

## Inferența causală

### Cuprins

1	Strategii de cercetare .....	1
2	Inferența causală.....	2
2.1	Cauza, efectul și relația causală.....	2
2.2	Condițiile inferenței causale în tradiția lui Donald Campbell.....	4
2.3	Modelul causal al lui Rubin (MCR).....	7
3	Experimentul .....	9
3.1	Experimentul adevărat.....	9
3.2	Quasi-experimentul .....	10
3.3	Avantaje și dezavantaje ale experimentelor și quasi-experimentelor.....	10
4	Întrebări recapitulative.....	12
5	Exercițiu .....	12
6	Referințe bibliografice.....	12

### 1 Strategii de cercetare

Cercetarea științifică se orientează după una din cele două paradigme majore: calitativă și cantitativă. Referindu-ne doar la abordarea cantitativă, aceasta poate urma diferite strategii care, la rândul lor, se concretizează în diverse modele de cercetare. În ceea ce privește strategiile cantitative, Muchinsky (2006) le clasifică în două categorii:

- *Metode de cercetare primare*, care produc informații directe în legătură cu subiectul cercetării, fiind mai frecvent întâlnite în practică. Din această categorie ar face parte experimentul, quasi-experimentul și chestionarele.
- *Metode de cercetare secundare*, care analizează date rezultate în urma mai multor studii pe un anumit subiect sau temă de cercetare, de pe urma cărora extrag concluzii, utilizând diverse tehnici de sintetizare și integrare a datelor.

La rândul lui, Spata (2003), clasifică strategiile de cercetare în două mari categorii: experimentul și strategiile descriptive. Mai departe, în categoria metodelor descriptive include studiile non-experimentale, observația naturală, sondajul și studiile corelaționale. Wampold (2006) face distincție între modelele de cercetare experimentale, quasi-experimentale și ”pasive”. În această ultimă categorie intră studiile bazate pe modele corelaționale și comparațiile dintre grupuri independente (formate din subiecți diferiți).

Un sistem mai elaborat de clasificare este cel propus de Montero și León (2007), care identifică următoarele tipuri de cercetări psihologice:

#### a) Studii teoretice

- *Clasice* (trecerea în revistă a cercetărilor anterioare, fără analiza statistică a rezultatelor lor);
- *Meta-analiză* (analiza studiilor pe un anumit subiect, cu utilizarea unor proceduri de estimare a mărimii efectului, în vederea evidențierii suportului pentru ipotezele testate);

#### b) Studii empirice bazate pe metodologia cantitativă

- *Descriptive, bazate pe observație sistematică*, prin utilizarea unui cod convențional, în mediul natural sau artificial;
- *Descriptive, bazate pe sondaj și eșantionare probabilistică* (diferite tipuri de cercetări bazate de regulă pe chestionare);
- *Experimente* (manipularea variabilei independente, plus alte condiții specifice);
- *Experimente pe subiecți unici* (în cazuri speciale);

- *Quasi-experimente* (manipularea variabilei independente, dar fără respectarea altor condiții specifice experimentului);
  - Studii *ex post facto* (variabila independentă nu este manipulată, ci doar măsurată);
  - *Studii instrumentale* (dezvoltarea unor noi proceduri de intervenție, aparate, instrumente sau teste psihologice, care necesită evaluarea proprietăților psihometrice; adaptarea tuturor acestora în alte medii culturale);
- c) Studii empirice bazate pe metodologia calitativă
- *Metoda etnografică* (studiul grupurilor umane în viața cărora cercetătorul se implică direct);
  - *Metoda studiului de caz* (analiza descriptivă a unui element unic, care poate fi o persoană, un grup, o organizație etc.);
  - *Metoda acțiune-cercetare* (un context social uman în care intervenția și cercetarea se asociază într-o secvență succesivă, până la rezolvarea problemei care a impus acest demers);

Dacă analizăm toate clasificările propuse pentru strategiile de cercetare cantitative putem observa că unul dintre criteriile centrale care stau la baza lor este potențialul de inferență causală. Din această perspectivă, cei mai mulți autori optează pentru existența a trei abordări majore (Alasuutari, Bickman, & Brannen, 2008; Mitchell & Jolley, 2001): *experimentul*, *quasi-experimentul* și cercetările denumite generic "*corelaționale*", acestea din urmă urmărind, fie asocierea dintre variabile, fie diferențele dintre grupuri independente.

## 2 Inferența causală

### 2.1 Cauza, efectul și relația causală

"Teoriile reprezintă explicații cauzale. Scopul oricărei științe este explicația, iar explicația este întotdeauna causală" (Schmidt, 1992, p. 1177). Forma cea mai înaltă a cunoașterii științifice înseamnă aflarea cauzei fenomenelor. Probabil nu există nici un cercetător care să nu accepte această afirmație. Problemele încep însă imediat ce ne punem problema ce se înțelege prin "cauză" și ce se înțelege prin "efect". La nivelul gândirii comune, recunoaștem cu relativă ușurință ceea ce este "cauză" unui anumit "efect". De exemplu, dacă se întrerupe curentul, iar liftul în care ne aflăm se blochează între etaje, ni se pare evident că întreruperea curentului este cauza blocării liftului. Această concluzie corespunde definiției filozofului și fizicianului englez John Locke (1632-1704): "O cauză, este ceva care face ca orice altceva, fie o simplă idee, o substanță sau un lucru, să înceapă a fi; iar un efect este ceea ce își are începutul din altceva" (Locke, 1690). Lucrurile se complică însă, dacă încercăm să identificăm acel "ceva" de la care efectul "a început a fi". Este oare oprirea curentului cauza imediată a blocării liftului? Dacă aprofundăm căutarea s-ar putea să aflăm că oprirea curentului a fost determinată de tăierea unui cablu de către lama unui buldozer. Asta înseamnă că "adevărată" cauză a blocării liftului constă într-o manevră greșită a buldozeristului. Numai că este posibil ca acesta să nu fi fost informat de șefii săi că în zona respectivă este plasat un cablu electric, ori că schema rețelei electrice nu fusese actualizată după o intervenție recentă, sau că cei care au montat cablul electric nu l-au pus la adâncimea prevăzută în proiect. Cu fiecare nouă explicație pe care o găsim, părem a ne îndepărta tot mai mult de "cauză" care era atât de evidentă în clipa în care liftul s-a blocat între etaje. Iar acest lucru nu se întâmplă doar într-o situație ca cea evocată în exemplul nostru, ci în orice situație în care ne punem problema aflării cauzei.

Pentru a rezolva această problemă, Mackie (apud Horsten & Weber, 2005) a propus conceptul de "*inus condition*", care definește cauza drept "o parte insuficientă, dar neredundantă, a unei condiții, care este prin ea însăși nenecesară, dar suficientă, pentru apariția unui efect". În contextul exemplului de mai sus, întreruperea curentului electric este o "condiție insuficientă" deoarece nu este singura care poate explica blocarea liftului. În același timp, manevra a buldozeristului, lipsa de informare a acestuia de către șefi sau plasarea incorectă a cablului electric sunt "condiții nenecesare, dar suficiente" care au contribuit la blocarea liftului. Oprirea curentului are în acest context rolul de "*inus condition*".

Nici în ceea ce privește noțiunea de efect problemele care se ridică nu sunt mai simple. În exemplul de mai sus am luat în considerare drept efect, blocarea liftului, dar acest lucru s-a produs ca urmare a opririi funcționării motorului, care a fost determinată de întreruperea funcționării alimentatorului acestuia. Ceea ce rezultă de aici este că aceeași cauză poate avea efecte variate. De exemplu, persoanele din lift pot avea o criză de panică. Altfel spus, într-un anumit context causal vom descoperi doar acele efecte pe care suntem capabili sau interesați să le descoperim. Într-un studiu cu privire la efectul recompenselor asupra performanței în muncă, performanța poate fi măsurată sub diferite aspecte (de ex., eficiența în sarcină sau performanța contextuală), efectul putând fi diferit de la un tip de performanță la altul. Mai mult decât atât, capacitatea de a observa un efect depinde de măsura în care condițiile cercetării permit o comparație cu situația în care cauza nu s-ar fi manifestat. Este ceea ce filozoful David Hume (1711-1776) numea *raționament contrafactual*. Shadish et al. (2001) observă că nu putem trage o concluzie de tip causal decât în situațiile în care putem crea un context în care cauza este prezentă și un alt context în care cauza nu este prezentă. Astfel "efectul" este diferența pe care o putem observa în starea (nivelul) variabilei dependente între cele două situații. Chiar și așa, raționamentul contrafactual este fundamental calitativ, asumarea unei relații causale fiind o decizie subiectivă, chiar dacă ne bazăm pe proceduri statistice pentru evaluarea diferenței dintre cele două condiții (Shadish, 1995).

Din cele spuse până acum, rezultă că relația dintre cauză și efect este rezultatul unui proces care implică un întreg lanț de factori, fiecare dintre ei având un anumit rol în apariția efectului. Unii dintre aceștia sunt *factori necesari* (în absența lor efectul nu se produce), alții sunt *factori suficienți* (prezența lor poate determina apariția efectului, dar același efect poate determinat și de alți factori), iar alții sunt *factori contributori* (nu sunt nici necesari, nici suficienți, dar prezența lor afectează modul în care se manifestă cauza în determinarea efectului).

Expresia determinării multiple a relației causale a fost inclusă de Reichardt (2006) în *principiul paralelismului*, conform căruia dacă există un efect ca funcție a unei cauze, atunci în această relație intervin în mod necesar următorii patru factori: subiectul, mediul, timpul și rezultatul. Dacă se poate realiza un model care să permită inferența causală pentru unul dintre acești factori, atunci se poate realiza un model care să susțină o concluzie causală pentru oricare dintre ceilalți trei. De exemplu, efectele sunt estimate prin evaluarea comparațiilor între mai mulți subiecți care primesc tratamente diferite. Principiul paralelismului implică faptul că un efect poate fi de asemenea surprins prin efectuarea unei comparații între diverse locații, momente sau variabile care descriu efectul respectivei cauze. Această tipologie deschide o perspectivă mai largă asupra opțiunilor pe care le are un cercetător atunci când își pune problema identificării unei relații causale.

În realitatea vieții, chiar dacă știm că un anumit efect este rezultatul unui complex de factori, nu putem niciodată să îi cunoaștem pe toți și nici cum interacționează ei în determinarea efectului. Din acest motiv, cunoașterea unei relații causale se limitează de regulă la inferarea probabilității de apariție a unui efect într-un anumit context. Iar dacă avem în vedere specificul contextului în care a fost pusă în evidență o anumită relație causală, rezultă că posibilitatea de generalizare a unui efect causal de la un context la altul este discutabilă și nesigură (Shadish et al., 2001). La aceeași concluzie ajunsese cu 80 de ani mai înainte și Niles (1922), care afirma că știința este limitată în posibilitatea ei de a explica complet cauza diverselor fenomene psihice, nu poate demonstra necesitatea absolută a unei secvențe causale și nici nu poate susține cu certitudine repetabilitatea acesteia, trebuind să se limiteze la concluzii probabilistice.

*Criteriile necesare pentru inferarea unei relații causale* au fost formalizate de John Stuart Mill (1806-1873) și cuprind trei condiții fundamentale (apud Holland, 1986):

- cauza trebuie precedă efectul;
- între cauză și efect trebuie să poată fi dovedită o relație;
- nu poată fi dovedită nici o altă explicație pentru apariția efectului.

Fiecare dintre aceste condiții ridică probleme majore. Precedența cauzei nu se reduce la măsurarea acesteia "înainte" de măsurarea efectului, ci la manifestarea ei înainte de apariția efectului, fapt care exclude posibilitatea de a fi influențată de acesta. Dacă presupunem că nivelul performanței este cauzat de nivelul satisfacției în muncă, ar trebui să putem dovedi că satisfacția, la rândul ei, nu este un efect al performanței. Măsurarea performanței și a satisfacției și calcularea corelației dintre ele nu este în măsură să susțină singură o concluzie causală. Pentru aceasta ar trebui să realizăm un model de cercetare în care să ne asigurăm că nivelul satisfacției nu poate fi influențat de performanță (de ex.,

prin blocarea feedbackului cu privire la performanță). Nu trebuie să uităm, de asemenea, că utilizarea corelației liniare nu permite identificarea corelațiilor de tip curbiliniu, ceea ce înseamnă că absența unei corelații liniare nu este neapărat echivalentă cu absența unei relații între două variabile. În același timp ar trebui să ne asigurăm că nici o altă variabilă în afară de satisfacție nu poate avea un impact asupra performanței. În acest scop, prima sarcină a cercetătorului este aceea de a identifica variabilele care ar putea interveni între *VI* și *VD*, iar a doua sarcină este aceea de a le controla efectul.

## 2.2 Condițiile inferenței cauzale în tradiția lui Donald Campbell

Una dintre abordările fundamentale în problema inferenței cauzale în cercetarea psihologică este aceea inițiată de Donald Campbell (Campbell, 1957; Campbell & Stanley, 1966). În esență, această orientare se focalizează pe identificarea amenințărilor la adresa validității interne, care subminează inferența unei relații cauzele dintre variabila independentă și variabila dependentă. Lista acestor amenințări, descrisă inițial de Campbell, a făcut obiectul unei analize aprofundate de-a lungul timpului, fiind practic general asumată în prezent (Holland, 1986; West, Biesanz, & Pitts, 2000). Pentru a le putea înțelege mai ușor, le vom prezenta în continuare în contextul unui studiu ipotetic cu privire la efectul unui program de instruire destinat reducerii comportamentelor contraproductive într-o organizație:

- *Stabilitatea temporală*, care implică faptul că răspunsul la tratament nu depinde de momentul aplicării acestuia, ci numai de tratamentul însuși. Altfel spus, trebuie să ne asigurăm că numărul conduitelor contraproductive nu a scăzut pentru că, de exemplu, programul de instruire s-a terminat într-o perioadă în care activitatea a organizației era redusă, iar oportunitatea conduitelor contraproductive era, din acest motiv, mai redusă și ea. În cazul cercetărilor psihologice, condiția de stabilitate temporală poate fi afectată de numeroși factori, cum ar fi ”maturizarea” participanților, intervenția unor evenimente ca oboseala, variația ciclică a potențialului psihic etc.

- *Tranziția cauzală* se referă la caracterul inductor al măsurării din condiția de control. Acesta nu trebuie să afecteze măsurarea care urmează tratamentului. Dacă, de exemplu, participanții la studiu sunt conștienți că organizația manifestă un interes special cu privire la comportamentele negative și că are loc o înregistrare a acestora, acest lucru poate reprezenta o violare a condiției de tranziție cauzală, deoarece reducerea numărului de comportamente negative nu ar mai fi atribuibilă programului de instruire, ci (și) efectului de ”alarmă” indus de efectuarea primei măsurări.

- *Omogenitatea unităților de studiu* este o condiție care trebuie întrunită atât în studiile în care se compară valori măsurate la nivelul aceluiși grup, cât și în studiile în care sunt comparate valori măsurate pe grupuri diferite. În primul caz, omogenitatea se referă la faptul că unitățile de studiu (individuale sau colective) fac parte din ”aceeași categorie”, dacă le raportăm la natura tratamentului. De exemplu, în cazul în care grupul supus instruirii este format din angajați ”model”, dar și din angajați ”recidiviști” în materie de comportamente contraproductive, atunci este posibil ca, pe ansamblu, efectul tratamentului să nu poată fi observat. În al doilea caz, omogenitatea ridică probleme încă și mai dificile, deoarece grupurile sunt formate din subiecți diferiți, ceea ce poate face ca ”răspunsul” la tratament să fie diferit din acest motiv, iar nu ca urmare a tratamentului (programul de instruire).

- *Stabilitatea valorii tratamentului la nivelul unității (Stable Unit Treatment Value Assumption - SUTVA)* este o condiție complexă, care poate fi evidențiată sub mai multe aspecte:

- ✓ *Stabilitatea valorii în raport cu includerea în grupul cercetării*, presupune ca performanța să nu fie afectată de însăși includerea în grupul de cercetare. De exemplu, dacă participanții din grupul supus tratamentului vor considera că au fost aleși ca urmare a ”calităților lor deosebite”, este posibil ca efectul tratamentului să fie stimulat în mod artificial, iar condiția SUTVA să fie încălcată.
- ✓ *Stabilitatea valorii în raport cu tratamentul primit de celelalte grupuri*. Cunoașterea tratamentului aplicat grupului de control de către cei din grupul de tratament (sau invers) poate afecta stabilitatea valorilor. Altfel spus, valorile obținute în condițiile specifice unui grup nu trebuie să fie influențate de condițiile în care sunt obținute valorile în grupul de comparație. Cel mai potrivit exemplu în acest context este ”efectul Hawthorne” (Morawski & Bayer, 2003), consemnat cu ocazia experimentelor conduse de Elton Mayo la uzinele Western Electric Company de lângă Chicago, în

perioada 1927-1939. Deși obiectivul cercetărilor a vizat evaluarea efectului condițiilor fizice asupra performanței muncitorilor, rezultatele au arătat că atât grupul de control (care lucra în condiții de iluminare corespunzătoare), cât și grupul experimental (care lucra în condiții de iluminare redusă cu 75%), nu au obținut doar o productivitate mărită, ci au resimțit și o satisfacție mai mare. Muncitorii din ambele grupuri au reacționat prin creșterea motivației ca urmare a percepției interesului din partea cercetătorilor pentru munca lor.

- ✓ *Stabilitatea valorii în raport cu tratamentul.* Această condiție presupune că efectul tratamentului nu variază în timp, fie prin "eroziune naturală" (de ex., oboseală, uitare), fie prin destructurare ca urmare a contactului cu alte surse de influență. Programul de reducere a comportamentelor contraproductive poate fi eficient, dar dacă înainte de măsurarea efectului acestuia intervine un eveniment personal sau colectiv (de exemplu, o reducere de salariu) efectul poate fi anihilat/atenuat și nu mai poate fi pus în evidență.

Printre cele mai eficiente strategii de creștere a șansei de asumare a condiției SUTVA pot fi citate: (i) păstrarea discreției asupra naturii cercetării față de participanți; (ii) mascarea tratamentului (metodă frecvent utilizată, de exemplu, în medicină, unde un nou medicament se administrează "orb", pacienții neștiind dacă iau medicament sau placebo); (iii) evitarea situației în care includerea într-un anumit grup de cercetare să fie percepută de participanți ca o sursă de beneficii/dezavantaje comparativ cu participanții din alt grup de cercetare.

- *Administrarea tratamentului la nivel de grup.* Aplicarea unui tratament se poate face la nivel individual (instruire cu fiecare angajat în parte), dar și la nivel de grup (instruire de grup). În acest al doilea caz, apar probleme de ordin statistic care trebuie avute în vedere. Atunci când se lucrează la nivel de grup, întregul grup este supus, fie condiției de control, fie condiției de tratament. Sub aspect statistic, rezultatul este acela că răspunsul (performanța) la nivelul fiecărui individ nu mai este independentă de a celorlalți membri ai grupului. Altfel spus, dacă selectăm la întâmplare doi participanți din grupul supus tratamentului, valorile lor vor fi mai apropiate decât ale unor participanți extrași la întâmplare din orice alte grupuri. Acest fapt nu afectează estimarea efectului cauzal, în spiritul modelului lui Rubin, ci eroarea standard, care este mai mică decât în cazul în care valorile ar fi independente. Consecința acestei situații este amplificarea artificială a probabilității de respingere a ipotezei de nul, prin creșterea erorii statistice de *tip I*. Chiar și în condițiile unui nivel relativ redus de dependență a valorilor intra-grup, atunci când volumul grupului atinge 25 de subiecți, eroarea de tip I (respingerea  $H_0$  când în realitate este adevărată) are o probabilitate de 0.19, față de nivelul teoretic admis de 0.05 (Barcikowski, apud Morgan-Lopez, 2007).

- *Complianța la tratament.* Pentru a avea efect, orice tratament (terapie psihică, program de instruire) trebuie să fie acceptat și asumat de participanți. Așa cum există pacienți care refuza medicamentele prescrise de medici, la fel este posibil ca participanții la cercetările psihologice să nu reacționeze la condițiile acestea, din motive care nu au nici o legătură cu natura tratamentului. De exemplu, participarea la un programul de instruire poate implica ore suplimentare, deplasarea într-o altă locație sau, pur și simplu, nu reprezintă un interes pentru unii dintre subiecți. Gradul de complianță la tratament este întotdeauna variabil într-o cercetare, iar acest lucru afectează identificarea efectului cauzal.

- *Stabilitatea participanților.* Această condiție poate fi privită sub două aspecte: *prima*, se referă la diminuarea numărului de subiecți în fază post-test față de faza pre-test; *a doua*, vizează situația în care unii dintre subiecți nu pot fi evaluați post-test. Ambele situații reprezintă surse de eroare în estimarea efectului cauzal. Jurs și Glass (apud West et al., 2000) au propus o strategie în doi pași pentru detectarea erorii generate de diminuarea diferențiată a subiecților între grupul supus tratamentului și cel de control. *La primul pas*, se compară procentul participanților care s-au retras din cele două grupuri. Dacă aceste valori diferă, atunci raportul de pierdere diferențială poate fi interpretat ca efect al tratamentului. *Al doilea pas*, presupune efectuarea unei serii de comparații ANOVA, pe datele pre-test, pentru combinația 2 (grup tratament/grup control) x 2 (subiecți compleți/subiecți pierduți), cu scopul de a identifica eventuale caracteristici de fond sub care diferă participanții (atitudini, personalitate, date demografice etc.). Obținerea unui efect principal semnificativ pentru grupul de tratament, indică o posibilă eroare de eșantionare a grupului, pe când un efect principal semnificativ pentru subiecți compleți vs. subiecți pierduți sugerează că rezultatele nu pot fi

generalizate. În fine, efectul interacțiunii grup tratament x status post-test este un semn că subiecții care au părăsit cercetarea diferă de cei care au rămas, ceea ce reprezintă o sursă de eroare în estimarea efectului causal.

- *Omogenitatea unității de tratament* presupune faptul că unitățile supuse tratamentului sunt identice sub toate aspectele. Acest lucru înseamnă că fiecare unitate (individuală sau colectivă) este un receptor identic pentru condiția de tratament și poate fi formalizat astfel:  $Y_t(u_1) = Y_t(u_2)$  și  $Y_c(u_1) = Y_c(u_2)$ , unde  $u_1$  și  $u_2$  sunt două unități (subiecți) diferite. Numai în aceste condiții  $Y_t(u_1) - Y_c(u_1)$  reprezintă o estimare corectă a efectului causal, în conformitate cu modelul causal Rubin. Din păcate, această condiție este aproape imposibil de susținut în studiile psihologice, dată fiind unicitatea configurației psihologice a ființei umane.

- *Randomizarea*. Ideea de bază a modelului causal propus de Rubin este aceea că există un efect potențial diferit în funcție de modul în care unitatea investigată se raportează la tratament. De exemplu, putem presupune că efectul programului de reducere a comportamentelor contraproductive este diferit la angajații din organizații de stat comparativ cu angajații din organizațiile private. Dat fiind faptul că este imposibil să măsurăm efectul pe aceeași persoană, odată ca angajat la stat și apoi ca angajat în mediul privat (*problema fundamentală a inferenței causale*) trebuie să ne asigurăm că în grupul supus programului de instruire există șanse egale de participare pentru angajați din ambele tipuri de organizații. Soluția acestei probleme este repartiția aleatoare a participanților. Desigur, s-ar putea spune că din moment ce știm care este variabila covariantă (tipul de organizație) am putea aranja ca în grupul cercetării să avem în mod egal angajați din ambele tipuri de organizații. Numai că în realitate numărul variabilelor covariante este imposibil de cunoscut, iar repartiția aleatoare are tocmai efectul de a omogeniza grupurile cercetării, prin egalizarea aleatoare a șanselor pentru orice subgrup definit valorile oricărei variabile covariante, cunoscute sau necunoscute (sex, vârstă, vechime în muncă, specificul locului de muncă, trăsături de caracter etc.). Mecanismul de repartizare aleatoare în grupurile cercetării (tratament vs. control) echilibrează grupurile comparate, făcându-le echivalente, și permite interpretarea eventualei diferențe dintre rezultatele lor ca fiind atribuibilă tratamentului și nu unor variabile covariante. Repartiția aleatoare face diferența dintre experimentul controlat aleatoriu și un studiu observațional pe aceeași temă.

O consecință a repartiției aleatoare o reprezintă posibilitatea de egalizare a nivelului de bază al grupului sub aspectul variabilei dependente. Este evident că orice tratament nu vine pe un loc complet gol. Fiecare angajat are un anumit mod de raportare la comportamentele contraproductive înainte de începerea programului, așa cum fiecare elev are un anumit nivel de cunoștințe de bază înainte de începerea unei noi discipline de studiu. Repartiția aleatoare oferă șansa ca "nivelul de bază" al grupurilor supuse comparației să fie echivalent, ceea ce aduce un argument în plus pentru susținerea unei inferențe causale cu privire la rezultatul cercetării.

Importanța repartiției aleatoare este evidențiată cu pregnanță un studiu de referință efectuat de Shadish și Clark (Shadish & Clark, 2008; Shadish, Clark, & Steiner, 2008). Obiectivul pe care și l-au propus cercetătorii a fost de a evidenția dacă experimentul cu repartiție ne-aleatorie (*cvasi-experiment*) are aceeași putere cu experimentul cu repartiție aleatorie (*experiment adevărat*). În acest scop, a fost realizată o cercetare având drept suport eficiența unui program de instruire în două tipuri de sarcină: matematică și vocabular (limba engleză). Subiecții, studenți la psihologie, au fost separați aleatoriu în două grupuri, unul care urma să fie repartizat aleatoriu pe tipul de sarcină, iar altul în care subiecții au putut să aleagă ei înșiși ce program vor să urmeze (fig. 6.3). Rezultatele au arătat că în experimentul non-randomizat, unde subiecții au ales din ce grup experimental să facă parte, în funcție de preferința pentru tipul de sarcină, efectul antrenamentului a fost supraestimat, cu 18.6% în cazul matematicii și cu 11% în cazul vocabularului.

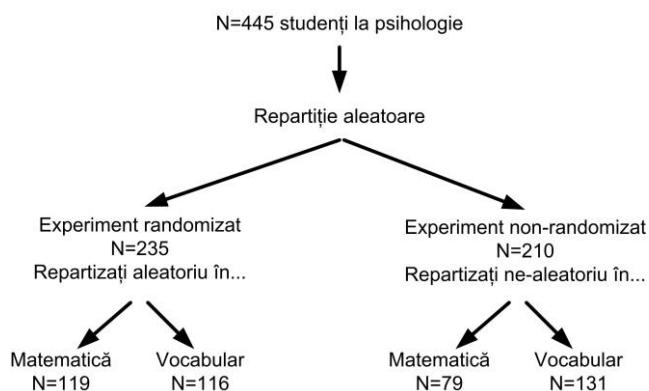


Figura 6.3. Repartiția subiecților în studiul Shadish (op. cit.)

Realizarea practică a procesului de randomizare se poate obține prin orice metodă care exclude subiectivismul repartiției, fie din partea cercetătorului, fie din partea participanților (auto-selecția. În acest scop, se poate ”da cu banul”, ori se poate utiliza un algoritm de selecție implementat într-unul din programele computerizate (de ex., SPSS). Importantă rămâne însă rigoarea procesului. Orice abatere de la regula de repartiție aleatoare este în măsură să afecteze fundamentarea causală a concluziei. Non-compliance la tratament, pierderea subiecților și încălcarea condiției *SUTVA*, sunt cei mai importanți factori care pot afecta în mod semnificativ eficiența repartiției aleatoare

O privire retrospectivă asupra condițiilor impuse pentru asumarea unei concluzii de tip causal ar fi de natură să descurajeze orice tentativă în acest sens, având în vedere evasi-imposibilitatea întrunirii simultane și perfecte a tuturor acestor condiții. Practic, este greu de imaginat un model de cercetare în domeniul psihologiei, al științelor socioumane, în general, în care să nu poată fi identificate o serie întreagă de amenințări la adresa posibilității de inferență causală. Nici chiar în cele mai riguroase experimente de laborator nu pot fi controlate toate variabilele și nici nu se pot asigura condiții perfecte de măsurare a efectelor. Nevoia cunoașterii științifice, fundamentată pe dovezi empirice, impune menținerea interesului pentru explicația causală, chiar și în condiții sub standardul ideal. Astfel, au fost propuse soluții pentru acceptarea unor concluzii causale chiar și în contextul unor cercetări non-experimentale (asupra cărora vom reveni cu alt prilej). Krause (2010) afirmă că psihologie este obligată să încerce să descopere relații causale, în ciuda dificultăților careia trebuie să îi facă față. Ca urmare, el propune conceptul strategic de ”condiție causală suficientă” (*sufficient condition cause*), capabil să pună în evidență sursele de influență causală într-un anumit context.

### 2.3 Modelul causal al lui Rubin (MCR)

Pornind de la abordarea inferenței causale propusă de Campbell, Donald Rubin a elaborat, după un deceniu și jumătate, o teorie proprie, cunoscută în prezent sub numele de ”*modelul causal Rubin*” (Rosenbaum & Rubin, 1983; Rubin, 1974, 1977, 1978, 1986, 1997, 2010). Principala diferență între aceste două modele este dată de faptul că, în timp ce Campbell s-a focalizat pe susținerea sensului efectului causal, obiectivul fundamental al lui Rubin a fost acela de a propune un model statistic prin care să poată fi estimată mărimea efectului causal (West & Thoemmes, 2010).

Pentru a înțelege acest model în liniile sale generale, să ne imaginăm că într-o organizație care urmărește creșterea eficienței, se introduce un program de instruire a angajaților cu scopul reducerii comportamentelor contraproductive (absenteism, furt, distrugere de bunuri etc.). Pentru evaluarea eficienței, numărul comportamentelor contraproductive este evaluat de două ori, înainte și după parcurgerea programului de instruire. Să presupunem că grupul supus cercetării înregistrează, în medie, mai multe conduite contraproductive înaintea programului de instruire decât după parcurgerea lui. Într-un astfel de context de cercetare, presupunând întrunirea tuturor condițiilor impuse de modelul Campbell, Rubin (1974, 2004) definește efectul causal prin diferența dintre cele două valori de performanță:

$$Y_t(u) - Y_c(u)$$

unde:

- $Y$  simbolizează valoarea variabilei dependente ("criteriul" din ecuația de regresie)
- $t$  se referă la condiția tratamentului
- $c$  se referă la condiția de comparație (de regulă, grupul de control)
- $u$  se referă la unitatea supusă tratamentului.

Practic,  $Y_t(u)$  reprezintă numărul de conduite contraproductive la nivelul unității de studiu supuse tratamentului (în cazul nostru, fiecare subiect), iar  $Y_c(u)$  este nivelul de control, măsurat înaintea tratamentului, de asemenea pe fiecare subiect în parte. De precizat că unitatea de studiu într-o cercetare poate fi de natură individuală sau colectivă. În primul caz, tipic pentru domeniul psihologiei, variabilele măsurate descriu diferite caracteristici ale unor persoane. În al doilea caz, mai frecvent întâlnit în sociologie, psihologia socială, educație, medicină, variabilele descriu colectivități (grupuri, școli, localități, unități sanitare etc.). De exemplu, efectul unui nou sistem de evaluare a performanței poate studiat la nivel "individual" (prin compararea performanțelor angajaților care au fost evaluați cu noul sistem, față de cei evaluați cu vechiul sistem), sau la nivel "colectiv", prin studierea comparativă a performanței unor organizații care au utilizat/nu au utilizat noul sistem de evaluare. În acest al doilea caz, valorile  $VD$  descriu fiecare organizație luată în studiu, nu pe fiecare angajat al acestora.

Dacă  $u$  se referă la o unitate cu un singur element (persoană, grup, organizație etc.), atunci ecuația de mai sus descrie un *efect causal individual*. Dacă cercetarea se referă la un eșantion de unități, atunci avem de a face cu un *efect causal mediu*, care descrie tendința centrală a efectelor înregistrate la nivelul fiecărei unități individuale componente. În fine, dat fiind faptul că cercetarea a măsurat efectul tratamentului la nivelul unui eșantion, atunci trebuie să avem în vedere un *efect causal estimat la nivelul populației*, a cărui valoare decurge din potențialul de generalizare pe care îl asigură gradul de reprezentativitate al eșantionului.

MCR ne oferă o definiție operațională a efectului causal, dar în același timp evidențiază ceea ce Holland (1986, p. 947) numește *problema fundamentală a inferenței causale*: "Este imposibil să observăm ambele valori  $Y_t(u)$  și  $Y_c(u)$  asupra aceleiași unități și, ca urmare, este imposibil să observăm efectul lui  $t$  asupra lui  $u$ ". În contextul exemplului nostru, această afirmație înseamnă că nu putem avea garanția absolută a faptului că unitatea de studiu (angajatul) este identic în cele două momente ale măsurării, singura diferență fiind aceea indusă de programul de intervenție. Spre deosebire de lumea fizică, în lumea naturală există numeroase aspecte, în afară de condiția tratamentului, care pot interveni între cele două măsurări, făcând inferența causală ne-sustenabilă. În exemplul nostru am evocat doar două astfel de surse (salarizarea și sistemul de supraveghere), dar putem adăuga cu ușurință și altele: maturizarea angajaților, creșterea atașamentului față de organizație, concedierea unor participanți care au manifestat conduite contraproductive pe durata programului de instruire, includerea în program pe bază de voluntariat, fapt care a favorizat prezența în eșantion a unor angajați "sensibili" în raport cu scopurile instruirii etc.

Lucrurile devin încă și mai complicate dacă programul de intervenție s-ar efectua pe două grupuri diferite de angajați. În acest caz, diferențele potențiale dintre membrii grupului experimental și cei ai grupului de control pot fi foarte numeroase. Evident, soluția obligatorie de luat în considerare în acest caz este repartiția aleatorie a angajaților în cele două grupuri.

Putem concluziona că problema cea mai dificilă a inferenței causale este aceea de a izola efectul variabilelor covariante cu variabila causală. În principiu, avem două soluții la îndemână două soluții indirecte, pe care Holland (op.cit.) le numește *soluția științifică* și *soluția statistică*. În timp ce soluția statistică se bazează pe proceduri de tratare a datelor care controlează variabilele al căror efect dorim să îl "eliminăm" (de ex. analiza de covarianță, regresia multiplă, corelația parțială ș.a.), soluția științifică se referă la condițiile și modul în care datele sunt obținute (controlul experimental al variabilelor).



### 3 Experimentul

Experimentul este singura modalitate de cercetare care întrunește cele trei condiții propuse de John Stuart Mill pentru asumarea unei concluzii cauzale:

- *precedența cauzei asupra efectului* – este aplicat un ”tratament”, după care se observă efectul acestuia;
- *existența unei relații între cauză și efect* – se analizează existența unei legături între ”tratament” (nivelurile VI) și variația VD;
- *unicitatea explicației* – se utilizează diverse metode prin care sunt controlate efectele altor variabile, ceea ce face posibilă reducerea posibilității altor explicații decât VI pentru variația VD.

Trăsătura definitorie a experimentului este dată, înainte de toate, de obiectivul de cunoaștere pe care și-l propune, și anume identificarea unei relații cauzale între variabila independentă (VI - denumită și *tratament*) și variabila dependentă (VD – denumită și *rezultat*). Experimentul reprezintă o procedură de interogare a realității care, departe de a se comporta consistent, este oscilantă, ambiguă și confuză. Cea mai importantă sursă de confuzie este variabilitatea, în particular, faptul că aceeași cauză are efecte diferite asupra unor receptori diferiți. Dar, comparând variabilitatea efectului asupra unor subiecți supuși unor condiții diferite de tratament se pot deduce concluzii fundamentate cu privire la relația cauzală dintre tratament și efect (Kirk, 2003).

Întrebările tipice care stau la baza experimentelor sunt de genul următor: *există un efect asupra performanței generat de introducerea unui program de instruire?*; *aplicarea unui program de combatere a fumatului produce o scădere consumului de țigări?*; *are consumul de alcool un efect asupra vitezei de reacție a conducătorilor auto?*; *așteptările profesorilor au un efect asupra performanței elevilor (efectul Pygmalion)?*; *are dispoziția afectivă un efect asupra memoriei?*. În fiecare dintre situațiile evocate avem de a face cu un anumit ”tratament” (programul de instruire, programul de combatere a fumatului, consumul de alcool, așteptările profesorilor, dispoziția afectivă) al cărui efect trebuie pus în evidență prin variația valorilor unei variabile dependente (performanța, consumul de țigări, viteza de reacție, performanța elevilor, memoria).

Experimentele au fost utilizate în psihologie încă de la sfârșitul sec XIX și începutul sec. XX, Woodworth fiind primul care descrie metoda experimentală în psihologie, denumind-o ”metoda cea mai de încredere pentru observarea fenomenelor în condiții controlate” (apud Fuchs & Milar, 2003). În anul 1903 Schuyler a introdus utilizarea grupului experimental și a grupului de control, care a devenit între timp o procedură clasică. În anul 1916, McCall a avansat ideea repartizării aleatoare în cele două grupuri comparate, iar în 1925, odată cu lucrarea *How To Conduct an Experiment*, a instituit comparația dintre grupuri ca procedură standard pentru experiment (apud Creswell, 2008). În aceeași perioadă Ronald A. Fisher a propus un aparat statistic de bază pentru analiza datelor experimentale, odată cu publicarea cărții sale *Statistical Methods for Research Workers* (Fisher, 1926).

În funcție de capacitatea de a susține o concluzie de tip causal, strategiile de tip experimental sunt clasificate în două categorii: experimentul adevărat (*true experiment*) și quasi-experimentul.

#### 3.1 Experimentul adevărat

Un experiment adevărat este un model de cercetare care se desfășoară de regulă în condiții de laborator și definit prin patru caracteristici de bază:

(i) *Manipularea uneia sau multor variabile independente*. Cel mai simplu experiment se bazează pe existența unei variabile independente cu două niveluri (condiții), care sunt suficient de puternice (distincte) pentru a genera efecte diferite asupra unei variabile dependente. În ciuda unei practici destul de răspândite, utilizarea unui grup de control, care nu este supus tratamentului, nu reprezintă o condiție pentru un experiment (Stone-Romero, 2004).

*Repartizarea aleatoare a unităților de studiu* (care pot fi indivizi sau grupuri) în raport cu condițiile experimentale. Presupunând că există suficiente unități de studiu, repartiția aleatoare asigură condiții de echivalență a grupurilor comparate și, prin aceasta, validitatea internă care să permită încrederea în rezultatul cercetării. Principala problemă pe care o are de înfruntat orice cercetare științifică este aceea a amenințărilor la adresa validității interne. Repartiția aleatoare a subiecților în grupurile comparate asigură cea mai puternică armă împotriva amenințărilor la adresa validității interne, deoarece, prin

aceasta, se elimină posibilele influențe ale surselor externe de variație. Totuși, este bine să reținem ca repartiția aleatoare nu face decât să asigure condițiile teoretice omogenizării grupurilor de studiu, fără a ne oferi și garanția că acest lucru este realizat, deoarece hazardul poate produce orice fel de repartiție a unităților de studiu, inclusiv repartiții non-echivalente între grupurile experimentale (McMillan, 2007)

(ii) *Măsurarea efectului condițiilor tratamentului* asupra uneia sau mai multor variabile dependente (atitudini, comportamente etc.), dar și efecte induse de variabile externe (covariante).

(iii) *Menținerea constantă a tuturor variabilelor externe*, care ar putea să inducă o ambiguitate în legătură cu relația cauzală dintre tratament și rezultatul acestuia. Acest obiectiv poate fi atins, fie prin aranjarea condițiilor cercetării astfel încât acestea să fie cât mai echivalente posibil, fie prin măsurarea variabilelor externe și eliminarea efectului lor prin proceduri statistice speciale (analiza de covarianță).

### 3.2 Quasi-experimentul

Quasi-experimentul este un model de cercetare care se desfășoară, de regulă, în condiții naturale. Atunci când suntem nevoiți să comparăm grupuri din mediul real, nu este întotdeauna posibilă formarea de grupuri experimentale artificiale, compoziția grupurilor este cea naturală, iar procedura de evaluare și comparare a rezultatelor celor două grupuri este identică cu situația experimentală. Quasi-experimentul are două caracteristici similare cu experimentul adevărat: (i) utilizarea a cel puțin două condiții de tratament și (ii) măsurarea efectului asupra variabilei dependente. În același timp, se deosebește de experimentul adevărat sub două aspecte: (i) repartizarea nealeatoare a unităților de studiu în grupurile supuse comparației și (ii) singura modalitate de control a variabilelor externe este de ordin statistic. În quasi-experiment, unitățile de studiu sunt repartizate în raport cu condițiile de tratament prin auto-selecție, decizii administrative sau prin alegerea unor grupuri deja constituite (elevii unor clase, angajații unor secții sau compartimente etc.). Acest fapt lasă loc de manifestare pentru amenințări asupra validității interne. Din cauza lipsei de echivalență între grupurile comparate, răspunsul la tratament (rezultatul) nu poate fi pus exclusiv pe seama tratamentului.

Aceeași temă de cercetare poate fi abordată atât cu o strategie experimentală, cât și pe calea quasi-experimentului. De exemplu, să presupunem că dorim să studiem efectul stresului asupra numărului de erori. În varianta experimentului adevărat, am putea constitui două grupuri de subiecți, obținute prin repartizarea aleatoare din totalul subiecților disponibili. Unul din grupuri ar urma să execute o anumită sarcină în condiții de stres, iar celălalt grup ar urma să execute aceeași sarcină în absența stresului. În acest caz, diferența dintre numărul de erori ale celor două grupuri ar constitui baza de inferență pentru o eventuală implicare cauzală a stresului în producerea erorilor. Dacă însă nu putem organiza un experiment adevărat (de exemplu, conducerea uzinei nu ne permite), atunci putem analiza numărul de erori produse de două grupuri de angajați care lucrează în condiții diferite de stres. Evident, în acest al doilea caz, dacă există o diferență a numărului de erori între cele două condiții, operează amenințările la adresa validității, astfel încât o explicație cauzală validă este ne-sustenabilă (Cook, Campbell, & Peracchio, 1990; Shadish & Clark, 2008; Shadish et al., 2008; Steyer, Partchev, Kröhne, Nagengast, & Fiege, 2008).

### 3.3 Avantaje și dezavantaje ale experimentelor și quasi-experimentelor

Atât experimentul cât și quasi-experimentul utilizează aceleași modele de cercetare, despre care vom vorbi cu ocazia temelor viitoare, fiind deosebite doar sub aspectul fundamentării explicației, ca urmare a amenințărilor la adresa validității.

Experimentele se pot desfășura în condiții artificiale, de laborator, sau în condiții naturale. În general, *experimentul de laborator* randomizat este considerat "standardul de aur" în materie de inferență cauzală (Sackett & Larson Jr., 1990; Shadish et al., 2001; Steyer et al., 2008; Stone-Romero, 2004), deoarece întrunește cele trei condiții fundamentale:

- posibilitatea de control al variabilei independente;
- observarea efectului asupra variabilei dependente;
- controlul variabilelor externe (covariante).

Totuși, experimentul de laborator nu este nici el lipsit de anumite dezavantaje și limite.

- Cea mai importantă dintre acestea constă în faptul că pot fi puțin relevante în raport cu mediul real, de care se îndepărtează foarte mult. Natura artificială a contextului de cercetare reprezintă prin ea însăși o amenințare asupra validității externe (transpunerea concluziilor în mediul real. Un exemplu în acest sens l-ar putea constitui cercetările asupra deciziei. Modelul clasic al deciziei pune accentul pe procesul de identificare și ierarhizare a variantelor de acțiune. Acest tip de abordare a condus la modele explicative bazate pe descompunerea și analiza proceselor cognitive pe baza unei metodologii de laborator adesea foarte sofisticate, dar îndepărtate de condițiile reale, operaționale (valoarea/utilitatea variantelor de acțiune, estimarea probabilității de succes, teoria jocului etc.). Acest model de studiu prezintă cel puțin trei deficiențe majore: (i) are un caracter stereotip, rigid și inflexibil, fapt care limitează transpunerea concluziilor de laborator în situații reale; (ii) nu permite luarea în considerare a unor factori care influențează în mod semnificativ procesul deciziei: experiența individuală, stresul, oboseala sau diferite caracteristici individuale (agresivitatea, impulsivitatea, atitudinea față de risc etc.); (iii) dificultatea de a lua în considerare limitele umane și erorile de raționament. Într-o tentativă de eliminare a acestor neajunsuri, a fost propus așa numitul *model naturalistic al deciziei* (*Naturalistic Decision Making / NDM*), care se bazează pe studiul modului în care oamenii utilizează experiența pentru a lua decizii în mediul operațional (Klein, 1997). Componenta cheie a acestui mod de abordare este luarea în considerare a dimensiunii personale, ignorată în studiile de laborator. În timp ce abordarea clasică pune accentul pe identificare variantelor de acțiune și pe momentul alegerii, modelul NDM mută centrul de greutate pe individ și pe contextul situației decizionale concrete. Ceea ce stă în centrul modelului clasic al deciziei, momentul alegerii, devine relativ neimportant în contextul NDM, deoarece profesioniștii din diverse domenii nu par a avea dificultăți cu alegerea variantei de acțiune ci, mai ales, cu înțelegerea situației.
- De cele mai multe ori, sarcina pe care trebuie să o îndeplinească participanții într-un experiment de laborator este una nouă, ceea ce necesită timp pentru a fi înțeleasă și învățată.
- Induce conduite dezirabile social. Chiar dacă se adoptă măsuri de mascare a obiectivului cercetării, este greu ca participanții să fie complet izolați de efectele produse de acesta, ceea ce poate face ca reacțiile și comportamentele lor să fie modificate în sensul unui "răspuns pozitiv".
- Durata experimentelor este de regulă scurtă, ceea ce face, pe de o parte, dificil să fie atins nivelul optim de cunoștințe legate de îndeplinirea sarcinii, iar pe de altă parte, efectele reflectă nivelul limitat al participanților în raport cu sarcina.
- Existența unor limite de ordin etic face ca intensitatea variabilei independente să fie limitată, iar pe cale de consecință, și nivelul efectului acesteia.
- Experimentele permit doar o explicație globală ("molară") prin care se concluzionează cu privire la existența unei relații cauzale, dar nu sunt foarte potrivite pentru o explicație analitică ("moleculară") a mecanismului relației cauzale (molecular) (Shadish et al., 2001).

*Avantajele quasi-experimentelor* provin din eliminarea dezavantajele experimentelor de laborator: se desfășoară în mediul real; cu participanți care efectuează sarcini naturale, familiare; au, de regulă, o durată mai lungă; intensitatea "tratamentului" corespunde situațiilor reale. Din acest motiv, experimentele de teren, chiar dacă presupun acceptarea unor compromisuri metodologice, pot fi mai relevante. De exemplu, Lee (2000), într-o cercetare cu privire la justiția organizațională, observă că cele mai multe studii au fost efectuate în condiții controlate, de laborator, ceea ce face ca validitatea lor externă să fie discutabilă. În ce privește *dezavantajele experimentelor de teren*, acestea țin de dificultatea controlului variabilei independente și a variabilelor externe (covariante), compoziția grupurilor nu este aleatoare (auto-selecție, selecție conjuncturală)

Fiecare dintre aceste abordări are avantaje și dezavantaje. Experimentul elimină amenințările la adresa validității interne, dar limitează posibilitatea de generalizare. Quasi-experimentul reproduce la bine condițiile reale, dar este limitat sub aspectul validității interne.

#### 4 Întrebări recapitulative

- Ce se înțelege prin conceptul de "inus condition" propus de Mackie?
- Ce se înțelege prin raționamentul contrafactual, utilizat în definirea efectului?
- Ce se înțelege prin factori cauzali *necesari, suficienți și contributori*?
- Care sunt criteriile fundamentale pentru susținerea unei inferențe cauzale propuse de John Stuart Mill?
- Care sunt condițiile fundamentale pentru asumarea inferenței cauzale?
- Care este formalizarea modelului cauzal Rubin?
- Ce se înțelege prin efectul cauzal *individual, mediu și estimat*?
- În ce consta problema fundamentală a inferenței cauzale?
- Care sunt caracteristicile de bază ale experimentului adevărat?
- Care sunt caracteristicile quasi-experimentului?
- Care sunt avantajele experimentului adevărat?
- Care sunt dezavantajele experimentului adevărat?
- Care sunt avantajele quasi-experimentului?
- Care sunt dezavantajele quasi-experimentului?

#### 5 Exercițiu

Căutați în literatura științifică cel puțin un experiment sau un quasi-experiment din domeniul psihologiei organizaționale și analizați posibile amenințări la adresa validității interne.

#### 6 Referințe bibliografice

- Alasuutari, P., Bickman, L., & Brannen, J. (2008). Social Research in Social Changing Conditions. In P. Alasuutari, L. Bickman, & J. Brannen (Eds.), *The SAGE Handbook of Social Research Methods* (pp. 1-8): SAGE Publications.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54(4), 297-312 (Accesat la 219.208.2011: <http://campus.murraystate.edu/academic/faculty/mark.wattier/Campbell1957.PDF>).
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Cook, T. D., Campbell, D. T., & Peracchio, L. (1990). Quasi Experimentation. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 1, pp. 491-576). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Creswell, J. W. (2008). *Educational Research. Planning, Conducting, and Evaluating Quantitative and Qualitative Research*: Pearson Education Inc.
- Fisher, R. A. (1926). *Statistical Methods for Research Workers*. London: Oliver&Boyd.
- Fuchs, A. H., & Milar, K. S. (2003). Psychology as a Science. In D. K. Freedheim (Ed.), *Handbook of psychology: History of Psychology* (Vol. 1, pp. 1-26). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of American Statistical Association*, 81(396), 945-970.
- Horsten, L., & Weber, E. (2005). INUS Conditions *Encyclopedia of Statistics in Behavioral Science*: John Wiley & Sons, Ltd.
- Kirk, R. E. (2003). Experimental Design. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of Psychology. Research Methods in Psychology* (Vol. 2): John Wiley & Sons, Inc.
- Klein, G. (1997). The current status of the naturalistic decision making framework w. In R. Flin, E. Salas, M. Strub, & L. Martin (Eds.), *Decision Making Under Stress; Emerging Themes and Applications* (pp. 11-27): Ashgate.
- Krause, M. S. (2010). Trying to Discover Sufficient Condition Causes. *Methodology*, 6(2), 59–70.

- Lee, H.-R. (2000). *An Empirical Study of Organizational Justice as a Mediator of the Relationships among Leader-Member Exchange and Job Satisfaction, Organizational Commitment, and Turnover Intentions in the Lodging Industry*. (PhD Thesis), Blacksburg, Virginia (Accesat la 01.07.2010: <http://scholar.lib.vt.edu/theses/available/etd-05012000-14210002/unrestricted/dissertation.pdf>).
- Locke, J. (1690). *An Essay Concerning Human Understanding* (Accesat la 10.06.2011: [http://oregonstate.edu/instruct/phl302/texts/locke/locke1/Essay\\_contents.html](http://oregonstate.edu/instruct/phl302/texts/locke/locke1/Essay_contents.html)).
- McMillan, J. H. (2007). Randomized Field Trials and Internal Validity: Not So Fast My Friend *Practical Assessment, Research & Evaluation*, 12(15). <http://pareonline.net/pdf/v12n15.pdf> Retrieved from <http://pareonline.net/pdf/v12n15.pdf>
- Mitchell, M. L., & Jolley, J. M. (2001). *Research Design Explained* (4th ed.): Wadsworth Pub Co.
- Montero, I., & León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862 ([http://www.aepc.es/ijchp/articulos\\_pdf/ijchp-256.pdf](http://www.aepc.es/ijchp/articulos_pdf/ijchp-256.pdf)).
- Morawski, J. G., & Bayer, B. M. (2003). Social psychology. In D. K. Freedheim (Ed.), *History of psychology* (Vol. 1, pp. 223-248): John Wiley & Sons.
- Morgan-Lopez, A. A. (2007). Analytic Methods for Modeling Longitudinal Data from Rolling Therapy Groups with Membership Turnover. *Journal of Consulting and Clinical Psychology*, 75(4).
- Muchinsky, P. M. (2006). *Psychology Applied to Work* (Eighth ed.): Thomson Learning Inc.
- Niles, H. E. (1922). Correlation, Causation and Wright's Theory of "Path Coefficients". *Genetics*, 7, 258.
- Reichardt, C. S. (2006). The Principle of Parallelism in the Design of Studies to Estimate Treatment Effects. *Psychological Methods*, 11(1), 1-18.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Rubin, D. B. (1974). Estimating causal effects of treatment in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66, 688-701.
- Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics*, 2(1-26).
- Rubin, D. B. (1978). Bayesian inference for causal effects. *The Annals of Statistics*, 6, 34-58.
- Rubin, D. B. (1986). Which if have causal answers. *Journal of the American Statistical Association*, 81, 961-962.
- Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores *Annals of Internal Medicine*, 127, 757-763. Retrieved from [http://www.annals.org/cgi/content/full/127/8\\_Part\\_2/757](http://www.annals.org/cgi/content/full/127/8_Part_2/757)
- Rubin, D. B. (2004). *Basic Concepts of Statistical Inference for Causal Effects in Experiments and Observational Studies*. Course materials used in Quantitative Reasoning, taught by Donald Rubin at Harvard University.
- Rubin, D. B. (2010). Reflections Stimulated by the Comments of Shadish (2010) and West and Thoemmes (2010). *Psychological Methods*, 15(1), 38-46.
- Sackett, P. R., & Larson Jr., J. R. (1990). Research Strategies and Tactics in Industrial and Organizational Psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 1, pp. 419-489). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis and cumulative knowledge in psychology. *American Psychologist*, 47, 1173-1181.
- Shadish, W. R. (1995). Philosophy of Science and the Quantitative-Qualitative Debates: Thirteen Common Errors. *Evaluation and Program Planning*, 18(1), 63-75.
- Shadish, W. R., & Clark, M. H. (2008, 17-19 July). *A Randomized Experiment Comparing Random to Nonrandom Assignment*. Paper presented at the Symposium "Causal effects - Designs and Analysis", Jena, Germany.
- Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random to Nonrandom Assignment (in press). *Journal of the American Statistical Association*.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* Boston: Houghton Mifflin Company.
- Spata, A. V. (2003). *Research Methods. Science and Diversity*: John Wiley&Sons, Inc.
- Steyer, R., Partchev, I., Kröhne, U., Nagengast, B., & Fiege, C. (2008). *Causal Effects in Experiments and Quasi-Experiments*. Jena: Springer.
- Stone-Romero, E. F. (2004). The Relative Validity and Usefulness of Various Empirical Research Designs. In S. G. Rogelberg (Ed.), *Handbook of Research Methods in Industrial and Organizational Psychology* (pp. 77-99): Blackwell Publishing.
- Wampold, B. E. (2006). Designing a Research Study. In F. Leong, T.L. & J. T. Austin (Eds.), *The Psychology Research Handbook. A Guide for Graduate Students and Research Assistants* (Second ed., pp. 93-103): SAGE Publications.

- West, S. G., Biesanz, J. C., & Pitts, S. C. (2000). Causal Inference and Generalization in Field Settings: Experimental and Quasi-Experimental Designs. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology*: Cambridge University Press.
- West, S. G., & Thoemmes, F. (2010). Campbell's and Rubin's perspectives on causal inference. *Psychological Methods, 15*(1), 18-37.