

Ebaluazioan egoten diren zehar-iritziak eta honetarako kalitate irizpideak

Juan Carlos Tójar Hurtado

Malagako Unibertsitatea

GAKO-HITZAK: Ebaluazioa. Gizarte-Zientziak. Baliozkotasuna. Neurria. Kualitatiboa.

1. SARRERA

Kalitatezko ebaluazioa lortzeko dagoen kezka garrantzi handiko gaia da Gizarte-Zientzien diziplinan. Gainerako zientzietan ez bezala, ebaluatzailearen paperak garrantzi handiagoa hartzen du, ebaluatzearen ekintza berari hertsiki lotuta baitago. Hau dela eta, lan honen lehen atala, batetik, gizabanakoan oinarritzen da, prozesuaren partaide den heinean eta posizio ezberdinetan eta, bestetik, zehar-iritzi gisa ebaluazioan eragin dezaketen arriskueta. Honela, (1) gizabanakoen eguneroko ebaluazioetako pertzepzio-fenomenorik eraginkorrenak zein dien gogoratuko ditugu (2) Gizabanakoa ikerketa subjektu huts gisa ikusteagatik ebaluazioan sor daitezkeen arriskuak ikusi eta (3) ebaluatzaileari berari eta honen ekintzari eragiten dieten arrisku iturri nagusienak osatuz bukatuko dugu. Lehen zati hau aurrez aipatutako eraginak saihesteari buruzko atal batekin amaitzen da. Atal honek jarraibide eta estrategia jakin batzuk ematen ditu balizko zehar-iritzi nagusienak, kasuan kasu, saihestu, neutralizatu edo murrizteko.

Arriskuak eta arrisku bakoitzean har daitezkeen neurriak aurkeztu eta aztertu ondoren, ebaluazioen baliozkotasunaren azterketan barneratuko gara. Bigarren atal hau hurbilketa klasiko bat eginez hasten da, hartu beharreko neurriei buruz aritu ondoren, ebaluazio prozesu osorako diren kalitate irizpideetan sartzeko. Lehenik eta behin, baliozkotasunaren balio zentralaren ezaugarriak analitikoki definitzen dituen ikuspegi didaktiko bat hartuko da. Ondoren, balioezintasunaren iturri nagusienak aztertuko dira eta, horretarako, lehenengo ataletan aipatutako kontzeptu eta efektuak har-

tuko dira, beti ere hauei kontra egiteko hartu beharreko neurrien jarraibidea ahaztu gabe. Bigarren atal hau ebaluazio kualitatiboen ikuspegitik balioz-kotasunari egindako analisi batekin amaitzen da. Ebaluatzeko modu ezberdin batek ebaluazioaren kalitatea arrazoitzeko modu ezberdin bat eskatzen du. Zentzu honetan ekarpenak dinamikoagoak dira eta garapen bidean daude, beraz, ezinbestekoa da hauek eztabaidara ekarri eta kontuan edukitzea etorkizunean ebaluazioak aurrerapenak izan ditzan.

2. GIZABANAKOEN PERTZEPZIOA

Yolanda eta Pedro seme-alabarik gabeko bikote arrunta zen. Hirian ibili ohi ziren paseatzen eta noizbehinka hiriko erdiguneko kafetegiren batean sartzen ziren. Eserita eta freskagarri bat edaten zuten bitartean, gustukoen zuten denbora-pasan aritzen ziren: Han zehar zebilen pertsonaren bat hautatu eta istorioak asmatzen zituzten. Hasteko, izaeraren ezaugarriak ateratzen zizkieten pertsonak zuen itxuran, keinuetan, ibileran, janzkeran eta abarretan oinarrituta eta pertsonaia guztiz gauzatuta zutenean, honen familia, lagunartea, heziketa, lana eta zaletasunak zer nolakoak liratekeen pentsatzen hasten ziren. Hau guztia egiten zuten, nahiz eta ordurako pertsona jakin hura begien aurrean ez izan, dagoeneko kale izkinan galdu zelako, esate baterako; hala ere, bikote honek keinuak eta egoerak pentsatzen eta ongi arrazoitutako egoera konplexuetan oinarritutako egiazko istorioak asmatzen jarraitzen zuten.

Ez dakigu Yolanda eta Pedrok egindako *inferentzia* eta *interpretazioak* zehatzak ote ziren ala ez. Beharbada asmatuko zuten baten batekin. Dena den, egindako balorazioak udako denbora-pasa huts bat besterik ez ziren.

Askotan, gizakiak besteengan edo hauek bizitzen dituzten egoeretatik sumatzen dituen seinaleekin egiten ditu interpretazio eta inferentziak. Funder eta Colvin (1988) autoreek erakutsi zutenaren arabera, badaude zenbait pertsona minutu gutxi batzuetako behaketarekin eta prestaketa berezirik gabe, ezagutzen ez duten pertsona bati buruzko juzkuak egiteko gai direnak pertsona horrek duen itxuran eta erakusten duen soziazibiltatean oinarrituta.

Egoera edo pertsonetatik hartzen ditugun inpresioak egiazko errepresentazio kognitiboekin lotzen ditugulako hartzen ditugu. Prozesu honetan, eskuragarritasunaren faktorea oso garrantzitsua da. Seinaleak eskura daukagun ezagutzaren arabera interpretatu ohi ditugu, egoera bati buruz dakiguna eta esan digutenaren arabera, esaterako. Trope (1986) autoreak erakutsita-koaren arabera, jendeak argazki bateko aurpegi baten espresio anbigua *egoeraren testuingurutik* dakitenaren arabera interpretatzen du.

Maila ekonomiko altuko ikasleak joaten direlako ospetsua den ikaste-txe batera doan haurra gizarte-maila altukoa dela pentsatuko da. *Korrespondentziazko inferentzia* gisa ezagutzen dena da hau. Modu sistematikoan

egindako inferentzia hauek sortzen dituzte *estereotipoak*: gizonezko guztiak berberak dira, andaluzak, kataluniarrak, galiziarrek honelakoak edo halakoak dira... Estereotipoa pertsona batek talde bati gizabanako baten edo bat baino gehiagoren ezaugarriak egokitu eta talde honetaz egiten duen errepresentazio kognitiboa da. Estereotipoen laguntzaz gizakioi errazago zaigu inguratzen gaituen mundua ulertzea. Dena den, gure inferentzietan akats oso larriak egitera ere eraman gaitzakete. Esate baterako, hiri jakin bateko zerbitzariekin esperientzi txarra izan eta hiri horretako taberna eta jatetxeetako langile guztiak ere desatseginak direla ondoriozta genezake. Gehiegi orokortzean hanka sartzen dugu.

Zientzialariak ez dira ikusten dituzten fenomenoen arrazoiak bilatzen dituzten bakarrak. Edozein pertsonak, zientzialari bat izan gabe, arrazoiak bilatu eta biziutako esperientziei egotzi egiten dizkie. Uste kausalak hainbat bidetatik egiten dira: *asoziazioagatik*, hots, bi errepresentazio kognitibo edo gehiagokoen arteko lotura (enpresaburu askok, esaterako, toxikomano ohi bati lana ematerako garaian zalantza egiten dute, drogen ideia maltzurkeriarekin lotzen baitute); *tradizio kulturalagatik* (adibidez, gure lagun bat judutarra dela jakin eta, hau dela eta, negozioetarako oso abila dela pentsatzea¹); *eskuragarritasunagatik* (Kahneman eta Tversky, 1973), burura datorrigun azalpenik erraz eta azkarrena (pertsona bat gure lagun batekin orain gutxi hitz zakarrak trukatzeko ikusi dugulako pertsona desatsegina dela pentsatzea, esaterako); *garrantziagatik* (trafikoak duen arriskuaz gertuko senideren batek istripu bat izan eta hiltzen denean ohartzen garenean); eta noski, horiek guztiak *nahasteagatik*.

Izan duen jarrera ikusita pertsona bati izaera jakin bateko ezaugarri bat egokitzen badiogu, *korrespondentziako inferentzia* egiten ari garela esaten da, hots, pertsona horren izaerari buruzko errepresentazio kognitiboa egiten ari gara. Gainera, ezaugarri horiek iraunkorrak direla pentsatu ohi da: behin trebetasunez aritu dena normalean ere azkarra da, pertsona bat aurkeztu eta «hanka sartu» zuena sozialki «oso abila ez» dela pentsatuko dugu. Inferentzia mota hauetan honelako akatsen bat egiten denean, *korrespondentzia zehar-iritziez* ari gara. Zentzu honetan, Tversky eta Kahneman (1983) autoreek *adierazgarritasunaren zehar-iritzia* definitu zuten. Izaera jakin baten ezaugarri bat egokitzeaz haratago, ageriko jarrera jakin batetik korrespondentzia zehar-iritzia zabaltzean datza. Adierazgarritasunak lagin baten eta biztanleria, banakako eta kategoria baten arteko, hau da, emaitzaren eta ereduaren arteko korrespondentzia eta antzekotasunari egiten dio erreferentzia.

¹ Uste hau Erdi Arotik dator, garai hartan diruaren kudeaketa judutarrek zuten lan gutxienetako bat baitzen. Gauza berbera gertatu da beste zenbait etniekin, txinatarrekin Indonnesia eta Malasian eta musulmanekin Afrika ekialde eta hegoaldean. Euren gizarteetan ekonomiako bitartekari aritzeak zorrotasun eta zuhurkeriarekin lotutako ezaugarriak hartu dituzte (Pettigrew, 1968).

Zenbait autorek zehar-iritzi hauen iraunkortasuna erakutsi dute. Lehenik hartzen ditugun iritziak, izan positiboak edo negatiboak, nahiko iraunkorrak dira, beraz, egindako inferentziekin bat datozen kasuak hautatzen saiatzen gara. Esate baterako, zerbitzariekin esperientzia desatsegin bat izan eta gero, iritzi honekin bat datozen kasuetan jarriko dugu arreta eta kasu hauek gordeko ditugu gure jakintza erregistroan eta ez oharkabean pasatzen diren aurkako guztiak (Lord *eta lag.*, 1984). Zenbait autore *autorrealizazioaren profeziaz* mintzatzen dira korrespondentziaren betierekotasun honetaz ari direnean (Merton, 1948; Darley eta Fazio, 1980). Adibidez, lehen esperientzia negatibo bat izan eta hiri jakin bateko hiritar guztiak desatseginak diren ustea dugu eta, beraz, hiri horretara goazen hurrengo aldiaren profetia ezartzen dugu: ziur behin baino gehiagotan egoera deseroso batean aurkituko naizela. Horrela, lehen inferentzia horiek betiereko egiten dituzten egoera guztiak sartuko ditugu gure jakintza erregistroan. Kolektibo jakin bateko (lurralde, herrialde edo lan sektore bateko) ezaugarri jakin bati buruzko (normalean ezaugarri negatibo bati buruzko) txisteak egiten direnean funtzionatzen duen mekanismo bera da.

Garrantzitsua ez den informazio gehigarria dugunean, adierazgarritasunaren zehar-iritzia indargabetzen dela ikusi da. *Disolbatze efektua* (Nisbett eta Ross, 1980) da. «Hanka sartu» zueneko pertsonaren adibidean, gertatutako zoritxarreko egoeraz gain informazio osagarria jasotzen baldin badugu (esate baterako bere izena, jatorria edo izaerarekin zerikusirik ez duen beste edozein datu), pertsona hori «sozialki oso abila ez» dela pentsatzeko probabilitatea murriztu egiten da.

Ezaugarri-jarrera harremana (korrespondentziaren zehar-iritzia) beste egoeretara extrapolatzen denan, *usteen oinarrizko akatsaz* ari gara. Adibidez, pertsona bat taberna batean ezezagunen aurrean urduri ikusten baldin badugu, uste hori oso erraz eraman daiteke beste egoera batzuetara. Nisbett eta Ross (1980) autoreen arabera, *usteen oinarrizko akats* hau adierazgarritasunaren eta eskuragarritasunaren heuristikoen ondorioa da eta askotan gertatzen den zerbait da, nahiz eta egoera batetik bestera jarreraren trinkotasuna oso konplexua den (ezaugarria-jarrera-egoera).

Ondorengoak gizabanakoek estimazioetan akatsak egitera bultzatzen dituzten beste pertzepzio fenomenoak dira²:

² Kontuan eduki behar da *efektua* eta *zehar-iritzi* terminoak, bata bestearen ordean erabilzen badira ere, esanahi ezberdinak dituztela. Zehar-iritzia akatsa da. Efekturen baten eraginez errealitateaz urruntzen den fenomenoaren interpretazio oker bati egiten dio erreferentzia. Efektuak, berriz, nahitaez negatiboa izan behar ez duen fenomeno bati egiten dio erreferentzia. Pertzepzio-efektuek, aurrerago deskribatuko diren beste batzuek bezala, hautemandakoa hobeto interpretatzen laguntzen digute, informazioa sinpletzen dute eta, beraz, prozesu kognitiboak errazten dute, nahiz eta batzuetan okerreko interpretazioetara edota akatsak egitera bultzatzen gaituzten. Bigarren zentzu hau da zehar-iritziaren zentzuarekin bat datorrena.

- Testuinguruaren efektua*: ebaluatutako subjektua/objektua aurkitzen den egoerako elementuek egiten den interpretazioan eragiten dute. Honela, estimulu berak ez du egoera ezberdin batean pertzepzio-erantzun bera sortzen. Beste pertzepzio efektu asko bezala, honako hau ere Gestalt Teoriak deskribatu du (ikus Sahakian, 1982). Haur bat ez da modu berean aintzakotzat hartuko puntako eskola batera doanean edota auzo txiro batetik datorrela jakiten denean.
- Halo-efektua*: Lehengo estimulu edota elementuek oraingoan pertzepzioan eragiten dute («forma ematen die»). Testuinguruaren efektuaren antzekoa da baina denborazko sekuentziaren berezitasuna dauka: iraganeko edo lehengo estimuluak, oraingoak eta etorkizunekoak. Subjektu baten aurreko ekintza *on* (edo *txar*) batek orainaldiko ekintza neutro baten pertzepzioan eragin dezake eta aurreko iritzi berera eraman gaitzake.
- Zentratzea*: ikus-eremuko fenomeno edo objekturen batek beste fenomeno edo objektu garrantzitsu batzuen pertzepzioa eragozten duenean gertatzen da. Hori bera gertatzen da, adibidez, gure ikus-eremuaren erdian irakaslearen azalpenei arreta handia jartzen ari den ikaslea dela eta arretarik jartzen ez duten ikasleez ohartzen ez garenean.
- Ainguratzea*: kanpoaldeko gertaera batek, zentrala ez den batek, interes nagusia duen fenomenoaren pertzepzioa eralda dezake. Efektu hau, esaterako, soinu arraro batek eragin dezake (hurbileko toki batean egiten ari diren konponketak, auto asko dituen errepidea,...), honek ebaluatzailearen pertzepzioa desbideratzen du aldi baterako. Hau batzutan ikus-eremuko pertsona baten edo batzuen ezaugarri fisikoak, janzkerak edo bitxi harrigarriren batek... ere eragiten du.
- Maskaramendua*: *Estalketa efektu* moduan ere ezaguna; soinu bat (estalitakoa) aldi bereko beste soinu baten eraginez (estaltzailea) gutxiago entzuten denean gertatzen da. Ainguratze efektua erakusteko lehen erabilitako adibideetan, konponketen edota autoz betetako errepidearen soinuek interesatzen zaigunaren pertzepzioa maskaratu eta aldatzen dutenean eta ezin dugunean inolaz ere entzun, orduan, sortutako efektua maskaramenduarena da eta ez ainguratzearena.
- Kontrastea*: Akats hau estimuluen artean gutxienerako ezberdintasunak gainestimatzeko direnean gertatzen da. Ezberdintasunak berez ez dauden tokian aurkitzen dira. Pentsa dezagun oso langileak diren ikasle talde baten aurrean gaudela. Kasu honetan, ikasle hauen artean, langilea izanda ere, besteak bezain beste nabarmentzen ez den bat baldin badago, hau oso langilea ez dela pentsa daiteke.
- Asimilazioa*: Kontrastearekin egindako akatsa ez bezala, asimilazioa estimuluen artean garrantzitsuak diren ezberdintasunak gainestima-

tzen direnean gertatzen da³. Estimuluen artean ez da nahikoa alboratzen. Hau da gertatzen dena, adibidez, talde berean ikasteko behar ezberdinak dauden subjektuak daudela ikusi eta arreta pertsonalizatuaren beharrik sumatzen ez denean.

—*Pantaila efektua*: Van der Berck (ikus Mitry, 1977) autoreak deskribatu zuen, pantaila bat erdian jarri eta ikusmen estimulu bat desagertzeak ez du esan nahi estimuluak ikuslearentzat hor jarraitzen ez duenik⁴. Hau oso erabilgarria da aldi batean hauteman ezin daitezkeen egoerei jarraipena ematerako garaian (ebaluatutako subjektua ikus-eremutik irten da, subjektuaren eta ebaluatzailearen erdian eragozpenen bat jarri da,...). *Tunel efektua pantaila efektuaren* kasu berezi bat da (mugitzen ari den elementu bat aldi baterako desagertzeak eta ondoren berriz ere agertzeak elementu horrek dirauela pentsarazten digu).

Zientzialariak ez daude zehar-iritzi mota hauetatik libre. Estereotipoek, aurreiritziek eta beste efektu posibleek ere eragiten diete, hauek ere pertsonak baitira. Baina ikertzaile diren neurrian, prest egon behar dute hauen eta bestelako zehar-iritzien aurrean eta berauek saihestu edota beraien eragina ahal den guztia murrizteko ahalegina egin behar dute.

3. EBALUATUTAKO SUBJEKTUEN ZEHAR-IRITZIAK

Metodologia esperimentalak deskribatutako fenomeno baten arabera, esperimentu batean parte hartzen duten subjektuek ikertuak izaten ariko ez balira izango luketen ez bezalako jarrera izan dezakete. Campbell eta Stanley (1966) autoreek azaldu zuten moduan, fenomeno hau prozesu bat neur-

³ Kontraste eta asimilazio efektuen arteko harreman osagarria oso ongi erakusten duten adibide harrigarri batzuk daude. *Saami* hizkuntzan, Laponiako jatorrizko hizkuntzan, elurra izendatzeko 100 modu baino gehiago daude. *Saami* batek iberiar penintsulako eski pistaren bat ikusi eta bi elur-hauts moten artean ezberdintzen baldin badu, penintsulako eskiatzaileen iritziz laponiarraren ezberdinketa gehiegizkoa izango da (kontraste akats bat egiten ariko da). Berriz, penintsulako eskiatzailea Norbegiara joan eta litekeena da honek *saamien* iritziz eurentzat oso ezberdinak diren bi elur motak nahastea (asimilazio akatsa). Malagan bada oso bitxia den ohitura bat ere, guztiek ere asko erabiltzen dutena, kafea izendatzeko ia hamar bat izen erabiltzea da, edalontzi edo kikan duen esne kopuruaren arabera (hauek dira izendatzeko modu ezberdinak kafe kopuru gehienetik gutxienera: «nube», «sombra», «corto», «semi-corto», «mitad», «semi-largo», «largo» edo «cortado», «solo», «crema»). Malagako pertsona batek beste probintzia batean «nube» eta «sombra» baten artean ezberdintzen duenean, gehiegi ezberdintzen duela pentsa daiteke (kontraste efektua). Baina kanpoko batek, berriz, Malagan kafe mota asko «kafesne» esanez nahasten dituenean, zehatzegia ez dela pentsa daiteke (asimilazio efektua)

⁴ Oso erabilia zinema-narrazioan. *Elipsiak*, edo azaltzen ez diren baina hala ere hasiera-ko edo bukaerako planoak ikusi direlako aurrez suposatzen diren eszenek, denbora aurreztu eta erreduzantziak eragozten dituzte narrazioan.

tzeko saiakerarekin batera datozen operazio guztien ondorio zuzena da. Hein batean, froga batean parte hartu duen subjektuak ikerketa baino lehen zena izateari uzten dio (Martin eta Bateson, 1986: 23). Deskribatu berri dugun fenomenoak *erreaktibotasuna* moduan ezagutzen da. Erreaktibotasuna sortzean ebaluaketa bat egitea ez da salbuespen bat.

Erreaktibotasunaz ebaluatzailearen interferentziak direla eta egoeraren naturaltasuna aldatzen denean hitz egin daiteke. Egoera naturala ebaluaketarik ez dagoen egoera bakoitza da eta, beraz, egoera horren edozein aldaketa erreaktibotasun moduan har liteke. Subjektuen jarrera aldatu izanak emaitza ezberdinak dakartza, ondorioz, erreaktibotasunak kolokan jartzen du emaitza horiek bestelako eskenatoki antzekoetara orokortu ahal izatea.

Askotan ez da beharrezkoa ebaluaketan zehar ikertzaile ezezagun bat bertan izatea, nahikoa da, adibidez, testuinguruan ohikoa den pertsonaia batek ebaluatzaile papera hartzea egoera erreaktibo bihurtzeko. Azken gertaera hau bera deskribatu dute pertsona batek bere ezkontidea behatzean (Jarret eta Nelson, 1984) edo baita irakasle batek bere ikaslea behatzean ere (Hay, Nelson eta Hay, 1980).

Erreaktibotasunaren adierazle diren hainbat ezaugarri deskribatu izan dira:

—*Elkarrekiko erreaktibotasuna*: Ebaluatzailearen jarrerak ebaluatutako subjektuaren efektu erreaktiboak jasaten dituenean. Honako hau ebaluatzaileak bere ekintza (nahigabe edo ez) ebaluatutako subjektuaren beraren erreakzioen arabera aldatzen duenean gertatzen da. Elkarrekiko erreaktibotasuna ebaluatzailea behatzailearen aldaketei iragazkorra denean eta subjektuarekin mimetizatzen denean gertatzen da (berarekin batera eta aldi berean urduri, alai edo haserre jartzen da, adibidez), edota jarrera osagarri bat hartzen duenean (urduri dagoela ikusten duelako lasaitzen, triste ikusten duelako alaitzen... saiatzen da).

—*Autoerreaktibotasuna*: Ebaluatutako subjektuak berak bere jarrera ebaluatutako dutela jakitean sortzen da, erantzunak bere buruaz duen itxurari egokitzearen beharretik (Nelson, Liponski eta Black, 1976).

Kontuan eduki behar da ebaluaketa eta erreaktibotasuna ez doazela beti batera. Zehar-iritzi hau ez da, adibidez, *hautematen ez den ebaluaketa* batean azaltzen. Kasu honetan, ez dago erreaktibotasunik ez urruneko nahiz modu ezkutuan gertatzen denean, mikrofono edo kamerak erabiliz (Kazdin, 1980; Patterson, eta Schrest, 1983), ez eta ebaluatutako subjektuak berak egoera artifiziala dela sumatzen duenean (haur txikiak, arazo neurologikoak, gizakiak ez diren espezieak,...). Halaber, ebaluatutako jarrera aldatzeko gutxieneko aukera dagoenean ere, erreaktibotasuna ez gertatzeko aukera dago (Riba, 1991).

Ex post facto egiten den ebaluaketa ere ez da erreaktibo izaten, hots, gertaerak gertatu eta gero gertatzen dena: aztarna fisikoak baloratzetik, era-

biltzearen poderioz objektuak higitzea, fitxategi dokumentalak edota oraingo lekukotasunetik eszenak berreraikitzea (Kazdin, 1980; Webb eta kol., 1981; Patterson eta Sechrest, 1983).

Beste hainbatetan, erreaktibotasunak, nahiz eta egon badagoen, ez da hain garrantzitsua ebaluaketarako. Hau *behaketa parte hartzaile aktiboaren* kasua da. Ikuspuntu hau kontuan edukiz, ebaluatzailearen eta subjektuen arteko interakzioak ez dio prozesuari aplikagarritasuna kentzen, alderantziz, zentzua eman eta aplikagarritasun horren oinarri da, helburua ekintzan parte hartzen duten subjektuen interakziotik sortutako sistema berraren deskribapena lortzea baita.

Ebaluatzailea ez den edozein motatako ikerketetan subjektuen efektuak honako hauek dira:

- Hawthorne*⁵ *efektua*: Ikertu edo ebaluatutako subjektuek (euren ikuspuntutik) ikerketa batean parte hartzen egote hutsagatik ahalik eta hoberen egiteko duten joera deskribatzen du. Hezkuntzaren esparruan Cook (1962) autoreak lehen hezkuntzako ikasleak efektu hauei oso sentibera zirela erakutsi zuen.
- John Henry*⁶ *efektua*: Ikerketa bateko kontrol edo konparaketa talde bateko subjektuek talde esperimentalak baino emaitza hobeak edo antzekoak lortzeko egiten dituzten esfortzuak adierazteko erabiltzen da.
- Placebo efektua*: Hau ere kontrol edo konparaketa talde baten subjektuekin gertatzen da, esku-hartzeren batek modu eraginkorren tratatzen dituela uste dutelako.

4. EBALUATZAILEEN ZEHAR-IRITZIAK

Kelley (1950) autoreak itxaropen efektuak argi eta garbi erakutsi zituen esperientzia bat egin zuen. Ikertzaile honek gonbidatutako irakasle batek psikologia ikastaro batean 20 minutuko eztabaida bat modera zezala eskatu zuen. Moderatzailea azaldu baino lehen eztabaidan parte hartu behar zuten ikasleek informazio ezberdina jaso zuten. Talde bati moderatzailea pertsona «*oso atsegina*, langilea, kritikoa, praktikoa eta ausarta» zela esan zitzaion. Beste taldeak informazio bera jaso zuen baina ñabardura batekin. Pertsona «*nahiko hotza*, langilea, kritikoa, praktikoa eta ausarta» zela esan zitzaion. Eztabaida bukatu ondoren ikasle guztiei moderatzailearen izaerari buruz euren iritzia esatera gonbidatu zitzaion. Pertsona oso atsegina zela esan zitzaion taldeak moderatzailea irakaslea hotza zela esan

⁵ Izen hau Illinois estatuko herri batetik dator, bertan, Elton Mayok Western Electric enpreskao argiztapen eta produktibitate mailei buruzko ikerketa batzuk burutu zituen.

⁶ Izen hau trenbidean lan egiten zuen langile batetik dator, langile hau bideetako lantek makina baino azkarrago jarri nahian hil zen (Borg eta Gall, 1983).

zitzaion taldeak baino xaloagoa, zentzudunagoa, gizatiarragoa, umore hobe zuela, informagoa, lagunartekoagoa eta ospe hobe zuela esan zuen.

Rosenhan (1973) haratago joan eta hainbat ospitaletako psikologo eta gainerako langileak ikerketan zorrozki diseinatu ziren itxaropenei jarraitu zieten *egoera bat sortu zuen*. Ikertzaile honek inolako arazo emozionalik ez zuten pertsona boluntarioak hamabi ospitale psikiatriko ezberdinetan ingresatzea lortu zuen. Ingresatu zituztenean subjektuek existitzen ez ziren ahotsak entzuten zituztela esan zuten, gainerakoan, euren jarrera ahalik eta naturalena izan zen. Euren egoera pertsonala zintzoki kontatu besterik ez zuten egin, hori bai, parte hartzen ari zireneko ikerketari erreferentziarik egin gabe. Esperientziak iraun zuen bi asteetan paziente faltsuetariko bat bera ere ez zuten harrapatu. Ziurrenik, ospitaleetako langileek benetako gaixo mentalak zirela jakintzat hartu zuten eta espero zutena ikusi besterik ez zuten egin.

Akats mota hau dela eta, ebaluatzaileak/ikertzaileak berak espero edo nahi duena ikusi edo aurreratzen du benetan gertatzen ari dena ikusi beharrean. Honi *itxaropen efektua* deritzo. Itxaropenak arrazoi askogatik sor daitezke: aurrez jakiten dena okerrekoa edo osagabea izatea, emaitza jakin batzuk lortzeko nahia; baina ikerketan hainbat alderdirekin ere lotuta zegoela aurkitu da, alderdi hauek, esate baterako, estatus sozioekonomikoa, gizarte maila edo kultur ingurunea (Mucchielli, 1974), aldagai aske jakin baten jarrera alde aurretik jakitea (O'Leary, Kent, eta Kanowitz, 1975), ebaluatzaileei alde aurretik ematen zaien informazioa, Rosenhanen kasuan ikusi dugun moduan (1973), lehen emaitzen eragina (Anguera, 1983) eta zenbaitetan erakargarritasun fisikoa eta aldartea bezalako ezaugarri biopsikosozial soilak⁷.

⁷ Ikerketa askok ebaluatzaileen eta epaileen erabakietan erakargarritasun fisikoak duen eragina frogatu dute. Adibidez, Eagly eta Makhijani (1991) eta baita Feingold (1992) autoreak ere erakargarria den jendea hain erakargarria ez dena baino interesgarriagoa, atsegina-goa, irekiagoa eta sozialki abilagoa dela pentsatzen dela erakutsi dute. Benson eta lag. (1976) erakargarriak diren pertsonen errazago laguntzeko joera dugula frogatu zuten. Cliffordek (1975) egindako esperientzian lehen hezkuntzako irakasleei haurren argazki eta notak erakusten zizkien hauek ebalua zituzten eta egiaztatu zuen irakasleek erakargarriagoak zirenak inteligentagoak zirela eta akademikoki aukera gehiago zituztela uste zutela. Ildo berari jarraituz, Berry eta McArthur (1985) eta baita McArthur eta Berryk (1987) haur tankerako aurpegia zuten gizonen helduak (haur tankerako begi handiak, bekain altuak eta kokots txikia) arpegiera helduegoa zutenak baino zintzoagoak, inuzenteagoak eta atsegina-goa zirela uste zela ikusi zuen. Heilman eta Stopeckek (1985) erakutsi zutenaren arabera, erakargarritasunak itxaropenetan duen eragina ezberdina da generoaren arabera: gizonaren kasuan erakargarritasuna lehiakortasunaren adierazle gisa ikusten zen, emakumezkoaren kasuan, berriz, emaitza alderantzizkoa zen, erakargarria zen emakumeak lehiakortasuna profesionala ez den bideetatik lortu zuela uste zen. Subjektuek erakusten duten aldarreak ere eragin handia du itxaropenetan eta, beraz, baita gainerakoetaz ikusi nahi denan ere. Isen (1987) eta Williamson eta Clarke (1989) erakutsi zuten aldarre oneko dauden pertsonen norberaren eta gainerakoan jarrera aldarre txarreko daudenak baino modu positiboagoan ikusten dituztela. Aldarte onaren eragina aldarre txarrean baino indartsuagoa da norma-

Itxaropenak ebaluatutako subjektuen jarrera ere aldatu dezake, hau da, errektibotasuna sor dezake. Metodologia esperimentalean, saguekin egindako esperientzien arabera, Rosenthalek (1966) ikertzaileek ikerketaren hipotesiarekin bat zetozen emaitzak lortzeko joera zutela sumatu zuen. Rosenthalek pentsatu zuen efektu hau ikertzailearen itxaropenak edo nahiak ikertutako subjektuei nahigabeko pista moduan transmititzen zitzaielako gertatzen zela eta subjektu hauek itxaropen horiekin bat egiteko moduan erantzuten zutela.

Handik gutxira, Rosenthal eta Jacobsonek (1980) irakasleen itxaropenak euren ikasleen jarrera aldatzeko gai zirela ikusi zuten. Zuk nola pentsatu, hala egingo du pertsonak. Fenomeno honi *Rosenthal edo Pigmalion*⁸ *efektua* izena eman zaio.

Kontrako fenomeno *Golem*⁹ *efektua* izenez ezagutzen da (Ball, 1988): irakasleek zenbait ikasleen jarrera akademikoaz itxaropen baxuak baldin badituzte errendimendua ere baxua da.

Barberrek (1976) ikerketa asko eta asko berrikusi eta Rosenthal efektua zalantzan jarri zuen; efektu hau prozeduretan gertatutako akatsetatik eratorritako arazoengatik, okerreko erregistroengatik, datuen azterketa desegokia egiteagatik eta esperimentatzaileak ihes egiteko efektuagatik gertatzen zela esan zuen.

Errektibotasunaz, itxaropenek eta pertsonen pertzepzioak ekar ditzaketen akatsez gain (pertsonen pertzepzioa ere ez da beste bi akatsez libre geratzen), gai honi buruzko literaturak aurrekoek gain beste zenbait akats iturri ere deskribatzen ditu.

Lehen aipatutako akatsez gain, badaude informazioa gaizki prozesatetik datozen akatsak. Esate baterako, *arreta* edo *motibazio akatsak* (ebaluatzaileek ahaztuak gutxi, nekea edo interes gutxi izatea). *Lehentasun* eta *berritasun efektuek* memoria eragiten dituzten akatsak bezalakoak ere deskribatu dira (Murdock, 1962), hauek zerbaitetaz modu diferituan (denbora jakin bat igaro eta gero) ohartzeko garenan dira batez ere garrantzitsuak. Lehentasunak hoberen gogoratzen dena lehenik ikusi duguna dela azaltzen du (lehen irudipenak) eta berritasun efektuak azkenekoz gertatutako gertaerak errazago kontatzen direla dio¹⁰.

lean. Isenek (1984) fenomeno hau animo baxua dutenek gauza atseginetan pentsatzeko duten beharagatik gertatzen dela azaldu zuen.

⁸ Mitologia erromatarren arabera, Pigmalion eskultoreak lan bat egin eta lan horretaz maintemindu zen. Venus jainkosak eskaerari kasu egin eta Galatea bidali zion, estatua itxura berbera zuen emakumea eta berarekin ezkondu zen. Bernar Shawk Pigmalionen mitoa 1913an idatzi zuen kritika sozialeko komedia batekin ezagutarazi zuen. Lan honen argumentuaren haritik musikal bat egin zen, *My fair lady* (1955ean esatreinatua) eta beranduago George Cukor zuzendariak zinema eraman zuen (1964).

⁹ Izen hau mito judutar batetik dator. Mito honen arabera, izaki mekaniko bati kableak gurutzatu eta izaki suntsikor batean bihurtu zen.

¹⁰ Berritasunak ikasle eta beste pertsonen gainerakoan aurrean ikasitakoa erakutsi aurretik azken unean errepasatzeko estrategia erabiltzera bultzatzen die. Era berean, atzizkiaren

Jatorri kognitiboa duten beste zenbait akats informazioaren *kodifikazio* edo *kategorizazio mentala* egiten denean gertatzen dira. Jarrerak gaizki ezberdintzeak asignazio akatsak ekar ditzake. Kategoriatu kopurua handituz doan neurrian, kategoriak ezberdintzeko akatsen kopurua ere handituz doa (Dorsey, Nelson eta Hayes, 1986). Akats batzuk halaberrez gertatzen badira ere (akats aleatorioak), hainbatetan kategorien sistema konplexua erregistroen zehaztasunarekin lotuta dagoela erakutsi da (Mash eta McElwee, 1974).

5. NOLA EKIDIN EBALUAZIOAN EGOTEN DIREN ZEHAR-IRITZI NAGUSIAK

Ebaluazioaren baliozkotasunari egindako mehatxuen kontrolari buruz aurrerago egingo diren gogoetak alde batera utziz, puntu honetan ebaluazioan eragiten duten zehar-iritzi nagusien alderdi batzuei buruz hitz egitea komeni da: subjektuen errektibotasuna eta ebaluatzailearen / ikertzailearen itxaropena. Helburua zehar-iritzi posible hauen eraginak ekidin, murriztu edo ezabatzea litzateke.

Errektibotasunari dagokionez, honako hiru ebaluazio egoera mota hauek ezberdindu daitezke: E1) inolako errektibotasunik *ez dagoen* egoera, E2) errektibotasuna *egon* bai baina ebaluazioaren helburuetarako *garrantzitsua ez* den egoera eta E3) errektibotasuna *egon* eta egoeraren prozesua eta/edo emaitzak *alda* ditzakeen egoera. Ikertzaileak, ebaluazioaren inguruko baldintzen arabera, egoera aldatzea posible den ala ez erabaki ahal izango du.

E1 egoerak dagoeneko komentatu dira errektibotasunari buruzko aurreko atalean. *Hauteman ez den ebaluazioa* da, *ex post facto* ebaluazioa edo gutxieneko aldagarritasuna duten fenomenoaren gogoeta. E2 egoerak ebaluazio *partehartzaile aktiboa* darabiltenak dira edo, gauza bera dena, ikerketa partehartzailearen modu posibleetako bat. Gainerako egoeratan (E3) aukera bakarra prozedura edota baliabide jakin batzuk erabiliz errektibotasuna murrizten saiatzea da. Hau lortzeko prozedura nagusienak honako hauek dira:

—*Ohitzea*: subjektuak ebaluatzailearen presentziara ohitu daitezten itxarotea da. Normalean hau lortzeko modua honakoa da: saio pilotu ugari egiten dira, hauetatik lortutako datuak subjektuen errektibotasunaren adierazpena baloratzeko bakarrik erabiltzen dira eta sumatzen denean edo, hobe, saioetako datuak aztertu ondoren egiaztatzen denean errektibotasuna nahikoa murriztu dela (ikerketaren helburuetarako naturaltasuna neurri handian eraldatzen ez duena), erregis-

efektua ere aipa daiteke (Crowder eta Morton, 1969); honen arabera, garrantzitsua ez den informazio edo esperientziak berritasunaren efektua murriztu dezake.

tratutako informazioari buruzko gogoetari ekiten zaio. Erreaktibotasuna sortzeaz lehenbailehen utzi nahi dion ebaluatzaileak modu iraunkorrean, uniformearean eta jarrerarean aldaketa handirik izan gabe egon behar du, ebaluatutako subjektuak uneoro bera hor dagoela sumatu ez dezan. Alde batetik, ikertzaileak ohitze hau erraz dezakeen plan jakin bat aplikatu dezake modu aktiboan, bestetik, besterik gabe ohitzea halaberrez gertatzea itxaron dezake. (Martin eta Bateson, 1986: 102-103). Zenbaitetan, ahalegin guztiak eginda ere, edo ebaluazio egoera mugagabe luzatzen baldin bada ere, gerta daiteke inoiz ez ohitzea (Kazdin, 1982).

- Prozedura baliabideak*: Bere horretan erabilgarriak izan edota ohitzeari lagun diezaioketen baliabideak, esaterako, subjektuari jasoko den informazioa konfidentziala izango dela eta anonimatu mantenduko dela esatea. Halaber, erreaktibotasuna murrizteko garrantzitsua izaten da subjektuari ebaluatutako izaten direnean naturaltasunez jokatzeko, egia esan dezatela eskatzerako garaian eta konfiantza emateko garaian ebaluatzaileak duen gaitasuna. (Kazdin, 1980; Webb eta kol., 1981). Zentzu honetan, kontuan eduki behar da, gainera, ez dela komeni subjektuei ikerketa hasteko beharrezkoa den informazioa baino gehiago ematea (Fiske, 1978). Gomendio hauekin batera ohitze prozedura errazteko aipatutakoak ere hartu daitezke kontuan (ebaluatzailearen jarrerarean uniformearean edo iraunkorra, gorabehera handirik gabekoa izatea).

Ebaluatzailearen/ikertzailearen itxaropenen aurka borroka egiteko hainbat estrategia posible daude. Nagusienetako bat *ebaluatzaile* nagusiaren eta ebaluazioaren *informazioa biltzen duen* pertsonaren edo pertsonen artean *ezberdintzea* da (behatzaileak, elkarriketatzailak, erregistratzaileak, kodifikatzaileak,...). Estrategia honen bigarren ezaugarria azken hauei ebaluazioaren helburu jakinak eta hipotesiak zein diren ez esatea da (Hobbs eta lag. 1979).

Beste estrategia nagusia *hipotesi irekiak ezartzearekin* dago lotuta. Ebaluatu beharreko fenomenoak deskribatu eta aztertzeke estrategia induktibo bat erabiliz gero (Anguera, 1983: 15-16), teoriari atxikitzeke efektua ekiditen da, beraz, zuzenean eragin daiteke itxaropenen murrizketan.

Ebaluatzailearen eta beste profesionalen rola ezberdintzen lan egiten denean, kodifikatzaileekin dagoen komunikazioa arina izan dadila kontuan edukitzea komeni da. Askotan, *elkar-elikatze informatiboa* erabiltzen baldin bada, zer egiten duen ongi edo zer gaizki esaten baldin bazaio, hau ebaluatzaile nagusiaren itxaropenak transmititzea izan daiteke (O'Leary eta lag., 1975) eta rol bereizketa desagertu daiteke.

Kodifikazio tresna edo sistema arruntak erabiltzeak kategoria bakoitzaren itxaropena murrizten du, bai kopuruan bai eta konplexutasunean ere (Dorsey eta lag., 1986). Erabilitako kodifikazio teknika edo tresna geroz eta konplexuagoa izan, orduan eta handiagoa da ebaluatzailearen itxarope-

nak gertatzeko aukera. Konplexua eta zehatza den eskema edo sistema bat ebaluatzailearen itxaropenen gidoi perfektuan bihurtzen da eta kodifikatzaile edo eremuko langileei ikertzailearen hipotesi edo teorien arabera ebaluatzen dena nondik eta nola interpretatu behar den adierazten die.

Kodifikazio tresna edo sistemaren erabilera errazteko *argibideak* egitea komeni da. Hauetan kategorien definizioak, adibideak eta kontrako adibideak egon behar dira. Azken honen adibide garbia kategoria sistemen eskuliburuak dira (ikus adibidez Herbert eta Attridge, 1975, eta Julian eta lag., 1989 lanak).

Tresna erabiltzeko eskuliburuak egon edo ez, informazioa jaso behar duten pertsonen formazio edo trebakuntzari arreta jartzea komeni da (Bijou eta kol., 1969; Dorsey eta lag., 1986; O'Leary eta lag., 1975). Reidek (1982) esan bezala, eremuko ebaluatzaileak, izpiritu analitiko eta zientifikoa eta motibazio handia izateaz gain, teknikari buruzko trebakuntza jaso beharko du lan egin beharko duen testuinguruaren barnean. Autore honek bere ekintza kontrolatzeko hasierako trebakuntzaren egitea eta beti ere gaitasunak mantentzen direla zaintzea proposatzen zuen¹¹.

6. ZER DA EBALUAZIO ZUZEN BAT¹²

Bildutako informazioak datu kuantitatiboak lortzea uzten baldin badu, normalean erabiltzen diren kalitate irizpideak *fidagarritasuna* eta *erabilgarritasuna* dira. Bi atributu hauek neurriaren ezaugarriak dira eta ez tresneak (Herbert eta Attridge, 1975). Hauek erabiliz fenomeno edo gertaera bati (izan fisikoa, soziala edo psikologikoa) emaitza batzuk lortzeko tresna jakin bat aplikatzen baldin bazaio, erabilitako neurriaren kalitatea zein den jakin daiteke. Orokorrean eta neurriaren testuinguruan fidagarritasunak neurriaren trinkotasunari, iragartzeko erraztasunari eta egonkortasunari egiten dio erreferentzia, erabilgarritasunak, berriz, neurriak neurtu nahi dena zehaztasunez adierazten duen ala ez. Bi ezaugarrien harremanaren inguruan Suen eta Aryk (1989) fidagarritasuna baliozkotasuna egoteko beharrezko atributua zela baina ez nahikoa proposatu zuten eta hau azaltzeko epaiketa bateko lekuko baten deklarazioaren adibidea erabili zuten. Lekukotasun batek fidagarria izan behar du. Gertaerek lekuko batek baino gehiagok hala izan zirela baieztatzen badute dira fidagarriak, baina fidagarritasun horrek ez du bere zehaztasun edo baliozkotasuna ziurtatzen: lekukoak nahita jar zitezkeen ados edota beste faktore batzuk direla eta norabide berera bultzatuta egon litezke (akats sistematikoa).

¹¹ Trebakuntza modu ezberdinak erkatzeko Killian (1982) kontsulta daiteke. Wilson eta Griswold (1985) autoreek esperimendu moduan behatzaileen trebakuntza ulermen maila ezberdinetan manipulatu zuten, maila hauek erregistroen fidagarritasun eta baliozkotasunean zuten eragina aztertzeko.

¹² Hau eta hurrengo atalak Tójar (2001) lanaren arabera egokitu eta eguneratu dira.

Ebaluazio kuantitatibo bat zehatz-mehatz diseinatzeko bi atributu nagusi ezberdindu behar dira baliozkotasunean: *kanpo baliozkotasuna* eta *barne baliozkotasuna*. Lehena datuak orokortzearen aukerarekin dago zuzenki erlazionatuta eta baita bigarrena ere baina kontuan edukiz mendeko aldagaian behatutako efektuak soilik aldagai askeari zor zaizkiola.

Era berean, kanpo baliozkotasuna hiru modu ezberdin hauetara uler daiteke:

- Biztanleriaren baliozkotasuna*, edo ebaluazioan ateratako emaitzak aplikagarriak litzazkiokkeen subjektuen identifikazioa (Bracht eta Glass, 1968).
- Baliozkotasun ekologikoa*, emaitzak beste testuinguru eta ingurumen baldintzetara orokortzea. Erantzun beharreko galdera honakoa litzateke: ze baldintzatan lortuko lirateke ebaluazioan lortutako emaitza berak?
- Konstruktua baliozkotasuna*, edota erabilitako neurriekin osatutako egitura teorikoa ze zehaztasunarekin ebaluatzen den eta nola ebaluatutako arrazoi-efektu harremana beste tokietara eta beste uneetan orokortu daitekeen.

Bestalde, barne baliozkotasuna beste modu askotara uler daiteke:

- Edukiaren baliozkotasuna*, behaketa batek interesatzen zaigun aldagaia domeinu teorikoa zenbateraino islatzen duen adierazten du.
- Baliozkotasun enpirikoa, irizpidearen baliozkotasuna* ere deitzen zaio. Frogan egindakoa irizpide batekin alderatzean psikometrian lortzen dena, hau da, «*frogak neurtu behar duenaren neurri zuzena eta askea*» (Anastasi, 1968: 109). Froga eta irizpidea aldi berean aplikatzen baldin badira, *aldi bereko baliozkotasuna* dela esaten da, berriz, irizpidearen neurria beranduago aplikatzen baldin bada, frogatik eta irizpidetik lortutako emaitzen arteko korrelazioari *aurre baliozkotasuna* deritzo. Irizpidea konplexua denean (adibidez, hainbat aukera edo froga ezberdin dauzkanean) *baliozkotasun sintetikoaz* ari gara (Balma, 1959, in Anastasi, 1968).
- Konstruktua baliozkotasuna, egiturazko baliozkotasuna* ere deitzen zaio, lortutako neurriak objektuak teorian zehaztutako erlazioarekin zenbateraino eta nola erlazionatzen dituen adierazten du (Suen eta Ary, 1989). Psikometrian *egiturazko baliozkotasuna* «*[froga batek] egindako zerbait edo ezaugarri teoriko bat neurtzeko maila*» da (Anastasi, 1968: 119). Baliozkotasun mota hau lortutako emaitzak ezberdindu beharko lituzkeen aldagaiekin erlazionatzen ez direla konturatzean egiaztatzen baldin bada, baliozkotasun alboratzailea deritzo (Campbell eta Fiske, 1959) eta teorikoki erlazionatu beharreko aldagaiekin bat egiten badute, *baliozkotasun konbergentea* deritzo (Campbell, 1969).

7. EBALUAZIOAREN BALIOZKOTASUNA KONTROLATZEA

Ebaluazioan agertzen diren hainbat egoera mehatxuak izan daitezke bai kanpo baliozkotasunarentzat bai eta barne baliozkotasunarentzat ere. Mehatxu hauek ikerketa baliogabe utz ez dezaten, akats iturrien ekintza kon-tuan edukitzea komeni da eta hauek ezabatzen edo kontrolatzen saiatzea.

Kanpo baliozkotasunaren ohiko mehatxuak lagin alboratuak erabiltzea-rekin daude erlazionatuta (kezkarriak, baldin emaitzak orokortu nahi badi-ra), ebaluazio baldintzen artifizialtasunarekin (manipulatua edo simulatua) eta baita aurrez aipatutako efektuen ekintzarekin ere (nagusiki erreaktibota-suna, itxaropena eta hauekin lotuta dauden beste efektu batzuk). Mehatxu hauek ekidin edo murrizteko lagin adierazgarrien (probabilistiko) tekni-kak erabil daitezke, testuinguru naturaletan egindako ikerketak lehenetsi eta aurreko atalean aipatutako ebaluazioan zehar-iritziak ekiditeko ahol-kuak erabili.

Barne baliogabetasunaren iturriak honakoak dira:

- *Historia*. Subjektuen bizitzan gertatutako esperientzia edo gertaerak dira eta ikerketaren koiunturan eragin dezakete.
- *Heldutasuna*. Testuinguruko aldagaien ekintzetan eragin edo haue-kin batera etor daitezkeen prozesu psikofisiologikoak.
- *Tresneria*. Datu bilketarako tresnen arazo teknikoak (zehaztapen fal-ta, fidagarritasunari edo emandako neurrien baliozkotasunari dagoz-kien arazoak, teknikak gutxiegi aplikatzea,...). Tresneriak *sabai* eta *lurzoru efektua* azaltzea bultza dezake. Bi efektu hauek subjektuen artean alboratzeko arazoak sortzen dituzte (aniztasun gutxi), kasu batean subjektuek ez dutelako gutxienero puntuazioa lortzen (froga zailegia da hauentzat) eta banaketa asimetriko positiboa gelditzen delako (lurzoru efektua), beste kasuan, berriz, batzuk gehiengo pun-tuazioa lortzen dutelako (errazegia) eta banaketa asimetriko negati-boa gelditzen delako (sabai efektua).
- *Administrazio errepikatuak*. Froga ondoko emaitzetan sumatutako hobekuntzak froga aurretik ikasitakoagatik izan daiteke (*froga au-rreko efektua*).
- *Bataz besterako erregresioa*. Lehen behaketatan kategoriaren batean muturreko balorazioak jaso dituzten subjektuek ondorengo ebalua-zioetan bataz bestekora hurbiltzeko joera izaten dute¹³.
- *Subjektuen hautapen diferentziala*. Konparatu behar diren taldeak baliokideak ez badira (funtsean aldagai garrantzitsuak berberak izate), emaitzek ezberdintasunei dagokien efektuak izango dituzte.

¹³ Klasikotzat jotzen den ikerketa batean Galtonen ikus zuen oso altuak ziren gurasoak seme-alaba oso altuak izaten zitzutela baina ez beraiek bezain altuak. Gauza bera gertatzen zen oso bajuak ziren guresoekin, seme-alabak ez ziren hain bajuak. Fenomeno hau lagin han-diekin kontrastatu zen eta ordutik aurrera bataz bestekorako erregresioa izena eman zitzaion.

- Hilketa esperimentalak*. Ebaluazioan zehar, dena delako arrazoiengatik parte hartzen ari diren subjektuetatik hainbatek prozesua bertan behera uzten dutenean esaten da.
- Interakzioa*. Baliogabetzeko iturri bi edo gehiagoren arteko interakzioak bata besteari gehitzen joaten diren efektuak sortu eta ebaluazioaren emaitzak kolokan jar ditzaketen efektuak sor daitezke.
- Bestelako efektuak*, aurrez kanpo baliogabetasunean aipatutakoak (latentzia, Pygmalion, John Henry, plazebo, Golem eta Hawthorne) eta beste batzuk (laterala¹⁴, bigarren mailakoa, ondorengoak¹⁵,...) ustez aldagai asketik etorritako emaitzei gehitu zatzazkieke (ikus Ball, 1988).

Baliogabetasunaren iturriak ekiditeko ezar daitekeen kontrolean subjektuak (edo bestelako unitateak) eta ebaluatu behar diren egoerak hautatu behar dira, ebaluazioak zenbat aldiz eta zein une jakinetan egingo diren zehaztu eta erabiliko diren teknika eta ebaluazio tresnak protokolo kontrastatu batek findutakoak eta honen mendekoak izatea saiatu beharra dago. Beraz, aurreko ataletan aipatutako zenbait teknika aplikagarriak dira, hots, ebaluatzailea trebatzearena, teknika konplexuegia ez izatearena eta prozedura eskuliburua egitearena.

8. BALIOZKOTASUN KUALITATIBOA

Azken atal honetan ezinbestekoa da baliozkotasunaren azterketa bat egitea ebaluazio kualitatiboen ikuspegitik. Lan honen sarreran baieztatu zen moduan, ebaluatzeko modu ezberdin batek ebaluazioaren kalitatea arrazoitzeko modu ezberdin bat eskatzen du. Beste ezberdintasunen artean, errealitateaz dugun ideiak eta ebaluatzaileen asmoak bat ez badatoz, ebaluazioek aurrez aipatutako irizpide klasikoak ez diren kalitate irizpide berriak izan beharko dituzte.

Autore batzuek fidagarritasuna eta baliozkotasuna datu kualitatiboetara aplikatzea proposatu baldin badute ere (Kirk eta Miller, 1986), gehien-gehoak ikuspegi aldaketa bat proposatzen du. Informazio kualitatiboen kalitaterrako irizpideak aurkitzerako garaian hezkuntzaren ikerketan ondorio gehien izan dituen proposamena Guba eta Lincoln autoreena izan da (Guba, 1981; Lincoln eta Guba, 1985).

¹⁴ Efektu lateral edo kolateralak (*side effects*) ustekabeen sortzen direnak dira, ebaluazioaren helburuetatik albo batera daude eta zeharka mendeko aladagaiari eragin diezkieke positibo edo negatiboki. Bigarren mailako efektuak efektu klateral mota berezi bat dira (normalean negatiboak)

¹⁵ Ondorengo efektua edo lo efektuak (*after edo sleeper*) esku-hartze batetako emaitza beranduei egiten dio erreferentzia, hobekuntzak esku-hartzea bukatu eta ebaluatu baino askoz geroago sortzen dira.

Autore hauen arabera, ikerketa naturalistarako irizpide nagusia «egiazkotasuna» (*trustworthiness*) da. Egiazkotasun edo zehaztasun hau informazio kualitatiboetan erabil daiteke, «egiazkotasun kuantitatibo» zentzuaren maila berean. Irizpide hau honako irizpide hauetan bereiz daiteke:

- a) *Sinesgarritasuna*. Ezaugarri honek ez du esan nahi aurkikuntzak sinesgarriak edo espero zirenak izan behar dutenik. *Gardentasuna* da beste edozer baino garrantzitsuagoa. Honek esan nahi du ebaluatzailearen interpretazioen eta gertaeren arteko kontrasteak interpretazio horiek sinesgarri eta onargarri egin behar dituela ahalik eta modurik gardenenean.
- b) *Transferigarritasuna*. *Aplikagarritasun* ideiaaren antzekoa. Ez dio inolaz ere orokortzeari erreferentzia egiten. Kontua ez da emaitzak orokortzea, baizik eta lan hipotesiak formulatzea, testuingururako garrantzitsuak eta beste antzeko egoerei transferigarri edo aplikagarri zaizkien lan hipotesiak.
- c) *Mendekotasuna*. Neurri batean aurrekoarekin lotuta dagoen ideia da. Mendekotasunak emaitza jakin batzuen *tinkotasunari* egiten dio erreferentzia, nolabaiteko ezegonkortasuna eragiten duten aldaketa instrumentalak direla eta egon daitezkeen aldaketa txiki batzuei izan ezik. Gainjartzen edo osatu egiten diren bilketa metodoak erabiliz egonkorak azaltzen diren elementu jakin batzuk egiaztatu daitezke.
- d) *Berresgarritasuna*. Berresgarritasuna ebaluatzailearen aurrejuzku edo itxaropenak alde batera uzten direnean sortzen da; ebaluatzaileak *neutraltasunez* jokatzeko du eta emaitzak askeak izaten dira.

Sailkapen hau zalantzan jarri du ikuspegi humanistiko-intrepretatiboak, aurreko irizpideak eraikitzeko logika ikuspegi enpriko-analitikoarena baita¹⁶. Kritika hauek onartu eta Guba eta Lincoln (1990; eta baita Lincoln, 1995) autoreek orientazio positiboaren irizpideak inoiz ikertu edo ezagutu ez zituzten ikertzaile edo ebaluatzaileek onar zitzaizketen kontzeptu berriak proposatu zituzten. Irizpide berri hauek logika konstruktibistatik datoz eta *egiazkotasunaren* kontzeptu zentralak antolatzen ditu:

- Inpartzialtasuna* (*fairness*), informazio berari buruzko ikuspegi ezberdinak aurkeztu ahal izateko.
- Egiazkotasun ontologikoa*, errespetua heldutasunean, egitura emikoak, pertsonak dituzten ikuskerak egitea eta ebaluatzea.
- Hezkuntza-egiazkotasuna*, norberak egindakotik aurrerago joan eta egindakoa besteekin kontrastatzeko aukera, honela balore sistema

¹⁶ Guba eta Lincoln (1994: 114) autoreek irizpide kualitatibo eta kuantitatiboen arteko paralelismoak esplizituki onartu zituzten: zehazki barne sinesgarritasun eta baliozkotasunaren artean, kanpo transferigarritasunaren eta baliozkotasunaren artean, mendekotasunaren eta fidagarritasunaren artean eta berresgarritasunaren eta objektibotasunaren artean.

ezberdinek irtenbide ezberdinak proposatzen dituztela azpimarratzeko.

- *Egiazkotasun katalizatua*, hezkuntza-ebaluazio prozesuek ekintza bat zein mailatan bultzatu eta arintzen duten adierazten du.
- *Egiazkotasun taktikoa*, ekintzara inplikaturako pertsonen potentzial eta trebakuntza (*empowerment*) gisa.

Ikerketa kualitatiboetako informazioari egiazkotasuna emateko prozedurarik erabiliena *triangulaketa* da oraindik orain. Denzinen hitzak erabiliz, triangulaketa fenomeno beraren azterketa-metodologia ezberdinak konbinatzea da (Denzin eta Lincoln, 1994). Kemmisen (1983) zehaztu zuen moduan, hau datu iturri ezberdinak erabiliz lortu daiteke (pertsonak, tresnak, dokumentuak edo guztiak batera). Baina azken aldi honetan metodo hau oso kritikatu izaten ari da, batez ere parte hartzaileen baliozkotasunari dagokionean (Bloor, 1999; Cutcliffe eta McKenna, 1996; Swanborn, 1998).

Teoria kritikotik ere zehaztu dira erabaki eta aldaketan dimentsio bereziak nahiz metodologia dialektikoa eta erreflexiboa kontuan hartzen dituzten ebaluazioak balioztatzeke irizpide espezifikoak zein diren.

Watermanek (1988) ikerketa-ekintza prozesuetan aplikagarri diren hiru baliozkotasun motaz hitz egiten du: *baliozkotasun dialektikoa*, *kritikoa* eta *gogoetazkoa*. *Baliozkotasun dialektikoak* praktikaren, ikerketaren eta teoriaren artean sortzen diren tentsioei egiten die erreferentzia. *Baliozkotasun kritikoak* ardura moralari, inplikazio etikoei eta ekintza eta teoriaren ondorioei egiten die erreferentzia. *Gogoetazko baliozkotasunak* ikertzailearen ikuskerak modu esplizituan aztertzea, onartzea eta forman jartzea dakar.

Erreflexibilitatearen kontzeptua eztabaidatzen zuten dagoeneko tradizio fenomenologikoan (Dukes, 1984) eta etnografikoan (Hammersley, 1992; Marcus, 1994). Hammersleyek (1992:64) txostenaren erreflexibilitateaz hitz egiten zuen, hau da, ikertzailearen efektuak, emaitzen informazioa biltzeko teknikak eta burututako prozesuari buruz interesatuei ematen zaien informazioa ere zein mailatan ebaluatzen den.

Ikuspegi kritiko edo soziokritikoei aplikagarri zaion irizpide hauekin batera, egindako ikerketaren *teoria maila* legoke (Morse, 1994). Azkenik, badirudi erabakiaren eta aldaketaren alde egiten duen ikuspegitik ezinbestekoa dela sortutako *transformazioa* baliozkotasun irizpide gisa hartzea eta *aldaketa prozesuaren intzidentzia eta eragin mailaren adierazle gisa kristalizatzea* (Sandín, 2000: 239).

Ikus daitekeen moduan, ikuspegi kualitatibotik ekarpenak dinamikoak izaten ari dira eta garapen bidean daude. Eztabaidara ekarri eta ebaluazio ez kualitatiboen kasuetan ere kontuan edukitzea (ikus Tójar eta Mena, 2004) oso aberasgarria izango litzateke etorkizunean ebaluazioa bultzatzeko.

Jasotze-data: 2004-02-02

Onartze-data: 2004-12-20

The concern about a quality assessment in Social Sciences is a subject of vital importance in teaching. As opposed to what happens in other sciences, the role of the assessor is more important in that it is inherently united to the action of assessment itself. For this reason the first part of this study has been dedicated to the individual as an integral part of the process, in different positions, and to the risks that, by means of possible biases the assessment may be affected. Once the risks have been presented and analysed, together with the measures which can be adopted in each case, a full analysis of the validation of the assessments is begun. This second part starts with a classical approximation starting from the properties of the measure, to from there extend the criteria of quality over the whole process of assessment. The study ends with an analysis of validity from the perspective of qualitative assessments. A different assessment method demands a different way of explaining the quality of the assessment. From this perspective, the contributions are more dynamic and are currently being developed more fully, for which reason it is indispensable to bring them to the debate and to take them into consideration for the advance of assessment in the future.

Keywords: Evaluation. Social sciences. Validity. Measurement. Qualitative.

La preocupación por una evaluación de calidad en Ciencias Sociales es un tema de vital importancia disciplinar. A diferencia de lo que ocurre en otras ciencias, el papel del evaluador toma mayor protagonismo al estar inherentemente unido a la propia acción evaluadora. Por esta razón, se ha dedicado la primera parte de este trabajo al individuo como parte integrante del proceso, en diversas posiciones, y a los riesgos que, en forma de posibles sesgos, pueden afectar a la evaluación. Una vez presentados y analizados los riesgos, junto con las medidas que pueden en cada caso adoptarse, se entra de lleno en el análisis de la validez de las evaluaciones. Esta segunda parte comienza con una aproximación clásica que parte de las propiedades de la medida para, a partir de ahí, extender los criterios de calidad sobre todo el proceso de evaluación. El trabajo termina con un análisis de la validez desde la perspectiva de las evaluaciones cualitativas. Una forma de evaluar diferente exige un modo distinto de razonar la calidad de la evaluación. Desde esta perspectiva, las aportaciones son más dinámicas y se encuentran en pleno desarrollo actual, por lo que resulta indispensable aportarlas al debate y tenerlas en consideración para el avance evaluador futuro.

Palabras clave: Evaluación. Ciencias sociales. Validez. Medida. Cualitativa.

L'inquiétude pour une évaluation de qualité en sciences sociales est un thème d'importance capitale pour la discipline. Contrairement à ce qui se passe dans d'autres sciences, le rôle de l'évaluateur devient plus important puisqu'il est intrinsèquement uni à l'action d'évaluation elle-même. C'est pour cette raison que la première partie de ce travail est consacrée à l'individu en tant que partie intégrante du processus, en différentes positions, et aux risques qui, sous formes d'éventuels travers, peuvent affecter l'évaluation. Une fois que les risques ont été présentés et analysés, ainsi que les mesures pouvant être adoptées dans chaque cas, on aborde en plein l'analyse de la validité des évaluations. Cette deuxième partie commence par une approche classique qui part des propriétés de la mesure pour, à partir de là, étendre les critères de qualité à tout le processus d'évaluation. Le travail se termine par une analyse de la validité de la perspective des évaluations qualitatives. Une autre forme d'évaluation exigerait un autre raisonnement sur la qualité de l'évaluation. Dans cette perspective, les contributions sont plus dynamiques et se trouvent en plein développement actuellement, et pour cette raison il est indispensable de les intégrer au débat et de les tenir en compte pour les futurs progrès de l'évaluation.

Mots clé : Évaluation. Sciences sociales. Validité. Mesure. Qualitative.

BIBLIOGRAFIA

- ANASTASI, A. (1968): *Tests psicológicos*. Madrid: Aguilar, 1982.
- ANGUERA, M. T. (1983): *Manual de prácticas de la observación*. México: Trillas.
- BALL, S. (1988): «Unintended effects in educational research». En J. P. KEEVES (ed.). *Educational Research, Methodology & Measurement. An International Handbook*. (pp. 490-493). Oxford: Pergamon.
- BARBER, T. X. (1976): *Pitfalls in human research. Ten pivotal points*. Nueva York: Pergamon Press.
- BENSON, P. L., KARABENIC, S. A. eta LERNER, R. A. (1976): «Pretty pleases: The effects of physical attractiveness on race, sex and receiving help». *Journal of Experimental Social Psychology*, 12, 409-415.
- BERRY, D. S. eta MCARTHUR, L. Z. (1985): «Some components and consequences of babyface». *Journal of Personality and Social Psychology*, 48, 312-323.
- BIJOU, S. W., PETERSON, R. F., HARRIS, F. R., ALLEN, K. E. eta JONHSTON, M. S. (1969): «Methodology for experimental studies of young children in natural settings». *The Psychological Record*, 19, 177-210.
- BLOOR, M. (1999): «On the analysis of observational data: A discussion of the worth and uses of inductive techniques and respondent validation». En A. BRYMAN eta R. G. BURGESS (Eds.). *Qualitative Research. Volume III*. (pp. 445-454). Londres: Sage.
- BORG, W. R. eta GALL, M. D. (1983): *Educational Research: A practical Guide for Teachers*. Nueva York: Logman.
- BRACHT, G. eta GLASS, G. V. (1968): «The external validity of experiments». *American Educational Research Journal*, 5, 437-474.

- CAMPBELL, D. T. (1960): «Recommendations for APA test standards regarding construct, trait, and validity». *American Psychologist*, 15, 546-553.
- CAMPBELL, D. T. eta FISKE, D. W. (1959): «Convergent and discriminant validation by multitrait-multimethod matrix». *Psychological Bulletin*, 15, 546-553.
- CAMPBELL, D. T. eta STANLEY, J. C. (1963): *Diseños experimentales y cuasi experimentales en la investigación social*. Mexico: Amorrortu, 1966.
- CLIFFORD, M. M. (1975): «Physical attractiveness and academic performance». *Child Study Journal*, 5, 201-209.
- COOK, D. L. (1962): «The Hawthorne effect in educational research». *Phi Delta Kappan*, 44, 116-122.
- CROWDER, R. G. eta MORTON, J. (1969): «Precategorical acoustic storage». *Perception and Psychophysics*, 5, 365-373.
- CUTCLIFFE, J. R. eta MCKENNA, H. P. (1999): «Establishing the credibility of qualitative research findings: the plot thickens». *Journal of Advanced Nursing*, 30 (2), 374-380.
- DARLEY, J. M. eta FAZIO, R. H. (1980): «Expectancy confirmation processes arising in the social interaction sequence». *American Psychologist*, 35, 867-881.
- DENZIN, N. K. eta LINCOLN, Y. S. (Eds.). (1994): *Handbook of qualitative research*. Londres: Sage.
- DORSEY, B. L., NELSON, R. O. eta HAYES, S. C. (1986): «The effects of code complexity and of behavior frequency on observer accuracy and interobserver agreement». *Behavioral Assessment*, 8 (4), 349-363.
- DUKES, S. (1984): «Phenomenological methodology in the human sciences». *Journal of Religion and Health*, 23 (3), 197-203.
- EAGLY, A. H. eta MAKHIJANI, M. G. (1991): «What is beautiful y good, but...: A metha-analytic review of research on the physical attractiveness stereotype». *Psychological Bulletin*, 110, 109-128.
- FEINGOLD, A. (1992): «Good-looking people are not what we think». *Psychological Bulletin*, 111, 304-341.
- FISKE, D. W. (1978): *Strategies for personality research*. San Francisco: Jossey-Bass.
- FUNDER, D. C. eta COLVIN, C. R. (1988): «Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgments». *Journal of Personality and Social Psychology*, 55, 149-158.
- GUBA, E. (1981): «Criteria for assessing the trustworthiness of naturalistic inquiry». *Educational communication and technology journal*, 29, 79-82.
- GUBA, E. L. eta LINCOLN, I. S. (1990): *Fourth generation evaluation*. (2.^a ed.). Londres: Sage.
- GUBA, E. L. eta LINCOLN, I. S. (1994): «Competing paradigms in qualitative research». En N. K. DENZIN eta Y. S. LINCOLN (eds.). *Handbook of qualitative research*. (pp. 105-117). Londres: Sage.
- HAMMERSLEY, M. (1992): *What's wrong with ethnography?* Londres: Sage.
- HAY, L. R., NELSON, R. C. eta HAY, W. M. (1980): «Methodological problems in the use of participant observer». *Journal of Applied Behavior Assessment*, 13, 501-504.
- HEILMAN, M. E. eta STOPECK, M. H. (1985): «Attractiveness and corporate success: Different casual attributions for males and females». *Journal of Applied Psychology*, 70, 379-388.

- HERBERT, J. eta ATTRIDGE, C. A. (1975): «A guide for developers and users of observation systems and manuals». *American Educational Research Journal*, 12, 1-20.
- HOBBS, S. A., WALLE, D. L. eta HAMMERSLY, G. A. (1979): «Effects of expectancy of outcome on the reactivity of self-monitoring». *Journal of Behavioral Assessment*, 1 (4), 282-288.
- ISEN, A. M. (1984): «Toward understanding the role of affect in cognition». En R. S. WYER eta T. K. SRULL (eds.), *Handbook of social cognition*. Vol. 3. (pp. 203-236). Hillsdale, NJ: Erlbaum.
- ISEN, A. M. (1987): «Positive affect, cognitive processes, and social behavior». En I. BERKOWITZ (ed.), *Advances in experimental social psychology*. Vol. 20 (pp. 203-253). New York: Academic Press.
- JARRET, R. B. eta NELSON, R. O. (1984): «Reactivity and unreability of husbands as participants observers». *Journal of Behavior Assessment*, 1 (4), 131-145.
- JULIEN, D. MARKMAN, H. J. eta LINDAHL, K. M. (1989): «A comparison of a global and microanalytic coding system: implications for future trends in studying interactions». *Behavioral Assessment*, 11, 81-100.
- KAHNEMAN, D. eta TVERSKY, A. (1973): «On the psychology of prediction». *Psychological Review*, 80, 237-251.
- KAZDIN, A. E. (1980): *Research design in clinical psychology*. Nueva York: Harper & Row.
- KAZDIN, A. E. (1982): «Observer effects: reactivity of direct observation». En D. P. HARTMAN (ed.). *Using observers to study behavior* (pp. 5-20). San Francisco: Jossey Bass.
- KELLEY, H. H. (1950): «The warm-cold variable in first impressions of persons». *Journal of Personality*, 18, 431-439.
- KEMMIS, S. (1983): «Action research». En T. HUSEN eta T. N. POSTLETHWAITE (eds.). *International Encyclopaedia of Education: Research and Studies*. Vol. 1. Oxford: Pergamon, 35-42.
- KILLIAN, J. E. (1982): «A Comparison of training methods for systematic observation». *Journal of Classroom Interaction*, 18 (1), 19-22.
- KIRK, J. eta MILLER, M. L. (1986): *Reliability and validity in qualitative research*. Londres: Sage.
- LINCOLN, Y. S. (1995): «Emerging criteria for quality in qualitative and interpretative research». *Qualitative Inquiry*, 1 (3), 275-289.
- LINCOLN, Y. S. eta GUBA, E. G. (1985): *Naturalistic inquiry*. Beverly Hills, Ca: Sage Publications.
- LORD, C. G., LEPPER, M. R. eta MACKIE, D. (1984): «Attitude prototypes as determinants of attitude –behavior consistency». *Journal of Personality and Social Psychology*, 46, 1254-1266.
- MARCUS, G. E. (1994): «What comes (just) after “post”? The case of ethnography». En N. K. DENZIN eta Y. S. LINCOLN (eds.). *Handbook of qualitative research*. (pp. 436-477). Londres: Sage.
- MARTIN, P. eta BATESON, P. (1986): *Measuring Behavior: An introductory guide*. Londres: Cambridge University Press.
- MASH, E. J. eta McELWEE, J. D. (1974): «Situational effects of observer’s accuracy: behavioural predictability, prior experience, and complexity of coding categories». *Child development*, 45, 367-377.

- MCARTHUR, L. Z. eta BERRY, D. S. (1987): «Cross-cultural agreement in perceptions of babyfaced adults». *Journal of Cross Cultural Psychology*, 18, 165-192.
- MERTON, R. (1948): «The self-fulfilling prophecy». *Antioch Review*, 8, 193-210.
- MITRY, J. (1977): «Psicología del cine». En N. SCHUMANN (dir.). *Enciclopedia de la psicología y la pedagogía. Vol. IV* (pp. 288-306). Madrid: Sedmay.
- MORSE, J. (1994): *Critical issues in qualitative research methods*. Londres: Sage.
- MUCCHIELLI, R. (1974): *L'observation psychologique et psychosociologique*. París: Libraires Techniques, ESF.
- MURDOCK, B. B. (1962): «The serial position effect of free call». *Journal of Experimental Psychology*, 64, 482-488.
- NELSON, R. O., LIPINSKY, D. P. eta BLACK, J. L. (1976): «The relative reactivity of external observation and self-monitoring». *Behavior Therapy*, 7, 314-321.
- NISBETT, R. E. eta ROSS, L. (1980): *Human inference. Strategies and shortcomings of social judgment*. Englewood Cliffs, New Jersey: Prentice Hall.
- O'LEARY, K. D., KENT, R. N. eta KANOWITZ, J. (1975): «Shaping data collection congruent with experimental hypothesis». *Journal of Applied Behavioral Analysis*, 8, 43-51.
- PATTERSON, D. R. eta SCHECREST, L. (1983): «Nonreactive measures in Psychotherapy outcome research». *Clinical Psychological Review*, 3, 391-416.
- PETTIGREW, T. F. (1968): «Race relations: Social and psychological aspects». En D. L. SKILL (ed.). *The international encyclopedia of the social sciences*. Vol. 13. (pp. 277-282). Nueva York: Mcmillan Company.
- REID, J. B. (1982): «Observer training in naturalistic research». En D. P. HARTMANN (ed.). *Using observers to study behavior*. (pp. 37-50). San Francisco: Jossey Bass.
- RIBA, C. (1991): «El método observacional. Decisiones básicas y objetivos». En M. T. ANGUERA (ed.). *Metodología observacional en la investigación psicológica. Vol. I. Fundamentación (1)*. (pp. 29-114). Barcelona: PPU.
- ROSENHAN, D. L. (1973): On being sane in insane places. *Science*, 179, 250-258.
- ROSENTHAL, R. (1966): *Experimenter effects in behavioral research*. Nueva York: Appleton.
- ROSENTHAL, R. eta JAKOBSON, L. (1980): *Pygmalion en la escuela*. Madrid: Marova.
- SAHAKIAN, W. S. (1982): *Historia y sistemas de la psicología*. Madrid: Tecnos.
- SANDÍN, M. P. (2000): «Criterios de validez en la investigación cualitativa: de la objetividad a la solidaridad». *Revista de Investigación Educativa*, 18 (1), 223-242.
- SUEN, H. K. eta ARY, D. (1989): *Analyzing quantitative behavioral observation data*. Hillsdale, N. J.: Erlbaum.
- SWANBORN, P. G. (1998): «Embracing ambiguities and valuing ourselves: issues of validity in action research». *Journal of Advancing Nursing*, 28 (1), 101-105.
- TÓJAR, J. C. (2001): *Planificar la investigación educativa. Una propuesta integrada*. Buenos Aires: Fundec / Paradox.
- TÓJAR, J. C. eta MENA, E. (2004): «Qué pueden aportar los criterios de calidad cualitativos a la investigación cuantitativa». *Metodología de las Ciencias del Comportamiento, vol. especial*.
- TROPE, Y. (1986): «Identification and inferential processes in dispositional attribution». *Psychological Review*, 93, 239-257.

- TVERSKY, A. eta KAHNEMAN, O. (1983): «Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment». *Psychological Review*, 90 (4), 293-315.
- WATERMAN, H. (1998): «Embracing ambiguities and valuing ourselves: issues of validity in action research». *Journal of Advancing Nursing*, 28 (1), 101-105.
- WEBB, E. J., CAMPBELL, D. T., SCHWARTZ, R. D., SECHREST, L. eta GROVE, J. B. (1981): *Nonreactive measures in the social Sciences*. Boston: Houghton and Mifflin.
- WILSON, F. R. eta GRISWOLD, M. L. (1985): «The effects of method and comprehensiveness of training on the reliability and validity ratings of counsellor empathy». *Measurement and Evaluation in Counselling and Development*, 18 (1), 3-11.
- WILLIAMSON, G. M. eta CLARK, M. S. (1989): «Providing help and desired relationship type as determinants of changes in moods and self-evaluations». *Journal of Personality and Social Psychology*, 56, 722-734.