagusiak hona ekartzea. Lehenik, egileek, *sasisaiakuntza* desberdinen aldean, *zorizko saiakuntzak* hartzen ari diren garrantzia azpimarratzen dute. Zorizko esleipenak arazo etiko, politiko eta legalak eragin baditzake ere (Coyle, Boruch eta Turner, 1991), estrategia hau laboratoriotik kanpo erabilgaitza denik esatea zalantzan baino kinkan jarri izan da, batik bat, testuinguru aplikatuetan egin diren benetako saiakuntza kopuru handia dela eta. Honela, gaur egun, badago egon testuinguru hauetan *zorizko saiakuntzak* egoki burutzea errazten duten faktoreei buruzko informazio nahikoa.

Saiakuntzek indar handia hartu duten arren, *sasisaiakuntzak* ere nabarmenki hobetu dira azken urteetan. Zentzu honetan ikertzaileak *subjektuen hautapen diferentzialaren* mehatxua, diseinu-mota honetan berezko nahasketa sorburutzat har daitekeena, ahal den neurrian kontrolatzearen beharraz ohartu dira. Zehazki, ildo horretan ahalegin handiak egin dira, besteak beste, erkaketa taldeak modu egokian parekatzeko, laginketa unitateen hautapenean erabilitako prozedura zehaztasunez deskribatzeko, tratamendu talde kopurua gehitzeko eta hautapen mehatxuaren izaerari buruzko hipotesi desberdinak argitu asmoz, analisi estatistiko ugari egiteko.

Cook eta Shadish planteatutako azken ekarpena *kausa inferentzien orokorpen prozesuari* eman zaion garrantzi handiari dagokio. Honi buruz, autoreek diote orokorpen hori aurrera erdiets ahal izatea, azken buru, ondorengoaren ondorioa dela, meta-analisiak izandako bilakaera batetik, eta, bestetik, eragile--eraginda erlazio zehatz baten itzalean dauden kausa mekanismoen azterketak eta azalpenak hartutako garrantzia.

BIBLIOGRAFIA AIPAMENAK

- ADAIR, J.G. (1973). *The human subject: The social psychology of the psychological experiment*. Boston: Little, Brown.
- ADAIR, J.G. (1978). Open peer commentary. *Behavioral and Brain Sciences*, 3, 386-387.
- ADAIR, J.G. eta SPINNER, B. (1981). Subjects' access to cognitive processes: Demand characteristics and verbal report. *Journal for the theory of Social Behavior*, 11, 31-52.
- ALWIN, D.F. (1974). Approaches to the interpretation of relationships in the multitrait-multimethod matrix. In H.L. Costner (Ed.), *Sociological methodology 1973-1974* (79-105 or.). San Francisco: Jossey-Bass.
- ANDERSON, D.Ch. eta BORKOWSKI, J.G. (1978). *Experimental Psychology: Research Tactics and their Applications* (2nd edition). Dallas, TX: Scott, Foresman and Company.
- ARNAU, J. (1981a). *Diseños experimentales en psicología y educación*. I. liburukia Méjico: Trillas.
- ARNAU, J. (1981b). *Diseños experimentales en psicología y educación*. II. liburukia Méjico: Trillas.
- ARNAU, J. (1990). Metodología experimental. In J. Arnau, M.T. Anguera eta J. Gómez Benito (Eds.), *Metodología de la investigación en ciencias del comportamiento* (9-124 or.) Murcia: Universidad de Murcia.
- ARNAU, J. (1995). *Diseños de investigación aplicada en esquemas*. I eta II. liburukiak Bartzelona: Universidad de Barcelona.
- ATO, M. (1991). *Investigación en ciencias del comportamiento. I: Fundamentos*. Barcelona: PPU.
- ATO, M. (1995). Conceptos básicos. In M.T. Anguera, J. Arnau, M. Ato, R. Martinez, J. Pascual eta G. Vallejo (Eds.), *Métodos de investigación en psicología*. (45-66 or.). Madrid: Síntesis.

- ATO, M. eta NAVALON, C. (1993). Validez de la investigación social aplicada: hacia un nuevo esquema conceptual. In C. Navalon eta M.E. Medina (Eds.) *Psicología y Trabajo Social*. (43- 56 or.). Barcelona: DM-PPU.
- BARBER, T.X. (1976). *Pitfalls in human research: Ten pivotal points*. New York: Pergamon Press.
- BARBER, T.X. eta SILVER, M.J. (1968) Fact, fiction, and the experimenter bias effect. *Psychological Bulletin Monograph Supplement*, 70 (6, Pt.2), 1-29.
- BARCIKOWSKY, R.S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics*, 6, 267-285.
- BERNIA, J. (1979). *Psicología experimental. I*. Valencia: NAU LLibres.
- BLALOCK, H.M. (1971). Causal models involving unmeasured variables in stimulus-response situations. In H.M. Blalock (Ed.), *Causal models in the social sciences* (335-347 or.). Chicago: Aldine.
- BORUCH, R.F. McSWEENEY, A.J. eta SODERSTROM, E.J. (1978). Randomized field experiments for program planning development and evaluation. *Evaluation Quarterly*, 2, 655-695.
- BOX, G.E. (1954). Some on quadratic forms applied in the study of ANOVA problems II. *Annals of Mathematical Statistics*, 25, 484-498.
- BRACHT, G.H. eta GLASS, G.V. (1968). The external validity of experiments. *American Educational Research Journal*, 5, 437-474.
- BREHM, J.W. (1966). *A Theory of Psychological Reactance*. New York, NY: Academic Press.
- BRENNER, M. (Ed.) (1981). *Social method and social life*. London: Academic Press.
- BRENNER, M., MARSH, P. eta BRENNER, M. (Eds.) . (1978). *The social contexts of method*. New York: St. Martin's Press.
- BROFENBRENNER, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513-531.

- BROWN, M.B. eta FORSYTHE, A.B. (1974). The ANOVA and multiples comparisions for data with heterogeneous variances. *Biometrics*, 30, 719-724.
- BRUNSWIK, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193-217.
- BRUNSWIK, E. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley, CA: University of California Press.
- BURNS, R.B. eta DOBSON, C.B. (1981). *Experimental Psychology: Research Methods and Statistics*. Lancaster: MTP Press Limited.
- CAMPBELL, D.T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54, 297-312.
- CAMPBELL, D.T. (1963). From description to experimentation: Interpreting trends as quasi-experiments. In C.W. Harris (Ed.), *Problems in measuring change* (212-242 or.). Madison: University of Wisconsin Press.
- CAMPBELL, D.T. (1969). Prospective: Artifact and control. In R. Rosenthal eta R.L. Rosnow (Eds.), *Artifact in behavioral research* (351-382 or.). New York: Academic Press.
- CAMPBELL, D.T. (1986). Relabeling internal and external validity for applied social scientists. In W.M.K. Trochim (Ed.), *Advances in quasi-experimental design and analysis* (67-77 or.). San Francisco: Jossey-Bass.
- CAMPBELL, D.T. (1988). *Methodology and epistemology for Social Science: Selected Papers*. Chicago, IL: University of Chicago Press.
- CAMPBELL, D.T. eta FISKE, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- CAMPBELL, D.T. eta STANLEY, J.C. (1963). Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (ed.) *Handbook of research on teaching* (171-246 or.). Chicago: Rand McNally.

- CAMPBELL, D.T. eta STANLEY, J.C. (1988) *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu. (jatorrizko argitalpena 1966).
- CHRISTENSEN, L. (1981). Positive self-presentation: A parsimonious explanation of subject motives. *The Psychological Record*, 31, 553-571.
- CHRISTENSEN, L.B. (1988). *Experimental Methodology* (4th edition). Boston, MA: Allyn and Bacon. (3rd edition, 1985).
- CLINCH, J.J. eta KESELMAN, H.J. (1982). Parametric alternatives to the Analysis of Variance. *Journal of Educational Statistics*, 7, 207-214.
- COCHRAN, W.G. (1983). *Planning and Analysis of Observational Studies* (Eds.: L.E. Moses eta F. Mosteller). New York, NY: John Wiley and Sons.
- COHEN, J. (1977). *Statistical power analysis for the Behavioral Sciences*. New York: Academic Press.
- COHEN, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.
- CONRAD, E. eta MAUL, T. (1981). *Introduction to Experimental Psychology*: New York, NY.: John Wiley and Sons.
- CONOVER, W.J. eta IMAN, R.L. (1981). Rank transformation as a bridge between parametric and nonparametric statistics. *The American Statistician*, 35, 131-133.
- COOK, T.D. eta CAMPBELL, D.T. (1976). The design and conduct of quasi-experiments and true experiments in field settings. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (223-325 or.). Chicago: Rand McNally.
- COOK, T.D. eta CAMPBELL, D.T. (1979). *Quasi-experimentation: Design & Analysis issues for field settings*. Chicago. Rand McNally.
- COOK, T.D. eta CAMPBELL, D.T. (1986). The causal assumptions of quasi-experimental practice. *Synthese*, 68, 141-160.

- COOK, T.D. CAMPBELL, D.T. eta PERACCHIO, L. (1990). Quasi-experimentation. In M.D. Dunnette eta L.M. Houghs (Eds.), *Handbook of industrial and organizational psychology* (491-576 or.). Palo Alto C.A.: Consulting Psychologist Press.
- COOK, T.D. eta SHADISH, W.R. (1994). Social experiments: some developments over the past fifteen years. *Annual Review of Psychology* 45, 545-580.
- COSTNER, H.L. (1969). Theory, deduction, and rules of correspondence. *American Journal of Sociology*, 75, 245-263.
- COYLE, S.L., BORUCH, R.F. eta TURNER, C.F. (1991). *Evaluating AIDS Prevention Programs: Expanded Edition*. Washington, DC: Natl. Academic Press.
- COX, D.R. (1958). *Planning of experiments*. New York: Wiley.
- CRONBACH, L.J. (1979). Las dos disciplinas de la psicología científica. In F. Alvira, M.D. Avia, R. Calvo eta J.F. Morales (Eds.), *Los dos métodos de las ciencias sociales* (93-124 or.). Madrid: CIS.
- CRONBACH, L.J. (1982). *Designing Evaluations of Educational and Social Programs*. San Francisco: Jossey-Bass.
- CRONBACH, L.J., AMBRON, S.R., DORNBUSH, S.M., HESS, R.D., HORNIK, R.C., PHILLIPS, D.C., WALKER, D.F. eta WEINER, S.S. (1980). *Toward Reform of Program Evaluation: Aims, Methods, and Institutional Arrangements*. San Francisco: Jossey-Bass.
- CRONBACH, L.J. eta MEEHL, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- CHAMBERS, J.M., CLEVELAND, W.S., KLEINER, B. eta TUKEY, P.A. (1983). *Graphical methods for data analysis*. Belmont, CA: Wadsworth.
- CHEN, H.T. eta ROSSI, P.H. (1983). Evaluating with Sense: The Theory-Driven Approach. *Evaluation Review*, 7, 283-302.
- CHEN, H.T. eta ROSSI, P.H. (1987). The theory-driven approach to validity. *Evaluation and Program Planning*, 10, 95-103.

- DAVIS, D. eta GAITO, J. (1984). Multiple comparison procedures within experimental research. *Canadian Psychology*, 25, 1-13.
- DENZIN, N.K. (1978). *The research act: A theoretical introduction to sociological methods* (2nd edition). New York: McGraw-Hill.
- DUNN, W.N. (1982). Reforms as arguments. *Knowledge Creation Diffusion Utility*, 3, (3) 293-326.
- DWYER, J.H. (1983). *Statistical Models for the Social and Behavioral Sciences*. New York: Oxford University Press.
- FELDT, L.S. (1958). A comparison of the precision of three experimental designs employing concomitant variable. *Psychometrika*, 23, 335-354.
- FILLENBAUM, S. (1966). Prior deception and subsequent experimental performance: The "faithful" subject. *Journal of Personality and Social Psychology*, 4, 532-537.
- FILLENBAUM, S. eta FREY, R. (1970). More on the "faithful" behavior of suspicious subjects. *Journal of Personality*, 38, 43-51.
- FISHER, R.A. (1925). *Statistical methods for research workers*. Londres: Oliver eta Boyd.
- FLEISS, J.L. (1986). *The design and analysis of clinical experiments*. New York: Wiley.
- FRIEDMAN, N. (1967). *The social nature of psychological research: The psychological experiment as a social interaction*. New York: Basic Books.
- GARCIA, J.F., PASCUAL, J. eta FRIAS, M.D. (1993, Julio). *Determinación del tamaño de la muestra en diseños univariados*. Izenburuean aurkeztutako txostena. III Simposium de Metodología de las Ciencias Sociales y del Comportamiento Santiago de Compostela, Galiza.
- GARCIA, J.F., PASCUAL, J. eta FRIAS, M.D. (1992). *Psicología Experimental: organización y transformación de datos*. Valencia: Nau Llibres.

- GARNER, W.R., HAKE, H.W. eta ERIKSON, C.W. (1956). Operationism and the concept of perception. *Psychological Review*, 63, 149-159.
- GLASS, G.V., PECKHAM, P.D. eta SANDERS, J.R. (1972). Consequences of failure to meet assumptions underlying the analysis of variance and covariance. *Review of Educational Research*, 42, 237-288.
- GORDRAY, D.S. (1989). Optimizing validity in program research: An elaboration of Chen and Rossi's Theory-Driven Approach. *Evaluation and Program Planning*, 12, 379-385.
- GOTTSDANKER, R. (1978). *Experimenting in psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- GRIPP, R.F. eta MAGARO, P.A. (1971). A token economy program evaluation with untreated control ward comparisons. *Behaviour Research and Therapy*, 9, 137-149.
- HAASE, R.F., ELLIS, M.V. eta LADANY, N. (1989). Multiple criteria for evaluating the magnitude of experimental effects. *Journal of Counseling Psychology*, 36, 511-516.
- HARTER, H.L. (1957). Error rates and sample sizes for range tests in multiple comparisons. *Biometrics*, 13, (511-536 or.).
- HARWELL, M.R., RUBINSTEIN, E.N., HAYES, W.S. eta OLDS, C.C. (1992). Summarizing Monte Carlo results in methodological research; The one and two factor fixed effects ANOVA cases. Sepecial Issue: Meta-analysis. *Journal of Educational Statistics*, 17 (4), 315-339.
- HIGGS, W.J. (1970). Effects of gross environmental change upon behavior of schizophrenics: A cautionary note. *Journal of Abnormal Psychology*, 76, 421-422.
- HOCHBERG, Y. eta TAMHANE, A.C. (1987). *Multiple Comparison Procedures*. New York, NY: John Wiley and Sons.
- HOLMES, D.S. (1967). Amount of experience in experiments as a determinant of performance in later experiments. *Journal of Personality and Social Psychology*, 7, 403-407.

- HOLMES, D.S. eta APPLEBAUM, A.S. (1970). Nature of prior experimental experience as a determinant of performance in a subsequent experiment. *Journal of Personality and Social Psychology*, 14, 195-202.
- HOOKE, R. (1983). *How to tell the liars from the statisticians*. New York: Marcel Dekker.
- HUFF, D. (1954). *How to lie with statistics*. New York: Norton.
- HUITEMA, B.E. (1980). *The analysis of covariance and alternatives*. New York: John Wiley and Sons.
- HUYNH, H. eta FELDT, S.F. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *Journal of the American Statistical Association*, 65, 1582-1589.
- JOHNSON, R.A. eta WICHERN, D.W. (1988). *Applied Multivariate Statistical Analysis* (2nd edition). Englewood cliffs, NJ: Prentice-Hall.
- JUDD, C.M. eta KENNY, D.A. (1981). *Estimating the effects of social interventions*. Cambridge, England: Cambridge University Press.
- JUNG, J. (1982). *The experimenter's challenge: Methods and Issues in Psychological Research*. New York: McMillan.
- KAHNEMAN, D. eta TVERSKY, A. (1973). On the Psychology on Prediction. *Psychology Review*, 80, 237-251.
- KAZDIN, A.E. (1988). *Child psychotherapy: Developing and identifying effective treatments*. Elmsford, NY: Pergamon Press.
- KAZDIN, A.E. (1992). *Research Design in Clinical Psychology* (2nd edition). Boston: Allyn eta Bacon.
- KENNY, D.A. eta JUDD, C.A. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, 99, 422-431.
- KEPPEL, G. (1973). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.

- KEPPEL, G. (1982). *Design and analysis. A Researcher's Handbook*. Englewood Cliffs, NJ: Prentice-Hall.
- KERLINGER, F.N. (1986). *Foundations of Behavioral Research* (3rd edition). New York, NY: Holt, Rinehart and Winston.
- KIMBLE, G.A. (1978). *How to use (and misuse) statistics*. Englewood Cliffs, NJ: Prentice-Hall.
- KING, D.W. eta KING, L.A. (1991). Validity Issues in Research on Vietnam Veteran Adjustment. *Psychological Bulletin*, 109 (1), 107-124.
- KIRK, R.E. (1982). *Experimental Desing: Procedures for the Behavioral Sciences* (2nd edition). Belmont, CA: Brooks/Cole.
- KISH, L. (1987). *Statistical Desing for Research*. New York, NY: John Wiley and Sons.
- KLOCKARS, A.J. eta SAX, A. (1985). *Multiple comparisons. Series: Quantitative Applications in the Social Sciences*. 61. liburuki Beverly Hills, CA: Sage Publications.
- KRAEMER, H.C. eta THIEMAN, S. (1987). *How Many Subjects: Statistical Power Analysis in Research*. Beverly Hills, CA: Sage Publications.
- KRATHWOHL, D.R. (1985). *Social and behavioral science research: A new framework for conceptualizing, implementing, and evaluating research studies*. San Francisco: Jossey-Bass.
- KRATOCHWILL, T.R. (1978). *Single-subject Research: Strategies for Evaluating Change*. New York: Academic Press.
- KRUGLANSKI, A.W. (1975). The human subject in the psychology experiment: Fact and artifact. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. 8. liburuki New York: Academic Press.
- KRUGLANSKI, A.W. (1976). On the paradigmatic objections to experimental psychology: a reply to Gadlin and Ingle. *American Psychologist*, 31, 655-663.

- KRUGLANSKI, A.W. eta KROY, M. (1976). Outcome validity in experimental research: A reconceptualization. *Representative Research in Social Psychology*, 7, 166-178.
- KUSCHE, C.A. eta GREENBERG, M.T. (1983). Evaluative understanding and role-taking ability: A comparison of deaf and hearing children. *Child Development*, 54, 141-147.
- LILLIEFORS, H.W. (1967). On the Kolmogorov-Smirnov Test for normality with mean and variance unknown. *Journal of American Statistical Association*, 64, 399-402.
- LINDQUIST, E.F. (1953). *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin.
- LIPSEY, M.W. (1990). *Design Sensitivity. Statistical Power for Experimental Research*. Beverly Hills, CA: Sage Publications.
- LUBIN, A. (1961). The interpretation of significant interaction. *Educational and Psychological Measurement*, 21, 807-817.
- MAHER, B.A. (1978). Stimulus sampling in clinical research: Representative design reviewed. *Journal of Consulting and Clinical Psychology*, 46, 643-647.
- MAHONEY, M.J. (1976). *Scientist as subject: The psychological imperative*. Cambridge, MA: Ballinger.
- MARASCUILO, L.A. eta McSWEENEY, M. (1977). *Nonparametric and distribution-free methods for the Social Sciences*. Monterey, CA: Brooks/Cole Publishing Company.
- MARK, M.M. (1986). Validity Typologies and the Logic and Practice of Quasi-Experimentation. In W.M.K. Trochim (Ed.), *Advances in Quasi-Experimental Design and Analysis* (47-66 or.). San Francisco: Jossey-Bass.
- MARSH, H.W. eta BYRNE, B.M. (1993). Confirmatory Factor Analysis of Multitrait-Multimethod Self-concept Data: Between -group and within-group Invariance Constraints. *Multivariate Behavioral Research*, 28(3), 313-349

- MASLING, J. (1966). Role-related behavior of the subject and psychologist and its effects upon psychological data. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (67-103 or.). Lincoln, NE: University of Nebraska Press.
- MAXWELL, S.E. eta DELANEY, H.D. (1990). *Designing Experiments and Analyzing Data. A Model Comparison Perspective*. California: Brooks/ Cole Publishing Company.
- McCALL, W.A. (1923). *How to experiment in education*. Nueva York: Macmillan.
- McGUIGAN, F.J. (1983). *Experimental Psychology: Methods of Research*. Englewood Cliffs, NJ: Prentice-Hall.
- MENDOZA, J.L. (1980). A significance test for multisample sphericity. *Psychometrika*, 45, 495-498.
- MILLER, R.G. (1966). *Simultaneous statistical inference*. New York: McGraw-Hill.
- MURRAY, L.W. eta DOSSER, D.A. (1987). How significant is a significant difference? Problems with the measurement of the magnitude of effect. *Journal of Counseling Psychology*, 34, 68-72.
- MYERS, A.K. (1980). *Experimental Psychology*. New York, NY: Van Nostrand.
- NEALE, J.M. eta LIEBERT, R.M. (1980). *Science and Behavior: An Introduction to Methods of Research* (2nd edition). Englewood Cliffs, NJ: Prentice-Hall.
- NUNNALLY, J.C. (1978). *Psychometric Theory*. New York, NY: McGraw-Hill.
- O'GRADY, K.E. (1982). Measures of explained variance: Cautions and limitations. *Psychological Bulletin*, 92, 766-777.
- OLSON, J.C. eta PETER, J.P. (1984). External validity?. In P.F. Anderson eta M.J. Ryan (Eds.), *1984 AMA winter Educators' Conference: Scientific method in marketing* (81-84 or.). Chicago: American Marketing Association.
- ORNE, M.T. (1962). On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *American Psychologist*, 17, 776-783.

- ORNE, M.T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal eta R.L. Rosnow (Eds.), *Artifact in behavioral research* (143-179 or.). New York: Academic Press.
- PALUMBO, D.J. eta OLIVERIO, A. (1989). Implementation theory and the theory-driven approach to validity. *Evaluation and Program Planning*, 12, 337-344.
- PASCUAL, J. eta CAMARASA, C. (1991). Diseños: supuestos, potencia y tamaño del efecto. In J. Pascual, M.T. Anguera, G. Vallejo eta F. Salvador (Eds.), *Psicología Experimental* (75-106 or.). Valencia: Nau Llibres.
- PASCUAL, J., GARCIA, J.F. eta FRIAS, M.D. (1994). *Investigación en psicología: Planificación y validez*. Valencia: Ediciones Cristóbal Serrano.
- PATTON, M.Q. (1989). A context and boundaries for a theory-driven approach to validity. *Evaluation and Program Planning*, 12, 375-377.
- PEDHAZUR, E.J. eta PEDHAZUR SCHMELKIN, L. (1991). *Measurement, Design and Analysis. An Integrated Approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- REICHARDT, C.S. (1985). Reinterpreting Seaver's Study of Teacher Expectancies as a Regression Artifact. *Journal of Educational Psychology*, 77, 231-236.
- REICHARDT, C.S. (1986). "Estimating Effects". Unpublished manuscript. Department of Psychology, University of Denver.
- RIBA, M.D. (1990). *Modelo lineal de análisis de la variancia*. Barcelona: Herder.
- ROBINSON, P.W. eta FOSTER, D.F. (1979). *Experimental Psychology: A Small-N Approach*. New York: Harper Row.
- ROETHLISBERGER, E.S. eta DICKSON, W.J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.
- ROSCOE, J.T. (1975). *Fundamental research statistics for the behavioral sciences*. New York: Holt, Rinehart and Winstons (2nd edition).

- ROSENBERG, M.J. (1965). When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology*, 1, 28-42.
- ROSENBERG, M.J. (1969). The conditions and consequences of evaluation apprehension. In R. Rosenthal eta R. L. Rosnow (Eds.), *Artifact in behavioral research* (279-349 or.). New York: Academic Press.
- ROSENTHAL, R. (1963). On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. *American Scientist*, 51, 268-283.
- ROSENTHAL, R. (1966). *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts.
- ROSENTHAL, R. (1969). Interpersonal expectations: Effects of the experimenter's hypothesis. In R. Rosenthal eta R.L. Rosnow (Eds.), *Artifact in behavioral research* (181-277 or.). New York: Academic Press.
- ROSENTHAL, R. (1976). *Experimenter Effects in Behavioral Research*. New York, NY: Irvington.
- ROSENTHAL, R. (1978). How often are our numbers wrong? *American Psychologist*, 33, 1005-1007.
- ROSENTHAL, R. eta ROSNOW, R.L. (Eds.) (1969). *Artifact in behavioral research*. New York: Academic Press.
- ROSENTHAL, R. eta ROSNOW, R.L. (1984). *Essentials of Behavioral Research: Methods and Data Analysis*. New York, NY: McGraw-Hill.
- ROSENTHAL, R. eta RUBIN, D.B. (1978). Interpersonal expectancy effects: the first 345 studies. *The Behavioral and Brain Sciences*, 3, 377-415.
- ROUANET, H. eta LEPINE, D. (1970). Comparison between treatments in a repeated measures design: ANOVA and multivariate methods. *British Journal of Mathematical and Statistical Psychology*, 23, 147-163.
- RYAN, T.A. (1959). Multiple comparisons in psychological research. *Psychological Bulletin*, 56, 26-47.

- SAXE, L. eta FINE, M. (1981). *Social Experiments: Methods for Design and Evaluation*. Beverly Hills, CA: Sage Publications.
- SEAVER, W.B. (1973) Effects of Naturally Induced Teacher Expectancies. *Journal of Personality and Social Psychology*, 28, 333-342.
- SCARIANO, S.M. eta DAVENPORT, J.M. (1987). The effects of violations of independence assumptions in the one-way ANOVA. *The American Statistician*, 41, 123-129.
- SCHEIDT, R.J. (1981). Ecological Valid inquiry: faith accompli? *Human Development*, 24, 225-228.
- SECHREST, L. eta YEATON, W.H. (1981a). Assessing the effectiveness of social programs: Methodological and conceptual issues. In S. Ball (Ed.), *Assessing and interpreting outcomes* (41-55 or.). San Francisco: Jossey-Bass.
- SECHREST, L. eta YEATON, W.H. (1981b). Empirical bases for estimating effect sizes. In R. F. Boruch, P.M. Wortman eta D.S. Cordray (Eds.), *Reanalyzing program evaluations*. San Francisco: Jossey-Bass.
- SECHREST, L. eta YEATON, W.H. (1982). Magnitudes of experimental effects in social science research. *Evaluation Review*, 6, 579-600.
- SHESKIN, D. (1984). *Statistical Tests and Experimental Design: A Guidebook*. New York, NY: Gardner Press.
- SCHMITT, N., COYLE, B.W. eta SAARI, B.B. (1977). A review and critique of analyses of multitrait-multimethod matrices. *Multivariate Behavioral Research*, 12, 447-478.
- SCHMITT, N. eta STULTS, D.M. (1986). Methodology review: Analysis of multitrait multimethod matrices. *Applied Psychological Measurement*, 10, 1-22.
- SIEGEL, S. (1956). *Nonparametric statistics for the Behavioral Sciences*. New York: McGraw-Hill.
- SIEGEL, S. eta CASTELLAN, N.J., jr. (1988). *Nonparametric Statistics for the Behavioral Sciences* (2nd edition). New York, NY: McGraw-Hill.

- SILVERMAN, I. (1977). *The human subject in the psychological laboratory*. New York: Pergamon Press.
- SMITH, P.L. (1982). Measures of variance accounted for: Theory and practice. In G. Keren (Ed.), *Statistical and methodological issues in psychology and social science research* (101-129 or.). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- SNOW, R.E. (1974). Representative and quasi-representative designs for research in teaching. *Review of Educational Research*, 44, 265-291.
- SOLOMON, R.L. (1949). An extension of control group design. *Psychological Bulletin*, 46, 137-150.
- STEVENS, J. (1992). *Applied Multivariate Statistics for the Social Sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.
- STRUBE, M.J. (1988). Some comments on the use of magnitude-of-effect estimates. *Journal of Counseling Psychology*, 35, 342-345.
- TANG, P.C. (1938). The power function of the analysis of variance tests with tables and illustrations of their use. *Statistics Research Memorandum*, 2, 126-149.
- TOOTHAKER, L.M. (1991). *Multiple Comparisons for Researchers*. London: Sage.
- TROCHIM, W.M.K. (1986). *Advances in Quasi-Experimental Design an Analysis*. San Francisco: Jossey-Bass.
- VALLEJO, G. (1986a). Procedimientos simplificados de análisis en los diseños de series temporales interrumpidas: Modelos estáticos. *Revista de Terapia del comportamiento*, 4(2), 114-148.
- VALLEJO, G. (1986b). Procedimientos simplificados de análisis en los diseños de series temporales interrumpidas: Modelos dinámicos. *Revista de Terapia del Comportamiento*, 4 (3), 323-349.
- VALLEJO, G. (1991a). La validez de las investigaciones en el ámbito experimental. In J. Pascual, M.T. Anguera, G. Vallejo eta F. Salvador (Eds.), *Psicología Experimental* (41-74 or.). Valencia: Nau Llibres.

- VALLEJO, G. (1991b). *Análisis univariado y multivariado de los diseños de medidas repetidas de una sola muestra y de muestras divididas*. Barcelona: PPU.
- VONESH, E.F. (1983). Efficiency of repeated measure designs versus completely randomized designs based on multiple comparisons. *Communication in Statistics-Theory and Methods*, 12, 289-302.
- WEBER, S.J. eta COOK, T.D. (1972). Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. *Psychological Bulletin*, 77, 273-295.
- WELCH, B.L. (1951). On the comparison of several mean values: an alternative approach. *Biometrika*, 38, 330-336.
- WILCOX, R.R., CHARLIN, V.L. eta THOMPSON, K.L. (1986). New Monte Carlo results on the robustness of the ANOVA F, W and F* statistics. *Communications in Statistics-Simulation and Computation*, 15, 933-943.
- WILLSON, V.L. (1981). Time and the external validity of experiments. *Evaluation and Program Planning*, 4, 229-238.
- WINER, B.J. (1971). *Statistical Principles in Experimental Design* (2nd edition). New York: McGraw Hill.
- WINER, B.J., BROWN, D.R. eta MICHELS, K.M. (1991) *Statistical Principles in Experimental Design* (3rd edition). New York, NY: McGraw-Hill.
- WITHERSPOON, D. eta WILSON, B.R. (1990). Evaluating research in worksite health promotion. *Health Values, Health Behavior, Education and promotion*, 14 (4), 22-28.
- WUEBBEN, P.L., STRAITS, B.C. eta SCHULMAN, G.I. (Eds.) (1974). *The experiment as a social occasion*. Berkeley, CA: Glendessary Press.
- ZIMNY, G.H. (1961). *Method in Experimental Psychology*. Chicago, IL: Ronald Press.

«Liburu hau Hezkuntza, Unibertsitatea eta
Ikerketa Sailaren laguntzaz argitaratu da»

ISBN 84-86967-75-9



9 788486 196775