

LA INTERFERÈNCIA DEL PODER SOBRE L'AVALUACIÓ: QUATRE CASOS DE PSEUDOAVALUACIÓ

**LLUÍS GARCIA I SEVILLA,¹ JOSEP REIG I VILALLONGA,² RAFAEL TORRUBIA I BELTRI,¹
GUILLEM VERGER I GARAU³**

¹Departament de Psicologia Mèdica i Psiquiatria, Universitat Autònoma de Barcelona

²Departament de Ciències Morfològiques, Universitat Autònoma de Barcelona

³Departament de Medicina Interna, Universitat Autònoma de Barcelona

El text present aprofundeix les idees generals exposades al Llibre de Ponències del Congrés sota el títol «La politització de l'avaluació o la psedoavaluació». Aquest aprofundiment és resultat de l'estudi de quatre casos flagrants de pseudoavaluació del nostre àmbit professional, que van des de les notes dels estudiants fins a la pràctica mèdica. Per tal de donar coherència a tot el text, hem tornat a incloure ací les idees generals expressades al Llibre de Ponències.

INTRODUCCIÓ

Els metges i biòlegs sabem que els processos vitals, per complexos que siguin, funcionen gràcies a mecanismes autoregulators, on les accions dels efectors sobre els receptors són modulades per les retroaccions, o informacions, que els receptors envien als efectors sobre llurs canvis d'estat causats per les accions efectores. Quan falla aquesta retroacció, per les raons que sigui, es generen greus errors de regulació i, finalment, l'aturada del procés vital.

En els processos socials també existeixen retroaccions que regulen les accions, des de la crítica dels altres fins a mecanismes autoregulators més complexos com el mercat. Molt sovint, a la nostra complexa societat, i a causa d'aquesta complexitat, és molt improbable que les retroaccions de les decisions preses a un cert nivell organitzatiu arribin als autors de les decisions. D'ací la pràctica impossibilitat de detectar els greus errors que malauradament massa sovint es cometen, una de les causes fonamentals de l'encariment, de la lentitud de reacció, del burocratisme, i de la ineficàcia dels sistemes de direcció de les grans institucions o corporacions, estatals o civils. Els directius

d'aquestes organitzacions, ben conscients del problema, tracten d'obtenir aquesta necessària informació retroactiva però malauradament, i també massa sovint, els procediments d'informació retroactiva usats no són correctes, amb les conseqüències nefastes que hom pugui imaginar.

D'altra banda, aquesta incorrecció no sols desprestigia injustament tots els procediments d'avaluació social, no sols fa que hom els dediqui menys recursos, tancant així el cercle de llur degradació absoluta, sinó que també fa que siguin elegits com a instrument de justificació de polítiques directives mancades d'argumentacions racionals i empíriques. És allò que, com a catalans, i a tall d'exemple, patim: ens conculquen els nostres drets polítics i lingüístics a tots els nivells i després hom fa enquestes a la nostra població on es «veu clarament» que els catalans renuncien a emprar llur llengua en moltes situacions; aquestes enquestes són usades per demostrar «científicament» que els catalans, en l'aspecte lingüístic, encara tenen més del que els cal. Aquesta pseudodemostració científica ens revolta, no sols per la seva falsedat, sinó també perquè ens humilia profundament. Hem d'evitar, emperò, que el nostre rebuig, o acceptació, d'una avaluació sigui basat en motius emocionals, perquè ens desagradia o no. Cal que, com a científics, usem les armes de la ciència: tota mesura, com explicitarem més endavant, és part inextricable d'una teoria, i la teoria implícita d'aquest cas no fa referència a cap concepció filosòfica de «lliure albir», sinó que es refereix als mecanismes, tan coneguts d'altra banda, de la repressió socio-política, de la «domesticació» d'una població per part d'un poder polític; per tant, allò que detecten les dites enquestes són els efectes de la repressió i no cap altra cosa.

En definitiva, l'avaluació és una situació social asimètrica i moltes vegades qui mana diu que avalua per tal de justificar les seves decisions. Ara bé, en democràcia hom tria qui mana, i per tant al triat no li cal més legitimitat, ni científica ni tècnica, per exercir el seu poder, i si cerca avaluar, mitjançant procediments científicament incorrectes o simplistes, les conseqüències dels seus actes, llavors entendrem que polititza l'avaluació, sostraiant-la del seu àmbit científic-tècnic. Les relacions entre política, poder, ciència, tècnica, etc., que convergeixen en els procediments d'avaluació, són difícils d'aclarir. A partir del seguit de pressuposicions que explicitem un poc més endavant, ens proposem presentar de manera simplement expositiva i narrativa una sèrie de casos que, per ordre de complexitat creixent, fan referència a: (i) l'avaluació acadèmica dels estudiants universitaris, (ii) l'avaluació del professorat mitjançant les enquestes als alumnes, (iii) l'avaluació dels metges d'hospital, i (iv) l'avaluació de la pràctica mèdica.

1. PRESSUPOSICIONS SOCIALS, POLÍTIQUES, CIENTÍFIQUES I MÈTRIQUES

L'organització social humana és jeràrquica, cosa que significa que si bé hi ha rols coordinats, del mateix nivell o poder, com són ara els establerts entre companys o entre col·legues, la majoria de rols comporta una relació asimètrica

de poder entre els rols supraordinats, com el de pare o patró, i els infraordinats, com el de fill o d'assalariat. Quan les relacions són asimètriques cal que el supraordinat, qui detenta el poder, rebi informació sobre el comportament de l'infraordinat per tal d'exercitar el poder; altrament la relació no funcionaria. Aquesta necessitat peremptòria d'informació, del supraordinat sobre la conducta de l'infraordinat, es mitiga en gran part pel gran nombre de regulacions implícites, normes socials, o explícites, lleis, que delimiten el comportament de l'infraordinat. Tanmateix, per atapeïda que sigui la teranyina de les normes socials o enfarfegada la selva legislativa d'una societat complexa com la nostra, aquests controls, el social i el polític, no abasten tota la casuística de comportaments. D'ací que els poderosos necessitin, malgrat tot, avaluar els rendiments individuals en el compliment de rols dels seus infraordinats.

Parlant socialment, poderós és qui pot alterar el comportament d'altri. Tradicionalment el poder ha estat dels detentors de la força física, poder que han exercit bé sense limitacions si n'eren els únics detentors o bé amb unes limitacions com més detentors igualment poderosos són al joc social. Aquesta igualació en poder desvertebraria la societat, tot desjerarquitzant-la, si no hi hagués mecanismes de provisió del poder. Entre molts d'altres, el mecanisme democràtic comporta la delegació temporal dels rols supraordinats a partir d'un cert grau d'acord entre persones que es consideren iguals entre elles, entre individus amb rols coordinats. Allò que es delega democràticament és la capacitat per alterar el comportament d'altri, el poder, i no cap altra capacitat o qualitat que puguin tenir els individus delegadors, tals com la saviesa, la ciència o el coneixement tecnològic. De manera que el supraordinat democràticament pot equivocar-se tant com el tirà. De fet, la provisionalitat de la delegació democràtica de poder és un reconeixement prou clar de la seva manca d'immunitat a l'error.

La ciència, a diferència d'altres sistemes de pensament, té com a punt de partença l'observació objectiva i com a punt d'arribada la demostració també objectiva, i l'objectivitat no és sinó consens o acord pràcticament total entre persones que es consideren iguals entre elles, els científics. Certament, aquest ambient igualitari i de lliure discussió entre individus amb rols coordinats, que cal perquè es generi el pensament científic, és pràcticament igual a l'ambient que cal per desenvolupar l'organització democràtica d'un grup o societat. Ara bé, malgrat aquesta semblança «ambiental», el pensament científic disposa d'una metodologia pròpia i privativa.

Així, a partir de l'objectivitat, en ciència, hom elabora teories o explicacions, teories que són refusades o acceptades per la comunitat de científics, no per les característiques personals o de poder de qui les propugna, sinó per les proves empíriques i objectives que avalen cada teoria. Malauradament, massa sovint qualche científic ha abusat de la seva condició de supraordinat per fer acceptar idees que no podien ser demostrades seguint una metodologia científica. Justament per evitar aquests abusos cal extremar les exigències de

la metodologia de la ciència amb les proposicions pretesament científiques dels poderosos.

El coneixement científic de qualsevol fenomen es compendia en la seva mesura. No hi ha teoria científica sense mesura i no hi ha mesura científica que no es fonamenti en una teoria. La mesura científica és, doncs, l'expressió més nítida de l'objectivitat científica, tant de la simple observació descriptiva d'un fenomen com de les proves experimentals i controladores de les seves teories explicatives, i l'objectivitat, remarquem-ho de nou, és acord i consens inter-subjectiu. D'ací la capacitat conciliadora de la mesura científica, d'ací l'ànsia per la mesura científica dels propugnadors d'un sistema democràtic d'organització social. Del punt de vista científic, proposar-se mesurar un fenomen equival a emprendre el seu estudi amb la metodologia de la ciència. D'ací la capacitat generativa d'explicacions teòriques de les mesures científiques, d'ací l'ànsia per la mesura científica dels propugnadors de l'estudi científic de l'organització social. Ara bé, la metodologia de la mètrica dels comportaments o dels fenòmens socials no és diferent de la d'altres mesures biològiques, perquè les societats animals i el comportament dels seus membres són també l'expressió de la seva biologia. Així doncs, hem d'exigir a les mesures comportamentals els mateixos requisits exigibles a d'altres mesures biològiques, p. e. la glucèmia o la pressió arterial:

(i) Una teoria que expliqui convenientment els fonaments de la mesura, relligant científicament el fenomen a mesurar i el procediment de mesura.

(ii) Determinació dels procediments de mesura adequats per distingir entre una característica permanent, p. e. glucèmia o pressió basals d'un individu, de les variacions momentànies o situacionals; en psicometria hom distingeix en aquest sentit, entre «tret» i «estat».

(iii) Demostració que la mesura en qüestió és suficientment estable com per considerar que hom avalua una característica permanent, o tret, i per tant atribuïble a variables de l'individu avaluat i no a variables ambientals i situacionals; en psicometria hom parla, en aquest sentit, de fiabilitat test-retest de la mesura.

(iv) Demostració que la mesura en qüestió varia en funció de variables ambientals i situacionals que la teoria prediu que modificaran allò mesurat, com l'exercici modifica la pressió arterial, o la ingesta modifica la glucèmia; hom parla, en aquest sentit, de sensibilitat de la mesura.

(v) Demostració que la mesura en qüestió covaria amb d'altres procediments, com més independents i diferents millor, d'estimació d'allò mateix mesurat, com l'ús de reaccions químiques alternatives per determinar la glucèmia o la mesura de la pressió arterial mitjançant un catèter intraarterial; en psicometria hom parla, en aquest sentit, de validesa de la mesura.

(vi) Demostració que la mesura en qüestió covaria, i en el sentit predit per la teoria, amb d'altres mesures d'altres variables dins un funcionalisme més global, com la pressió arterial amb la venosa, el ritme respiratori o el cardíac,

etc., o la glucèmia amb el nivell d'insulina en sang o amb la sensació subjectiva de gana o de desanament; aquestes demostracions són les més potents científicament parlant i reben el nom, en psicometria, de validesa científica o de construcció.

(vii) Obtenció de taules de la mesura en qüestió en funció de característiques de l'individu, explícitament biològiques, com l'edat, el sexe o la constitució, o psicològiques, com la intel·ligència o la personalitat, o socials, com la classe social o els anys d'escolarització, que puguin modificar, segons els tipus de fenòmens mesurats, els resultats. Aquestes taules són imprescindibles quan hom vol diagnosticar, és a dir, interpretar uns resultats, com ara determinar la presència d'una diabetis o d'una hipertensió en una persona. Aquestes taules reben el nom, en psicometria, de barems, taules que serveixen per satisfer els requisits per interpretar els resultats o per diagnosticar a partir d'ells.

2. DOS MOLT BONS EXEMPLES

Sembla mentida l'esforç que cal fer per convèncer la gent, encara que es tracti de persones intel·ligents i formades científicament, que les mesures psicològiques o socials són metodològicament **iguals** que qualsevol altre tipus de mesures d'un fenomen físic. Aquesta dificultat d'enteniment és perquè els científics, com a col·lectiu, encara no ens hem desempallegat de l'efecte Humpty Dumpty (rellegiu, si us plau, les nostres Pressuposicions Socials i les Polítiques) en aquestes àrees de saber.

Humpty Dumpty és un personatge d'en Lewis Carroll:

«No veig que vol dir això de *glòria*», digué Alícia.

Humpty Dumpty somrigué desdenyosament. «Ni ho veuràs fins que jo no t'ho digui. Significa que "és un argument que et deixa bocabadada"».

«Però *glòria* no significa "un argument que et deixa bocabadada"», objectà Alícia.

«Quan jo uso una paraula», féu Humpty Dumpty, en un to més aviat desconsiderat, «vol dir exactament el que jo decideixo que digui —ni més ni menys».

«La qüestió és», féu Alícia, «si es pot fer que les paraules signifiquin tantes coses diferents».

«La qüestió és», féu Humpty Dumpty, «qui mana —amb això n'hi ha prou».

(Lewis Carroll, *A través de l'espill*, 1872)

Així, i per exemple, si el **rectorat** d'una universitat **decideix** que el seu professorat ha d'ésser avaluat a partir d'**enquestes fetes als alumnes** de cada professor mitjançant una variant, adaptada a la situació de classe, de sociograma (una tècnica que permet determinar les afinitats i distàncies socials que es donen entre els individus que conformen un grup), doncs així es fa **amb el silenci de la majoria dels professionals universitaris que saben que la seva**

feina és fonamentalment una activitat intel·lectual i no una tasca purament relacional amb els alumnes, que és allò que mesuren els sociogrames. És cert que una classe és un grup, com també és cert que el professor el lidera, dit en termes tècnics, i que aquest lideratge és important, però *avaluar* la feina professoral a partir just d'aquest aspecte, deixant a banda totes les implicacions intel·lectuals, fonamentals i definitives dins el rol de professor universitari, és un pur disbarat. L'efecte Humpty Dumpty és màxim quan el poderós, «qui mana», que el provoca no és un simple superior jeràrquic sinó **el superior** en tots els ordres, el **sobirà estat**. En efecte, el poder polític, qualsevol poder polític, usa el llenguatge com allò que és fonamentalment: una tecnologia social; i el poder polític usa aquesta tecnologia social per dominar, en definitiva, per constituir-se com a poder. El poder polític té múltiples vies d'influència social molt poderoses: des de les lleis fins als continguts educatius, des de les inscripcions de les monedes i bitllets de banc fins al nom dels carrers, des dels reglaments administratius fins als sistemes electorals, des del sistema de selecció dels seus funcionaris fins a la declaració de les festes laborals, etc. Encara hi ha més «maquiavel·lisme», perquè aquest conjunt de creences fabricades pel poder polític modern són presentades, pel propi poder, com a creades espontàniament per la societat. Així, i per exemple, l'«opinió pública» diu que si hom **no usa el català** per etiquetar productes o no hi ha més programes televisius en català és **per «raons de mercat»**, perquè no hi ha demanda, quan **en realitat és just al contrari**: ha estat l'esforç, la insistència i els diners de moltes persones allò que ha fet possible, per exemple, que TV3 arribi fora de les cotilles geogràfiques que el poder li ha imposat per llei, i és per llei que tots els productes que circulen per l'estat espanyol, fins i tot els importats, han d'ésser etiquetats en castellà, i és per llei, i no per raons de mercat, que tots els súbdits de l'estat espanyol han de saber castellà i no poden al·legar la seva ignorància. En canvi, és perquè tots nosaltres volem, per demanda social, que ens hem reunit a Lleida, i si els nostres congressos han estat interromputs en dues ocasions, no ha estat perquè hagués minvat el seu «mercado», sinó perquè ens els prohibien fins i tot físicament.

A les Taules 1 i 2 trobareu exemplificats, seguint les consideracions abans exposades sobre les Pressuposicions Mètriques, aquests dos casos tan nostres, un per a alguns de nosaltres —l'avaluació dels professors universitaris—, i un altre per a tots nosaltres i d'una manera molt fonda: l'avaluació del desgavell lingüístic de casa nostra. En aquest darrer cas, especialment, cal adonar-se que l'efecte Humpty Dumpty, o deformació perceptiva de la realitat social a causa del poder polític-social, és esborronador perquè, en certa mesura, encara som sota els efectes deformadors del règim hispano-feixista de Franco. Recordem que el poble alemany patí el seu feixisme durant uns 12 anys, i que l'italià en patí el seu durant uns 20 anys. Doncs bé, nosaltres hem patit un feixisme aliè, l'espanyol, durant més de 40 anys! Imagineu com ens ha deformat la visió de la nostra pròpia realitat social, un poder polític, que com tots els feixistes, era

TAULA 1. Avaluació del professorat universitari

	PSEUDOTEORIA	TEORIA CIENTÍFICA
AFIRMACIONS GENERALS	L'aprenentatge acadèmic no és sinó el resultat d'una relació social especial entre les característiques personals del professor i les necessitats evolutives de l'alumne.	La tasca principal d'un professor és tant la transmissió de nova informació a l'alumne com la reorganització d'aquesta nova informació amb la que ja disposa aquest.
DEDUCCIONS QUE PODRIEN VERIFICAR-SE O S'HAURIEN DE VERIFICAR	La bondat d'un professor s'avalua mitjançant una mesura sociomètrica de la seva persona feta pels seus alumnes. En la selecció dels professors cal primar les seves potencialitats afectives i actitudinals. El rendiment acadèmic dels alumnes depèn fonamentalment dels aspectes emotius i relacionals del professor. No es pot aconseguir un bon rendiment acadèmic amb procediments «deshumanitzats». A més companyonia entre professor i alumne, més aprenentatge. L'aprenentatge depèn més del docent que del mètode. Motivar i interessar l'alumne per la matèria és un dels objectius fonamentals. Cal avaluar l'alumne en funció del seu esforç i motivació.	L'eficàcia d'un professor s'avalua examinant els coneixements i la capacitat de síntesi que ha generat en l'alumne. Cal primar en la selecció dels professors les seves capacitats intel·lectives i aptitudinals. El rendiment acadèmic de l'alumne depèn fonamentalment de la intel·ligència del professor. L'ensenyament programat per ordinador és un bon procediment per transmetre i reorganitzar coneixements. La companyonia no necessàriament augmenta l'aprenentatge, sinó que en certs casos el pot perjudicar. L'aprenentatge depèn més del mètode que del docent. Assolir un determinat nivell de coneixements de la matèria és un dels objectius fonamentals. Cal avaluar l'alumne en funció dels coneixements adquirits.
PRODUCTE FINAL	PROFESSORS AGRADABLES I SIMPÀTICS, QUE «CAUEN» BÉ ALS ALUMNES	PROFESSORS DURS I EXIGENTS, QUE NO NECESSÀRIAMENT «CAUEN» BÉ ALS ALUMNES

xenòfob, menyspreador de la intel·ligència i de tots els seus productes i valors (ciència, tecnologia, cultura, etc.), i que semblava tenir com a única obsessió l'espanyolitzar-nos.

Les taules resumeixen aquests dos casos avaluatius paradigmàtics, el dels professors universitaris (Taula 1) i el de la nostra situació lingüística (Taula 2),

TAULA 2. Avaluació de la nostra situació lingüística

	PSEUDOTEORIA	TEORIA CIENTÍFICA
AFIRMACIONS GENERALS	Les llengües tenen entitat i vida pròpia, de manera que n'hi ha de poderoses, d'elegants, de cultes, d'intel·ligents, etc.	La llengua és un comportament complex, causa i efecte de la socialització, procés determinat per les relacions socials.
DEDUCCIONS QUE PODRIEN VERIFICAR-SE O S'HAURIEN DE VERIFICAR	Les llengües s'estimen, així que només individualment es pot fomentar el seu apreci i la seva adopció. Defensar una llengua consisteix a fer veure les seves virtuts a fi de promoure un lligam afectiu entre la persona i la llengua. Diverses llengües poden conviure dins una mateixa societat, especialment quan són germanes i estan ben repartits els papers entre elles. El castellà s'imposa al català per la seva superioritat en tots els camps. Hi ha llengües glotòfagues que creixen perquè s'alimenten d'altres llengües. El castellà s'aprèn fàcilment, sense coaccions, sense esforç, perquè és present ambientalment pertot arreu. Els catalans «educats» usen el castellà amb els castellano parlants, com si els agradés compartir amb ells l'agradable «sonoritat» del castellà. La llengua castellana és més sòlida i apta per a la vida, mentre que la catalana és més intimista i evocadora d'afectes familiars. El castellà, com a estructura lingüística, és «més perfecte» que el català perquè s'escriu quasi com es parla.	Les llengües s'usen, de manera que només col·lectivament hom pot fer propalar i fer prevaler el seu ús. La defensa d'una llengua és un afer col·lectiu i és una de les tasques irrenunciables dels representants de la societat: els polítics. L'ús social de dues llengües dins un mateix territori indica la barreja de societats diferents: zones frontereres o una societat dominada per una altra. La legislació espanyola no sols imposa l'ús del castellà, sinó que en frena el del català. Els estats sobirans gasten recursos ingents en la llengua pròpia, tant dins com fora de les seves fronteres. El jovent no aprèn anglès malgrat viure amarat d'aquesta llengua, en canvi una forta estructura funcional fa que de ben petits aprenguin castellà a l'escola. Després de molts episodis de càstig, els catalans han estat educats a «girar la llengua» en conversar amb un castellanoparlant, en un acte de submissió apresca. La «minorització» política i legal a què són sotmesos els catalano parlants fa que associïn el català amb l'arraconament familiar i no amb les situacions de poder. Els catalans aprenen el castellà literari, una forma escrita molt invariant, mentre que, per a ells, el català és, lògicament, una de les seves múltiples variants orals.
PRODUCTE FINAL	ACCIONS VOLUNTARISTES I POSSIBILISTES	LLUITA LEGAL I POLÍTICA

confrontant una visió pseudocientífica amb una de científica. La pseudocientificitat és determinada per la falsedat empírica de les deduccions de la pseudoteoria, com ara la impossibilitat de provar que les llengües glotòfagues tinguin un «estómac» o un cert aparell digestiu. Hem de remarcar que, per simplificar la Taula 2, amb «castellà» volem referir-nos no sols a l'espanyol, sinó també al francès, i que, també per simplificació, les referències a l'organització política del Regne espanyol també es refereixen a les de la República francesa. Val a dir que hauríem d'exceptuar, dins aquestes correspondències hispano-franceses, les relatives a l'ús oral o escrit de les llengües: als països espanyols —els estats de llengua oficial espanyola— els poders públics fomenten primordialment un ús *oral* de la seva llengua, mentre que l'estat francès des de sempre ha fomentat també, i amb molta intensitat, l'ús *escrit* de la seva llengua. Així, si parlem de l'edició de llibres, seria fals dir «les idees publicades en llibres en castellà tenen una difusió més universal que les publicades en llibres catalans» perquè «les tirades dels llibres en castellà són molt magres —per exemple, en relació a l'anglès—, i més de la meitat dels exemplars són consumits a països de parla catalana». Canviant, en aquest cas suara esmentat «castellà» per «francès», els resultats no serien els mateixos.

3. ESTUDI DE CASOS

Com ja hem comentat, el nostre propòsit és simplement narrar les vicissituds que hem experimentat i les dificultats que hem trobat en l'ús i la defensa de l'avaluació científica que, com pot imaginar-se el lector, han provingut totalment de la interferència de diversos nivells de poder social. De la gran majoria d'aspectes narrats, tenim proves documentals, encara que no les mostrarem per no enfarfegar la narració. Quant a les persones implicades, no les esmentarem pel seu nom sinó pels seus càrrecs i això no sols per motius de cortesia o amistat, de fet amb moltes d'elles mantenim una llarga i fonda companyonia, sinó per la irrellevància de les seves particulars individualitats: és el càrrec que proporciona el poder i és el càrrec, com a peça d'una organització més extensa, el que determina en molta gran part l'actuació de la persona concreta. Fins i tot nosaltres, els autors d'aquest escrit, potser hauríem actuat igual en les mateixes circumstàncies de poder. No volem jutjar persones sinó fer que reflexionem, a partir de casos ben reals, sobre la capacitat distorsionadora del poder sobre la percepció i l'avaluació objectiva, ço és, científica, de la realitat.

3.1. L'avaluació acadèmica dels estudiants universitaris

Respecte a l'avaluació dels coneixements teòrics, de sempre s'ha sabut que els exàmens dits d'assaig o de redacció, aquells que consisteixen a desenvolupar un tema, són de difícil qualificació objectiva, dificultat que comparteixen amb els anomenats exàmens orals. Són ben conegudes les experiències que si

un examen d'assaig és qualificat per diferents professors del tema, o pel mateix professor però en moments diferents, els resultats són ben dispers. Dit tècnicament, els exàmens consistents a desenvolupar un tema, tant si són escrits com si són orals, no posseeixen prou fiabilitat, ni fiabilitat de corrector ni fiabilitat entre correctors; ras i curt, vulneren la més elemental de les pressuposicions científiques: l'objectivitat. Aquesta manca d'objectivitat genera en l'examinador una gran inseguretat a l'hora de qualificar, i per això hom ha adoptat una escala molt primària d'avaluació: primer una dicotomia nominal entre «Suspens» («No apte», etc.) i la superació dels mínims exigibles, superació que és graduada mitjançant una ben simple escala ordinal: «Aprovat», «Notable» i «Excel·lent». En conjunt és, doncs, una escala ordinal, si considerem el «Suspens» com el graó mínim, que, com totes les escales ordinals, no manté iguals les distàncies mètriques entre els diversos graus, és a dir, no hi ha la mateixa distància entre, per exemple, l'«Excel·lent» i el «Notable» que entre el «Notable» i l'«Aprovat». Aquests tipus d'avaluacions, per tant, no discriminen adequadament i per consegüent perjudiquen els avaluats, bé els superiors o bé els inferiors, per no haver-hi la mateixa proporcionalitat al llarg de tota l'escala entre grau de la característica qualificada i nota. Aquest sistema injust i obsolet, encara és l'oficial a les nostres universitats, cosa que pot indicar llur «grau» de desenvolupament científic, si més no a l'hora de qualificar llurs estudiants.

És obvi que quan un professor llegeix un examen d'assaig el seu judici és influït, i molt, pel significant, pel suport del missatge lingüístic: cal·ligrafia, tipus de lletra, disposició i disseny formals del full, ortografia, sintaxi, etc. Consemblantment succeeix a l'examen oral amb l'aspecte físic, el to i timbre de veu, l'actitud corporal, la gesticulació i la mirada, etc. Els exàmens d'elecció múltiple (EEM) eviten aquests esbiaixaments de judici i per això mostren una perfecta fiabilitat tant de corrector com entre correctors. Això no vol dir, però, que aquests tipus d'examen siguin adequats per avaluar la creativitat ni d'altres components que no siguin els coneixements i la seva relació entre ells o raonament, sempre i quan els ítems estiguin ben construïts. I d'EEM, per descomptat, n'hi ha de bons i de dolents; que siguin un bon sistema per avaluar coneixements no significa que **tots** els exàmens amb aquest format siguin correctes científicament parlant: hi ha tota una disciplina, la docimologia, dedicada a l'estudi de llur adequació mètrica. En definitiva, si hem d'avaluar coneixements, els EEM són els millors, sempre i quan siguin ben construïts. Però això no vol dir que d'altres aspectes que els **EEM de coneixements** explícitament no avaluen, com ara els enunciats anteriorment (ortografia, cal·ligrafia, tipus de lletra, etc.), no siguin ben importants per a la pràctica mèdica. Paga la pena remarcar, finalment, que els EEM proporcionen una escala d'interval, una escala amb distàncies mètriques iguals al llarg de tota l'escala, cosa que no perjudica, per discriminar diferentment, cap tipus d'estudiant, ni els bons ni els dolents.

A partir del curs 1980-1981 hom engegà, al «departament» de Psicologia Mèdica de la Facultat de Medicina de la Universitat Autònoma de Barcelona (UAB), una decidida política de millora dels sistemes d'avaluació. Amb aquesta política hom pretenia fer una feina útil per al conjunt de tots els professors i estudiants, i fer una tasca tècnicament impecable i modèlica per fer comprendre a la resta de professors que la psicologia no era el regne de la xerrameca sinó una disciplina tan rigorosa com les altres. Aquesta política, que fou assumida amb entusiasme pel degà, F. González Sastre, tenia com a primera meta la creació d'un Programa per a l'Avaluació d'Exàmens per Ordinador (PAEPO) com a eina imprescindible per a la creació d'un Servei d'Avaluació Acadèmica (SAC), servit per un professional de la psicometria amb dedicació exclusiva, amb els objectius següents: *a)* assessorament i difusió dels sistemes òptims d'avaluació acadèmica, en funció del tipus de docència, matèria, etc.; *b)* estudis sobre els sistemes d'avaluació usats per la Facultat amb els seus propis alumnes (exàmens de llicenciatura), o que classifiquen globalment la Facultat (exàmens MIR i semblants); *c)* avaluació total, des de la lectura dels fulls de resposta fins a un informe escrit, dels EEM fets a la Facultat, mitjançant el PAEPO; *d)* suggeriments als professors perquè puguin millorar els EEM a partir de les avaluacions de les preguntes usades; *e)* creació de bancs, per assignatures i programes, de preguntes d'examen depurades i equivalents, cobrint tota la matèria; *f)* confecció automàtica, a partir dels bancs anteriors, d'exàmens de longitud i fiabilitat adequades per avaluar les diverses parts del programa d'una assignatura; *g)* recerca i creació de sistemes per objectivar la vàlua dels exàmens d'assaig i els exàmens pràctics.

Aviat el marc de la Facultat resultà massa estret i la política engegada es dirigí a tota la universitat. La idea central era la creació, mitjançant el PAEPO, de bancs, classificats per assignatures o matèries, de preguntes d'EEM que complissin perfectament tots els requeriments psicomètrics exigibles. Arribar en aquesta meta comportava assolir diversos objectius previs: fomentar l'ús dels exàmens objectius, crear la necessitat de l'avaluació científica dels exàmens, obtenir suports, informàtics i administratius, per a les tasques programades, etc. En veritat, la finalitat científica era ultrapassada per d'altres beneficis que se'n derivarien: *a)* la comparabilitat de les notes entre assignatures, fins i tot de facultats diferents, amb totes les conseqüències: concessió de beques, de premis, etc.; *b)* l'homogeneïtzació de la credibilitat i validesa socials dels títols expedits; *c)* l'agilitació i estandardització dels procediments d'examen, amb clars avantatges administratius i d'entrenament generalitzat per als examinands; *d)* la possibilitat de confegir instantàniament, per a estudiants concrets i a qualsevol moment, un examen mitjançant una tria aleatòria de qüestions tretes del banc de preguntes; *e)* la generació de cooperació, de discussió i d'integració teòrica entre diferents professors per mor de l'objectivació de les preguntes; *f)* l'explicitació, en termes ben formals, dels objectius docents; *g)* la possibilitat d'avaluar no sols l'aprofitament dels

discents sinó també els cursos, els mètodes i la pròpia tasca dels ensenyants: avaluació del grau d'acompliment dels objectius docents i avaluació d'una de les activitats més rellevants dels professors, la d'examinar. És a dir, avaluant els exàmens que posa un professor l'avaluem doblement: tant pel rendiment que obtenen els seus alumnes com per la qualitat de la feina feta en confeigir l'examen.

La pràctica dels «exàmens»

Amb les finalitats de sensibilitzar els professors de la Facultat i d'introduir progressivament el rigor en l'avaluació acadèmica, ens proposàrem ensenyar als nostres propis alumnes de Psicologia Mèdica a avaluar exàmens. Era una idea ben «subversiva», en tot el sentit etimològic del mot, aquesta d'iniciar els nostres alumnes en l'avaluació científica dels exàmens. Amb aquesta finalitat, hom edità una monografia pràctica sobre avaluació del rendiment acadèmic (Pérez i Torrubia, 1981) i es programà una sessió de pràctiques, d'unes 4 hores de duració, amb el següent contingut:

- a) Construcció d'una prova d'assaig.
- b) Construcció d'una prova d'elecció múltiple.
- c) Administració de les dues proves i correcció.
- d) Càlcul de l'objectivitat de la prova d'assaig.
- e) Anàlisi d'ítems de la prova d'elecció múltiple.
- f) Càlcul de la fiabilitat Kuder-Richardson i coeficient de dificultat de tota la prova d'elecció múltiple.

Els alumnes eren dividits en dos grups: els «estudiants» anaven al bar mentre els «professors» preparaven dos exàmens, un d'assaig i un altre d'elecció múltiple, sobre coneixements de bàsquet. Després els «estudiants» responien els dos exàmens i, posteriorment, «estudiants» i «professors» avaluaven conjuntament els diversos exàmens generats. Aquesta pràctica era molt valorada pels nostres alumnes per raons òbvies. Tanmateix només es féu durant tres cursos seguits. Durant el tercer curs d'impartició d'aquesta pràctica, 1983-1984, un professor de clíniques demanà, a la secció de pres d'una junta de Facultat, «què c... ensenyaven els de Psicologia als alumnes sobre avaluació d'exàmens?» Explicà que després d'un examen d'assaig de la seva matèria, on sembla que es carregà molta de gent, el delegat de curs li demanà públicament quina era la fiabilitat de l'examen en qüestió. El petit aldarull de protesta dels estudiants es reproduí, menor i en sentit contrari, a la junta de Facultat. Llavors decidírem de suprimir la pràctica dels «exàmens», no per manca de combativitat o per deferència cortesa als nostres companys de Facultat, sinó perquè tampoc no havíem obtingut comprensió i suport dins el marc general universitari per al nostre projecte d'avaluació automàtica d'exàmens.

El PAEPO (Programa d'Avaluació d'Exàmens Per Ordinador)

Durant un parell d'anys ens dedicàrem, juntament amb el matemàtic i informàtic Joan Llopart, a la confecció d'un programa capaç d'avaluar ràpidament i psicomètrica qualsevol EEM (Garcia i cols., 1984). Si bé la Facultat encoratjà tot el projecte del SAC, no disposava dels mitjans per a la seva operativitat. En Llorenç Guilera, llavors director del Centre de Càlcul de la UAB, ens facilità la col·laboració dels informàtics Eugeni Gil i Manuel Peña per implementar el sistema de lectura de fulls d'examen i el PAEPO al VAX-11, llavors el «cervell» del Centre de Càlcul. El PAEPO implementat feia les operacions següents:

- 1) Lectura òptica o per teclat dels fulls de resposta dels examinands.
- 2) Introducció per teclat o lectura òptica de les plantilles de les diverses combinatòries de l'examen.
- 3) Correcció segons les diverses plantilles.
- 4) Estudi de la fiabilitat total de l'examen i de la validesa de cada qüestió en particular.
- 5) Càlcul iteratiu de la màxima fiabilitat total assolible amb els ítems de l'examen, com a funció del nombre d'ítems i el rebuig dels ítems psicomètricament més dolents.
- 6) Càlcul, amb menú d'opcions, de la puntuació de cada examinand en funció de la plantilla, la validesa de les qüestions, els ítems no rebutjats i les respostes a l'atzar.
- 7) Llista dels ítems de l'examen amb els índexs de dificultat i de discriminació i la freqüència de respostes de cada alternativa.
- 8) Llista alfabètica dels examinands amb llur nota bruta, corregida i decimal.
- 9) Informe escrit i estandarditzat sobre l'examen en qüestió, tot mirant de transmetre una informació psicomètrica complexa (forma de la distribució, fiabilitat interna, discriminabilitat, etc.) d'una manera fàcil i entenedora per a un no expert en psicometria.

Les dues innovacions del PAEPO, que no tenien d'altres programes semblants, com l'usat llavors a la Universitat Politècnica de Barcelona (ara UPC), eren la creació del millor examen possible amb les preguntes posades i l'informe global final. Mitjançant la iteració del punt 5), llevat de casos molt extrems, hom sempre podia disposar del millor examen possible amb les preguntes proposades. L'informe final, creat mitjançant algorismes que ara en diríem d'«intel·ligència artificial», assegurava la tecnificació progressiva dels professors en una matèria molt àrida però imprescindible per als bons professionals de l'ensenyament universitari: la docimologia.

Tot era a punt per al gran moment, només calia el suport i l'impuls de les nostres autoritats acadèmiques. Si bé el rector de llavors considerà molt interessant la proposta d'anar generalitzant l'avaluació de tots els exàmens

fets a la UAB, finalment el vicerector corresponent ens digué simplement NO. De primer vàrem pensar que no havia entès perfectament el nostre projecte. «I tant que ho entenc!», replicà. Segons ell, el projecte al voltant del PAEPO trauria el poder als professors, així de clar ho va dir. «Chapeau!» Efectivament, i en termes tècnics, el professor de la nostra universitat té dos poders simultàniament: el poder d'expert, aquella capacitat de convocatòria que tenen els qui saben una matèria i la volen ensenyar, i el poder coercitiu, de funcionari que pot aprovar o suspendre. Tots nosaltres, com a professors, ens il·lusiona de pensar que els alumnes ens vénen a classe perquè volen aprendre; la realitat crua, però, és que la majoria d'ells hi vénen per aprovar. Aquesta capacitat, d'aprovar o no, és exercida mitjançant l'examen. Si llevem aquesta arma al professor, aquest resta, en molts dels casos, completament indefens, sense cap altre poder, sense capacitat ni per manar dins l'aula. El vicerector s'estimava els professors. Nosaltres trobarem que el poder es justificava, una vegada més, mitjançant una pseudoavaluació, avortant una avaluació científica d'una realitat, si més no aparentment, senzilla i clara: l'avaluació acadèmica dels estudiants universitaris.

Molt posteriorment, els autors de la monografia pràctica sobre els exàmens, i amb el suport de l'ICE (Institut de Ciències de l'Educació) de la UAB, engegaren a partir del 1992-1993 un curs anual sobre «Avaluació del rendiment acadèmic», curs de nou hores de durada adreçat als professors de la UAB. L'esquema del curs es basa en la pràctica dels «exàmens» esmentada. Molt a poc a poc aquest curs va canviant les idees que el professorat té sobre els exàmens. Els seguidors del curs hi aprenen els aspectes tècnics de l'avaluació acadèmica, però per a molts d'ells més que un curs de «continguts» sembla una sèrie de sessions de canvi d'actituds, canvi aconseguit mitjançant demostracions empíriques i personalitzades de la manca de rigor de les avaluacions acadèmiques «tradicionals» i el debat obert sobre el tema, de manera que els alumnes-professors amb falses preconcepcions sobre la matèria abandonen els seus prejudicis. Al ritme actual, uns vint professors per curs, cap a l'any 2125 hi hauran passat tots els professors actuals de la UAB.

La pràctica dels «exàmens» ha estat reintroduïda durant el curs 1996-1997 dins el programa de l'assignatura, quinze anys després de la seva estrena i dotze de la seva supressió. La seva reintroducció ha estat feta dins una programació general sobre les múltiples facetes de l'avaluació mèdica: selecció d'aspirants a estudiar medicina, sistemes de qualificació acadèmica, avaluació de les habilitats pràctiques, sistemes d'acreditació, etc. Els alumnes ja han començat a «contestar» amb argumentacions tècniques els exàmens de l'assignatura. El projecte del SAC acabà foradat i al cap de sis anys, durant el curs 1986-1987, hom creà el Gabinet Tecnicopedagògic, sense assignació de personal, amb l'objectiu fonamental de millorar l'ensenyament de la medicina a la Facultat de la UAB.

3.2. L'avaluació del professorat mitjançant les enquestes als alumnes

A la majoria d'universitats catalanes s'administren enquestes per avaluar el professorat. Què pretenen mesurar les autoritats acadèmiques? Què mesuren realment les enquestes? Quina utilització es fa o es farà dels resultats? Les enquestes compleixen els requeriments psicomètrics propis d'aquests tipus de mesures? Per analitzar-ho, hem triat el cas de la UAB, que és el que coneixem més de prop. Aquesta història no és un fet únic i aïllat. Llevat d'algunes peculiaritats, el procés podria ser més o menys similar en altres universitats.

Sobre la preocupació per la qualitat de l'ensenyament o de com fent enquestes als alumnes podem llevar-nos de sobre la mala consciència

A finals de la dècada de 1980 es registrà a la UAB un moviment de preocupació per la qualitat de la docència que cristal·litzà en el document «**La Universitat Autònoma de cara als anys 90**». En aquest document es proposaven un seguit d'actuacions a càrrec d'experts amb l'objectiu d'obtenir un informe fiable de la situació de la docència, una formulació d'indicadors de qualitat docent individual i col·lectiva, i uns procediments d'avaluació de la tasca docent en funció d'aquells indicadors. De forma genèrica se suggerien actuacions tals com iniciar una campanya de sensibilització adreçada als professors, departaments i centres; revisar la infraestructura de l'ensenyament i completar els serveis de suport a la docència; crear un servei tecnicopedagògic; i fomentar les estades de professors a l'estranger.

Fruit d'aquell estat d'opinió, el Claustre General de la UAB, amb data 13 de març de 1992, va encomanar al vicerector d'afers acadèmics i al vicerector d'investigació i de professorat la dinamització del procés de discussió sobre el tema i la creació d'una comissió específica. Amb aquest objectiu, es redactà un «**Document-marc per a la reflexió sobre la instauració d'un sistema d'avaluació i estímul de la qualitat de la docència**» que fou enviat a tots els departaments i facultats.

En el document es proposaven dos nivells d'anàlisi: l'avaluació individual del professor i l'avaluació de la titulació. A tall orientatiu s'hi suggeria un ampli ventall de possibles indicadors interns: l'informe sobre l'avaluació global dels semestres de cadascun dels estudis; l'informe dels tutors i dels coordinadors de titulació; l'anàlisi de la permanència del professor; les avaluacions per col·legues; els projectes d'innovació docent; les relacions docència / recerca; i les enquestes als alumnes. Com a possibles indicadors externs es proposaven: l'avaluació per externs (col·legis professionals, etc.); l'avaluació per ex-alumnes; l'anàlisi del servei de borsa de treball i de les diverses vies d'accés a la professió (MIR, oposicions, etc.); l'anàlisi de dades de preinscripció en primera i segona opció; i d'altres, com mirar indicadors de tercer cicle i de cursos de postgrau.

Pel que fa als mètodes d'avaluació, s'hi recomanaven entre d'altres: enquestes i informes elaborats pels alumnes i redactats pels delegats de classe, informes de tutors, informes de caps de titulació i/o de les comissions d'avaluació de centre, dades objectives referides a l'alumne, a les instal·lacions, al nombre de professors i de PAS (personal d'administració i serveis), enquestes externes i avaluacions externes. S'estimava convenient involucrar l'ICE en la confecció de les enquestes (especialment en el vessant de com determinar si un professor és bo o no).

La discussió del «**Document Marc**» en els departaments i centres, i dins de la mateixa comissió, donà lloc a l'informe «**La UAB i la millora de la qualitat de la docència: Propostes de mesures al Claustre General**». En aquest document, que va ser aprovat el 28 de maig de 1993, s'hi definien tres nivells d'avaluació: els estudis, els departaments i el professorat, i s'hi argumentava que l'avaluació, ja sigui individual o col·lectiva, ha d'afectar els sistemes de promoció i/o la distribució de recursos.

En el capítol de l'avaluació dels estudis es proposaven un seguit de factors a mesurar com ara l'adequació dels currículums, els programes, la innovació docent, els mitjans humans, l'estructura horària, l'adequació dels nivells dels estudiants i l'avaluació del conjunt dels estudis. Pel que fa als departaments, els suggeriments eren molt poc precisos.

Quant a l'avaluació del professorat, en l'apartat «**què s'avalua**», no es feia cap especificació sobre quin havia d'ésser l'objectiu de l'avaluació i només hom precisava que «cal redefinir amb claredat les obligacions docents i establir mecanismes clars per fer-ne el control i l'avaluació». Pel que fa al «**com s'avalua**», es proposava tenir en compte els resultats de les enquestes semestral als estudiants, la informació individualitzada del coordinador de titulació i el coneixement del grau de compliment de les obligacions mínimes de classes i tutories, que hauria de ser subministrat pels equips responsables dels diferents centres. Sobre «**qui avalua**», es proposava, per fer-ho, la Comissió de Professorat, delegada de la Junta de Govern.

En un altre apartat del document, s'hi deia que l'enquesta hauria de tenir com a objectius: a) conèixer l'opinió dels usuaris de l'ensenyament —no s'especificava sobre què; b) fer-ne un element de reflexió, de l'anàlisi semestral i anual que el coordinador de la titulació ha d'elaborar; c) constituir una memòria de la docència anual des del punt de vista dels usuaris, en què es disposés de la radiografia de la qualitat de la docència; i d) donar elements al professorat per tal de millorar la seva docència.

Breu, d'una declaració de bones intencions, en el primer document, es va passar, en el segon, a la proposta ambiciosa d'un ampli conjunt d'indicadors de qualitat. La concreció de tot allò en un tercer document acabà, emperò, en una proposta «pragmàtica» d'avaluació del professorat, basada en les enquestes dels estudiants (per valorar quelcom encara no definit), els informes dels coordinadors de titulació, i el grau de compliment o no de les obligacions. Les

enquestes als estudiants permetien tranquil·litzar les consciències dels universitaris preocupats per la qualitat de la docència i es convertien així en una peça clau de l'engrenatge polític-burocràtic-avaluatiu de la qualitat de l'ensenyament.

Avalua que fa fort

La primera enquesta: un intent sense continuïtat. Com que calia passar a l'acció el més aviat possible, la força implacable de l'aparell administratiu es va posar en marxa. Durant aquell mateix curs, de manera paral·lela a tot el procés, es procedí a la confecció i a l'administració de la primera enquesta generalitzada d'avaluació del professorat: l'Enquesta d'Avaluació de la Docència (EAD). Promoguda per l'equip rectoral, constava de nou ítems amb escales de cinc opcions, i va ser realitzada per l'Oficina d'Avaluació i de Programació Acadèmica (OAPA). Els resultats individuals van ser enviats als avaluats, i els resultats col·lectius foren lliurats als directors de departament i als responsables de titulacions. Just acabat aquest primer intent d'avaluar els professors mitjançant els estudiants, ja fou objecte de contundents crítiques metodològiques des d'alguns sectors de professionals de l'avaluació, com ara la de Garcia Sevilla (1994), i aquesta primera enquesta no va tenir continuïtat el curs següent probablement a causa del canvi d'equip rectoral.

La segona enquesta: posada a punt de l'instrument. El 19 de maig de 1994, amb nou equip al rectorat, es crea la Comissió d'Avaluació i de Control del Professorat. El 17 de novembre de 1994, la Junta de Govern acorda aprovar la modificació del procés d'avaluació de l'actuació docent mitjançant l'enquesta als estudiants. En l'informe al Claustre del 13 de desembre de 1994 es comunica que aquest acord és en fase d'implementació i que l'avaluació del curs 1994-1995 és experimental. Durant aquell curs es procedeix al desenvolupament d'una nova enquesta anomenada «Avaluació de l'Actuació Docent Mitjançant l'Opinió dels/de les Estudiants». El qüestionari constava de tretze ítems i diverses dades d'identificació mostral i estructural. Es van administrar 4384 enquestes vàlides. L'objectiu, com hem remarcat, durant aquell curs «no era avaluar el professorat sinó posar a punt els instruments d'avaluació».

Línies generals de política de professorat. Període 1995-1998: Cal mantenir les enquestes. El Claustre General de 4 de desembre de 1995 aprovà un document sobre «Línies generals de política de professorat. Període 1995-1998». En l'apartat 4.3.1 sobre avaluació del compliment docent, l'avaluació i l'incentiu de la qualitat de la docència, es reafirmaven les línies marcades en el document aprovat pel Claustre el 28 de maig de 1993 i es destacava la necessitat de «mantenir el procés d'avaluació de la docència mitjançant l'enquesta als estudiants actualment en vigor». També es manifestava: «Tot i que hem de mantenir-la, és obvi que l'enquesta és només un indicador parcial, excessivament centrat en l'aspecte individual de la docència. Cal perfeccionar

l'avaluació individual, complementant-la amb altres indicadors; però, sobretot, cal avançar cap a un dispositiu d'avaluació de la **qualitat dels ensenyaments** (aspecte **col·lectiu** de l'avaluació de la docència)» (pp. 21–22). Aquest dispositiu d'avaluació hauria de ser efectiu, en fase pilot, al llarg del curs 1996-1997.

Malgrat les bones intencions sobre la necessitat d'abordar l'avaluació de la qualitat dels ensenyaments des d'un punt de vista més global, tot sembla indicar que tot el «bagatge avaluatiu» de què es disposava a la UAB a desembre de 1995 eren dues enquestes: una ja «desestimada», la realitzada per l'anterior equip rectoral, i una de nova, «amb un format senzill i metodològicament vàlid», segons el nou equip rectoral. No hi havia encara evidència de cap tipus d'avaluació de la qualitat de les titulacions ni de la qualitat dels departaments (a nivell docent). Però el més greu de tot és que en aquell document de política de professorat tampoc no quedava definit què es pretenia avaluar amb les enquestes als estudiants. Per posar-ne just un exemple ben bo d'entendre: no és el mateix avaluar la qualitat de la docència, que la qualitat del docent, o que la satisfacció dels usuaris. Tampoc no s'avançava en la identificació d'altres indicadors individuals. Malgrat totes les fortes crítiques alienes i les greus mancances reconegudes, el missatge de les autoritats acadèmiques era molt clar: calia perseverar en l'administració de les enquestes.

L'enquesta definitiva: a la tercera va la vençuda? Paral·lelament a l'aprovació d'aquest document, el dispositiu administratiu continuà funcionant de manera implacable. Començà a administrar-se una versió «depurada» de l'enquesta «experimental» formada per set ítems sobre les classes de teoria i set sobre les de pràctiques. Durant aquell curs es processaren un total de 90.124 (sic!) enquestes, i aquesta mateixa enquesta s'ha continuat administrant durant el curs 1996-97. La raó és molt senzilla: **les enquestes unipersonals dels professors fan forts aquells que en tenen els resultats i els poden utilitzar amb finalitats político-administratives.**

Un bon vernís pseudocientífic pot ser útil per donar credibilitat als actes político-administratius

El desembre de 1995 l'Assemblea de professors de la Unitat Docent de Bàsiques (Bellaterra) de la Facultat de Medicina de la UAB elabora, per unanimitat, un escrit de protesta per la forma com s'havien administrat les enquestes aquell trimestre i per la manca d'informació científica sobre tot el procés. En aquell document se sol·licita al vicerector de professorat informació sobre fiabilitat, validesa i normes d'interpretació de l'enquesta administrada durant les darreres setmanes (Unitat Docent de Bàsiques de la Facultat de Medicina, UAB, 1995).

El febrer de 1996 es rep la resposta del vicerectorat on s'adjunta l'informe tècnic de construcció de l'enquesta definitiva realitzada durant el curs anterior (Vicerectorat de Professorat, UAB, 1996). L'informe estava elaborat a partir de

la versió experimental de tretze ítems i no de la versió definitiva administrada durant el curs 1995-96. En el document es deia, entre altres coses, que «(el qüestionari)... s'ha centrat, només en l'aspecte del grau de satisfacció de l'usuari de la docència» o que «les comparacions entre titulacions no tenen cap sentit» i que «la nostra enquesta no avalua la qualitat de la docència ni, tan sols, la qualitat del docent. Per avaluar aquestes dimensions calen molts més paràmetres». També s'hi deia que «l'avaluació "individualitzada" és una etapa que cal superar per avançar cap a l'avaluació de la qualitat de les titulacions. De moment, però, és una etapa necessària encara».

Pel que fa a les dades tècniques, cal remarcar de l'informe del vicerectorat alguns aspectes metodològics ben preocupants: *a)* un nombre molt elevat de respostes en blanc; *b)* un gran nombre d'individus o d'enquestes que no quadraven entre les diverses taules; *c)* una notable absència d'informació sobre les dades bàsiques dels responents com edat, sexe, grup, convocatòria, etc.; *d)* sis de les tretze escales no presentaven cap resposta, entre milenars d'enquestes!, en l'opció 5 (central); i *e)* una bona part de l'informe es dedicava a justificar la bondat de la mostra emprada, però s'oblidava d'allò més important: l'estudi de la validesa.

Quant a la fiabilitat, només es proporcionaven dades referides a la consistència interna entre els diversos ítems, però no s'informava de cap procés de verificació de dades, del control de les tendències de resposta del respondent, de l'avaluació de la «mala fide», ni del nivell d'estabilitat temporal. Referent a la validesa, es presentava com a única evidència la correlació de cadascun dels ítems de l'escala amb l'ítem 13 de la mateixa escala. Cal remarcar que aquesta correlació no és res més que una altra forma de calcular la consistència interna. Qui estigui interessat en aquests aspectes tècnics pot consultar: Unitat Docent de Bàsiques de la Facultat de Medicina, UAB (1995, 1996) i Vicerectorat de Professorat, UAB (1996).

La fiabilitat és una condició necessària però no suficient per verificar la validesa d'una mesura. En el cas que ens ocupa, l'evidència empírica sobre la fiabilitat era escassa, però, el que és pitjor, l'evidència sobre validesa era inexistent. Com ja hem dit molt més amunt, l'anàlisi de la validesa d'un instrument de mesura implica necessàriament: *a)* la definició operativa de l'objecte de mesura, *b)* la demostració de la seva fiabilitat, *c)* l'elaboració de prediccions sobre la seva relació amb altres variables, i *d)* la comprovació del grau de compliment de les prediccions. No podem concloure que un instrument és vàlid per mesurar un constructe només perquè els autors hagin tingut la intenció de mesurar-lo en construir-lo, o pel fet d'haver-li posat un nom determinat: cal demostrar-ho. Podríem afirmar que l'enquesta en qüestió mesura simpatia, si li haguéssim posat el nom «Enquesta de simpatia del professorat»? Podem afirmar que l'enquesta mesura satisfacció de l'usuari just perquè potser aquesta era la intenció que tenien els autors en construir-la?

Malgrat la baixa qualitat de les dades tècniques facilitades, aquesta resposta del vicerectorat ens va permetre veure per escrit per primera vegada

l'objectiu de l'enquesta: avaluar la satisfacció de l'usuari. En canvi, en el qüestionari que es lliurava (i que es continua lliurant) als alumnes, hi deia textualment: **«Pensa que les apreciacions dels/de les estudiants constitueixen un dels mitjans que permetran articular iniciatives adequades d'intervenció per tal de garantir la qualitat de la docència»**. Ens preguntem si el vicerectorat disposava d'altres indicadors, a més de l'enquesta i del compliment dels horaris, per avaluar la qualitat de la docència.

Resumint, l'informe elaborat pel vicerectorat no sols no demostrava que l'enquesta fos vàlida per avaluar l'actuació docent, sinó que tampoc no demostrava que valorés la satisfacció de l'usuari. Una vegada més, s'utilitzava el nom donat a un instrument de mesura com a indicador de la seva validesa. Totes aquestes consideracions, i d'altres de més específiques i massa tècniques per als nostres lectors, foren recollides en un document intitulat «Sobre el document enviat pel vicerectorat de professorat referent a l'Avaluació Docent feta per l'Oficina de Suport a la Docència», i foren enviades al vicerector durant el mes d'abril del 1996 (Unitat Docent de Bàsiques de la Facultat de Medicina, UAB, 1996). A hores d'ara, més d'un any després, encara no se n'ha rebut cap resposta.

¿Es pot considerar que l'enquesta recull amb prou fidelitat l'esperit del **Document-marc per a la reflexió sobre la instauració d'un sistema d'avaluació i estímul de la qualitat de la docència**? Creiem rotundament que no. Suposant que l'objectiu de les enquestes hagués estat l'avaluació de la satisfacció de l'usuari, no tenim evidència de la seva validesa, i es difícil veure quina podria ser la seva traducció en termes de qualitat de la docència. Segons el conjunt de documents consultats, els objectius de les enquestes són ambigus, no hi ha evidència empírica de què és allò que mesuren, i no se sap quina utilització se'n podrà fer en el futur. Malgrat això, les enquestes, construïdes a partir de criteris polític-administratius que poc tenen a veure amb la ciència, es continuen utilitzant com un procediment «científicament correcte» per a l'avaluació del professorat. Paradoxalment, amb data 7-11-96, l'ICE de la UAB ens manifestà que en el moment actual no sembla oportú obrir un debat sobre l'avaluació del professorat universitari.

3.3. L'avaluació dels metges d'hospital

L'avaluació del metge d'hospital té una gran transcendència ja que sovint comporta l'acceptació o no com a treballador sanitari. També és un tema difícil i complex i nosaltres no tenim cap fórmula màgica perquè l'avaluació sigui perfecta. S'han d'avaluar no solament les habilitats i els coneixements mèdics sinó també, d'alguna manera, la seva personalitat i les seves capacitats de relació amb altres col·lectius de treball, i les de comunicació i de tracte amb els malalts i els familiars. En definitiva se l'haurà de jutjar, en certa manera, com a professional i com a persona, la qual cosa, naturalment, és plena de dificultats.

Aquí no podem entrar en les particularitats de l'avaluació ni segons la categoria del metge, des d'un cap de departament a un adjunt, ni segons l'especialitat, malgrat que el nivell de les diferents qualitats exigides tenen una importància diferent segons la categoria i l'especialitat a avaluar. També s'ha de tenir en compte el tipus i les particularitats de l'hospital.

Si bé tots els metges hospitalaris han de tenir aptituds desenvolupades per a l'assistència, la docència i la recerca i, probablement, tots o quasi tots els metges d'hospital les han d'exercir en aquest ordre, no és aplicable el mateix barem en tots els tipus d'hospital. L'únic que podem pretendre fer aquí és remarcar globalment les qualitats que, en opinió nostra, ha de tenir el metge d'hospital, deixant per a altres aportacions una concreció més detallada segons les particularitats esmentades.

Abans de tot, emperò, voldríem manifestar la nostra discrepància de com, no infreqüentment en alguns dels nostres hospitals, es porta a terme l'avaluació dels metges hospitalaris. No és rar que el procés vagi de la manera següent: en primer lloc, és publicat el perfil d'una plaça determinada de metge hospitalari amb les condicions essencials laborals i econòmiques; i en segon lloc, les sol·licituds són examinades per un o diversos comitès de l'hospital però finalment **qui realment avalua i escull el metge és el cap de servei corresponent o, quan es tracta d'avaluar metges per a alts càrrecs, són el gerent i el director mèdic i altres càrrecs polítics de l'hospital qui ho fan**. Creiem que en tot aquest procés sovint manca: a) una major publicitat de la convocatòria; b) una explicació més aprofundida de l'hospital i de la plaça convocada; c) certes condicions, per ara no exigides en general, però sí desitjables, com la dedicació exclusiva, per exemple; d) les incompatibilitats amb altres càrrecs, ocupacions o horaris; i e) hi manca sobretot professionalitat i democratització per designar els membres del tribunal avaluador definitiu, que sovint queda en mans, una vegada més, dels qui detenen el poder.

Així hem arribat fins i tot a la paradoxa que en el nostre país no sigui estrany trobar professionals que una part del dia treballen a l'hospital, sovint amb mitjans tècnics costossíssims, i que dediquen la resta de la jornada a fer-li la competència des d'un altre centre. Seria difícil defensar, en les actuals circumstàncies, que el metge d'hospital no tingui dret a l'exercici privat de la medicina, sobretot d'algunes especialitats, però s'hauria d'evitar el problema actual que comporta el que alguns metges hospitalaris tinguin el seu principal interès professional fora de l'hospital. És veritat que la nostra professió de metge, i sobretot als hospitals, té actualment poc prestigi i està mal pagada en relació a d'altres professions del mateix hospital (polítics, gerents, economistes, etc.), però això no ha de permetre perpetuar una vegada més el cercle viciós; s'han de cercar altres solucions, com han fet altres països propers al nostre. També és veritat que algunes places d'hospital estan interferides per convenis, reglaments i vincles amb altres institucions. Cal expurgar el més aviat possible els càrrecs de l'hospital de totes les càrregues obsoletes o inútils,

però les limitacions que actualment estan vigents no es poden aplicar o deixar de fer-ho segons les conveniències dels avaluadors.

Alguns dels defectes esmentats es podrien solucionar si el tribunal avaluador estigués format, amb l'ajut dels tècnics necessaris, per un bon nombre de professionals de l'especialitat de l'hospital corresponent i de procedents d'altres centres, per tal d'evitar una endogàmia injusta. Aquest tribunal que ha de disposar de recursos, i sobretot de temps suficient, ha d'avaluar en primer lloc els coneixements mèdics corresponents: la consideració de la trajectòria professional de l'aspirant és un bon mitjà per a una avaluació correcta; les referències o informes de tercers s'han d'usar amb la precaució deguda però sovint són imprescindibles si volem conèixer alguns aspectes com, per exemple, l'entusiasme i la capacitat per a la docència. S'ha d'evitar la tendència actual de jutjar quasi exclusivament els mèrits del metge hospitalari pels articles i llibres publicats o pel seu impacte, encara que sigui l'aspecte més fàcilment quantificable de l'avaluació.

Sovint les capacitats assistencials i docents no tenen relació directa amb les publicacions i no totes aquestes tenen el mateix valor. Per descomptat que cal avaluar el nombre de publicacions, però sobretot cal avaluar la vàlua personal i la seva adequació amb el tipus d'hospital i el lloc sol·licitat. Altres aspectes no estrictament mèdics que creiem que cal tenir en compte, perquè faciliten la labor mèdica, són el nivell de coneixement tant d'informàtica com d'idiomes estrangers i el grau de facilitat de contacte amb centres mèdics de prestigi. Sempre s'ha de tenir en compte si el sol·licitant ha fet la residència corresponent, ja que actualment la considerem un factor si no imprescindible, sí molt important. Si l'aspirant l'ha acabada recentment, atès que no existeix avaluació de l'especialitat en finalitzar el període de MIR no hi ha d'haver inconvenient en fer-la en el moment de l'admissió a un hospital.

Poques vegades és necessari avaluar un professional de la medicina en aspectes purament cognitius: aptituds intel·lectives, memòria, velocitat perceptiva, etc. Si calgués fer-ho, experts en aquesta matèria poden ajudar en l'avaluació, com ho fan sovint en la d'altres professionals de tots els nivells. Hem de remarcar, emperò, que en els metges d'hospital cal sempre avaluar, i en certa manera jutjar, la seva personalitat. No són suficients els coneixements mèdics per ésser un bon metge d'hospital, i per això també és necessari mesurar tot un seguit d'altres capacitats: per treballar en equip; per explicar, complaure, respectar i compadir els pacients; per relacionar-se amb els companys i amb d'altres persones que treballaran amb ell; per manar, per obeir; cal avaluar si és professionalment entusiasta o ensopit, tranquil o ansiós, si és imaginatiu, si és racista o classista, si és egocèntric o generós, si és constant i equilibrat, etc.

Totes aquestes característiques influiran sobre el seu rendiment i sobre el dels seus companys i el del personal que amb ell treballaran. Per molt difícil que pugui semblar avaluar totes aquestes qualitats humanes, i més encara

predir com evolucionaran, és fonamental tenir-les en compte. Això es pot aconseguir fent una o diverses entrevistes, o avaluant tècnicament la personalitat mitjançant proves psicològiques, o simplement fent el contracte definitiu després d'un període de prova. Tots aquests procediments són instruments que poden ajudar que la gran majoria de metges d'hospital admesos siguin mèdicament formats i humanament madurs i que el seu objectiu professional coincideixi amb la labor assistencial, docent i de recerca projectades pel seu centre.

3.4. L'avaluació de la pràctica mèdica

A les darreries del segle xx ningú no dubta que la Medicina, amb majúscula i sense adjectius, és una ciència i, com a tal, basada en l'observació de fenòmens, naturals o artificials, per tal de descobrir mitjançant l'experimentació les lleis que els regulen. Fins aquí, la definició pot servir per a qualsevol branca de coneixements sistematitzats, però en tractar-se d'una ciència amb un vessant aplicat, el seu cos de coneixements inclou també el coneixement dels mitjans de prevenir i guarir les malalties així com també la tècnica d'aplicar-los.

Tal i com ja hem remarcat molt més amunt, a les pressuposicions científiques, a partir de l'objectivitat, en ciència, hom elabora teories o explicacions, teories que són refusades o acceptades per la comunitat de científics, no per les característiques personals o de poder de qui les propugna, sinó per les proves empíriques i objectives que avalen cada teoria. I no està de més recordar que l'objectivitat científica té la seva màxima expressió en la mesura científica. Tanmateix, el poder, polític o social, interfereix activament molt sovint les mesures o avaluacions científiques, constituint allò que hem anomenat pseudoavaluació. N'hem exposat tres casos ben clars: l'avaluació dels estudiants, la dels professors i la dels metges d'hospital.

Per tot l'anterior, volem cloure la nostra aportació a la politització de l'avaluació amb l'efecte contrari, amb la **interferència del poder per omissió**, és a dir, amb l'eixordador silenci de les nostres autoritats pel que fa a les anomenades «medicines alternatives», encara que l'expressió «pràctiques alternatives» reflecteixi millor allò que representen. Dins d'aquest terme s'hi poden trobar diversos tipus de pràctiques de signe molt diferent. Bannerman i cols. (1983) les recolliren en un estudi de l'OMS sobre aspectes de la medicina tradicional. Tot i tractar-se de pràctiques molt variables, el que sí tenen en comú és no estar basades en cap tipus d'hipòtesi contrastada d'acord amb el mètode científic; és a dir, corresponen de ple a allò que Bunge (1969) definí com a pseudociències: «cos de creences i de pràctiques que llurs conreadors desitgen, de manera ingènua o maliciosa, donar com a ciència encara que no comparteix amb aquesta ni el plantejament, ni les tècniques ni el cos de coneixements».

És cert que hi ha un cert sector de la població que acudeix de manera ocasional, o fins i tot regular, a aquells que utilitzen algunes d'aquestes «pràctiques alternatives». Aquest és un fet que, per ell mateix, no és criticable des de la perspectiva que defensem, ja que pertany a l'esfera personal, de la mateixa manera que les idees polítiques o religioses del subjecte. Malgrat l'anterior, la decisió d'acudir-hi pot ésser influïda en una gran part pel coneixement de quina és la consideració que, d'aquest tipus de pràctiques, tenen aquelles persones, entitats o organismes, relacionats amb la supervisió de la pràctica mèdica, o en altres paraules, com són considerades aquestes pràctiques en ésser avaluades pels organismes o institucions competents.

És precisament per la importància que, envers la societat, té la opinió d'aquestes institucions que darrerament hom ha pogut assistir amb perplexitat a una sèrie de fets i declaracions sobre aquestes «pràctiques alternatives» sense que, per altra part, hagin estat explicitades ni la metodologia ni els resultats que s'han derivat del treball que, com és de suposar, s'haurà dut a terme abans d'adoptar decisions com ara la impartició d'alguns cursos de postgrau sobre homeopatia o medicina naturista dins d'alguna de les nostres universitats, o bé la declaració d'algun col·legi professional envers la intenció de donar reconeixement oficial a tres tipus de pràctiques alternatives: l'acupuntura, l'homeopatia i la medicina naturista. Davant d'aquestes decisions, és de suposar que aquestes institucions no tindran cap inconvenient a enviar els seus representants a col·laborar en actes com ara la «Fira internacional Màgic de ciències ocultes i medicines alternatives» que enguany ha celebrat la seva vuitena edició a Barcelona.

Darrera aquest fet s'hi troba, com a rerefons, la qüestió de qui ha d'avaluar aquest tipus de pràctiques i, sobretot, què és el que s'ha d'avaluar.

La societat posa en mans de diverses institucions i entitats el control i la supervisió dels coneixements, de l'habilitat i de la correcta pràctica professional. En el nostre país, la certificació que hom ha adquirit els coneixements i les habilitats suficients per garantir una activitat professional mèdica correspon a les universitats, que abans de donar el títol, reben les notificacions d'aptitud de les facultats en el cas del títol de llicenciat, o de les institucions sanitàries en què hom pot cursar l'especialització, per aquest darrer títol.

La supervisió d'una correcta pràctica professional, correspon actualment als col·legis professionals, encara que cada cop més es parla d'introduir criteris per permetre la recertificació dels coneixements i de les habilitats que permeten en el seu moment l'obtenció del títol, tal i com hom pot comprovar en altres apartats dins d'aquesta mateixa segona ponència. El control de l'activitat professional, que habitualment es fa «a posteriori», és a dir, després d'una determinada actuació, es realitza mitjançant els informes de diverses comissions i la consulta a diversos professionals o, millor encara, a les societats científiques corresponents.

Per tant, és evident que les diverses institucions docents i professionals tenen els seus mecanismes d'avaluació que, pel fet de tractar-se d'avaluacions

sobre temes científics, hom hauria de suposar que es basen en el mètode científic i, per tant basades en el contrast de les hipòtesis mitjançant proves empíriques i objectives. Això no obstant, en aquest cas, hom té la percepció que l'avaluació que algunes institucions haurien d'haver fet pel que fa al contingut, metodologia i resultats de les «pràctiques alternatives» o bé és inexistent o bé és absolutament deficient ja que, tal i com Telleria i cols. (1996) assenyalen, les pseudociències no presenten cap prova de les seves teories, mecanismes o hipòtesis explicatives. És cert que s'han publicat diverses metaanàlisis sobre algunes d'aquestes pràctiques però també es cert que els autors que les realitzaren reconeixien que les proves no són suficients per establir conclusions definitives a causa de la baixa qualitat metodològica dels assaigs.

Aquest és un exemple de com una avaluació inexistent o deficient realitzada de manera inconscient o no, pot modificar la conducta dels subjectes susceptibles d'utilitzar-la com a referència en el moment d'escollir una determinada opció per a la cura de la seva salut. De fet, cal remarcar que **una forma de realitzar una pseudoavaluació és precisament la de no avaluar.**

Les conseqüències d'escollir una opció terapèutica no són pas innòcues ja que, d'una banda, s'han descrit alteracions iatrogèniques per manipulacions o administracions de substàncies com a part d'un procés de «pràctica alternativa», i de l'altra —i això és més important— si un malalt afecte d'un procés patològic greu no rep un tractament científicament contrastat des del començament, o bé l'abandona per anar a cercar una «solució alternativa», quan davant de la previsible falta de resposta, retorni al tractament científic haurà perdut un temps essencial per a la bona marxa del seu procés, pèrdua de temps que pot arribar a ser fatal en el cas de determinades entitats patològiques.

Per evitar aquests greus efectes només cal que les persones o institucions a qui pertany el control i la supervisió de la pràctica mèdica compleixin la seva missió avaluant tot tipus de pràctiques mèdiques —pretenguin ésser científiques o no— i donant-ne publicitat per tal que tothom tingui elements de judici per tal de prendre una decisió basada en la realitat i no en un sentiment de màgia, de misteri o en el parany de «és efectiu, però la ciència actual no ho pot explicar...» Un cop informat, evidentment, tothom és lliure d'escollir l'opció que més li convingui, però cal tenir present que sense tota la informació ningú no pot escollir amb llibertat.

Per tant, des d'ací proposem que, dins del marc del Quinzè Congrés de Metges i Biòlegs de Llengua Catalana, hom faci una declaració¹ sobre l'estat del

¹L'Assemblea General del Quinzè Congrés de Metges i de Biòlegs de Llengua Catalana, recollint aquest suggeriment, aprovà pronunciar-se clarament i de manera pública sobre aquest tema, encarregant la redacció de la declaració a les entitats convocants del Congrés, entitats que debateren la declaració següent que se'ls proposà: «Les autoritats científico-sanitàries adverteixen que les anomenades "pràctiques mèdiques alternatives" no compleixen els estàndards científics de rigor i d'eficàcia, i que el seu ús pot endarrerir l'aplicació de tractaments eficaços, perjudicant així seriosament la recuperació de la salut.»

coneixement de les anomenades «pràctiques alternatives», a la llum del mètode científic, i seguint les directrius de diversos estudis sobre les «pràctiques alternatives», i demanem que prenguem consciència que **l'única alternativa a la medicina no és sinó una medicina millor**.

No voldríem acabar aquesta part, dedicada a l'avaluació de la pràctica mèdica, sense descriure un cas que pot ésser demostratiu d'una pseudoavaluació realitzada sobre un tipus de «pràctica alternativa». La comissió de cultura d'una facultat de medicina d'una universitat catalana va organitzar un acte sobre l'hipnosi, amb assistència d'un autoanomenat «expert». Els organitzadors comptaven amb el suport de certs professors que pensaven que potser sí que hi ha fenòmens que la ciència no pot estudiar. En el decurs de l'acte, «l'expert» va oferir una demostració pràctica de les seves habilitats utilitzant un estudiant voluntari, a qui va sotmetre a un procés d'hipnosi. El resultat va ser força espectacular, ja que el subjecte «hipnotitzat» recollia amb gran entusiasme trossos de paper que l'«expert» llençava a terra fent-li creure que eren bitllets de 5000 pessetes. Posteriorment, el mateix subjecte va ser induït a creure que, dins de l'aula, tothom estava despullat, amb gran enrenou dels alumnes presents, especialment de les alumnes, que eren fitades a fons per l'«hipnotitzat» delerós de visions anatòmiques directes. Un cop acabat l'espectacle, el subjecte, que declarà no recordar res, fou sotmès a una exploració electroencefalogràfica pels professors «oberts» a nous fenòmens, tots ells membres del Departament de Fisiologia de la facultat, sense que es detectés cap canvi en el seu funcionalisme cerebral. És a dir, que hom va intentar realitzar una avaluació del procés a què havia estat sotmès l'estudiant. Aquesta avaluació «a posteriori» es va revelar com a pseudoavaluació, ja que «l'experiment», per anomenar-lo d'alguna manera, estava viciat de principi, com alguns descobriren posteriorment, ja que es demostrà que l'estudiant-voluntari estava conxorxat, des de bell començament, amb l'hipnotitzador, i la realitat era que havia simulat un estat hipnòtic.

D'aquí es pot extreure la conseqüència que en avaluar qualsevol tipus de procés o fenomen, cal realitzar una avaluació completa, de principi a final, i valorant totes les variables. Si no és així, hom pot quedar ben retratat.

CLOENDA

Com hem vist al llarg d'aquest estudi de casos, el poder interfereix clarament l'avaluació del rendiment tant dels estudiants universitaris com la dels seus professors, obstaculitza la selecció dels millors metges per a les feines hospitalàries i, per contra, s'inhibeix quan es tracta de qualificar científicament certes preteses terapèutiques. El remei per a aquests «trastorns» avaluatius és el rebuig a qualsevol idea que justifiqui, falsament, una metodologia per a les ciències dites «humanes» diferent i enfrontada a la de les ciències «naturals» i, en conseqüència, l'assumpció decidida que les mesures psicològiques

o socials són metodològicament **iguals** que qualsevol altre tipus de mesures d'un fenomen físic. El lector interessat trobarà a la bibliografia següent bases més que suficients per a aconseguir aquest objectiu.

BIBLIOGRAFIA

- Banerman, R. H. (1983). *Médecine traditionnelle et couverture des soins de santé*. OMS, Ginebra.
- Bonboir, A. (1972). *La Docimologie*. Presses Universitaires de France, París.
- Bunge, M. (1969). *La investigación científica: su estrategia y su filosofía*. Ariel, Barcelona.
- Bunge, M., Ardila, R. (1988). *Filosofía de la Psicología*. Ariel, Barcelona.
- Carroll, L. (1985). *A través de l'espill*. Quaderns Crema, Barcelona.
- Castell, E., García Sevilla, L. (1988). Las dimensiones básicas de la práctica profesional del médico de Atención Primaria. *Atención Primaria* 5, 12-16.
- García Sevilla, L. (1976). La Psicología: Entre la ciencia y la ideología. *Perspectiva Social* 17, 26-31.
- García Sevilla, L. (1979). «Castellà» i «Espanyol» al diccionari de la Reial Acadèmia Espanyola (i 2). *L'Avenç* 17, 26-31.
- García Sevilla, L., Gomà, M., Martí, A., Muntaner, C., Pérez, J., Tobeña, A., Torrubia, R., Trullàs, R. (1983). Treballs en curs del Departament de Psicologia Mèdica. (1983-1984). Treballs del Departament de Psicologia Mèdica, 3. Universitat Autònoma de Barcelona.
- García Sevilla, L. (1985). Anàlisi de la Psicoanàlisi. Monografies de la Secció de Ciències 3. Institut d'Estudis Catalans, Barcelona.
- García Sevilla, L. (1986). Psicología del proceso de minorización de las lenguas sin estado propio. In Ruiz Olabuenaga, J. I., Ozamiz, J. A. (Dir.) *Hizkuntza Minorizatuen Soziologia*, pp. 343-346. Ttartalo, Donostia.
- García Sevilla, L. (1991). Taller sobre bilingüisme passiu. In Diversos autors (ed.), *II Congrés Internacional de la Llengua Catalana. Àrea 2. Sociologia de la Llengua*, pp. 402-403. Diputació de Girona, Girona.
- García Sevilla, L. (1994). Avaluació de l'avaluació de la docència universitària. 1992-93. (NOTA: l'ICE de la UAB en proporciona una fotocòpia a qui la demani.)
- García Sevilla, L. (1995). Les dimensions ideològiques dels balears. *Sus, Balears!* 3, 35-62.
- García Sevilla, L., Balada, F., Muntaner, C. (1982). Els no-catalanoparlants i l'ensenyament en català a la Universitat. Actes del 1er symposium sobre l'ensenyament del català a no-catalanoparlants.
- García Sevilla, L., Jiménez, S., Muntaner, C., Balada, F. (1984). El xoc cultural dels immigrants a Catalunya: Resultats qualitatius. *Annals de Medicina* 70, 108-111.
- García Sevilla, L., Llopart, J., Muntaner, C., Pérez, J., Torrubia, R. (1984). Examen de los exámenes. I Congreso de Evaluación Psicológica. Madrid, 26-29 de setembre.
- García Sevilla, L., Montaña, E., Muntaner, C., Torralba, L. (1981). La problemàtica de la reinstauració del català a la universitat: Estudi empíric d'una facultat. *Ciència* 11, 68-71.
- García Sevilla, L., Vallès, A. (1986). Un sistema econòmic de cens lingüístic de llengües minoritzades. *Treballs de Sociolingüística Catalana* 6, 135-138.
- Gross, P. R., Levitt, N., Lewis, M. W. (1996). The fligth from science and reason. *Annals of the New York Academy of Sciences* 775.
- Kamp, L. J., Langerak, W. F., Gruijter, D. N. M. (ed.) (1980). *Psychometrics for Educacional Debates*. Wiley, Chichester.

- Kleijnen, J., Knipschild, P., Riet, G. (1991). Clinical trials of homoeopathy. *Brit. Med. J.* **302**, 316–323.
- Kline, P. (1993). *The Handbook of Psychological Testing*. Routledge, Londres i Nova York.
- L'Autònoma. Monogràfic. Desembre de 1995.
- L'Autònoma. Monogràfic. Desembre de 1996.
- Magnusson, D. (1967). *Test Theory*. Addison-Wesley, Londres.
- Muntaner, C., Garcia Sevilla, L., Balada, F. (1991). La llengua catalana com a relació constitutiva d'una societat: proves empíriques. *In* Diversos autors (ed.), *II Congrés Internacional de la Llengua Catalana. Àrea 2. Sociologia de la Llengua* 348–352. Diputació de Girona, Girona.
- Pérez, J., Torrubia, R. (1981). Proves de rendiment acadèmic. Departament de Psicologia Mèdica, Universitat Autònoma de Barcelona.
- Rutman, L. (1984). *Evaluation Research Methods: A Basic Guide*. Sage, Beverly Hills.
- Serrano, P., Garcia Sevilla, L., Pérez, J. L., Pina, M., Ruiz, J. R. (1986). Municipal Police evaluation: Psychometric versus behavioural assessment. *In* Yuille, J. C. (ed.), *Police Section and Training: The Role of Psychology*, pp. 257–265. Martinus Nijhoff Publishers, Dordrecht.
- Sureda, B., Artigues, A., Bibiloni, G., Cuyàs, M., Garau, V., Garcia Sevilla, L., Garcias, F., Obrador, J., Oliver, G., Salvà, P., Serra, B. (1987). Enquesta sociolingüística a la població de Mallorca. Universitat de les Illes Balears i Consell Insular de Mallorca, Ciutat de Mallorca.
- Telleria, C., Sanz, V. J., Sabadell, M. A. (1996). La homeopatia: historia, descripción y análisis crítico. (Informe realitzat per encàrrec de l'Institut d'Estudis de la Salut de la Generalitat de Catalunya.) La Alternativa Racional, Saragossa.
- Unitat Docent de Bàsiques de la Facultat de Medicina. UAB. (1995). Resolució sobre l'avaluació docent feta per l'Oficina de Suport a la Docència.
- Unitat Docent de Bàsiques de la Facultat de Medicina. UAB. (1996). Sobre el document enviat pel vicerectorat de professorat referent a l'Avaluació Docent feta per l'Oficina de Suport a la Docència.
- Verger, G. (1994). L'assistència hospitalària a casa nostra. *Avui*. Barcelona. 15–11–1994.
- Verger, G. (1995). L'assistència hospitalària una altra vegada. *Avui*. Barcelona. 20–1–1995.
- Verger, G. (1995). La calidad y la gestión en los hospitales públicos. *La Vanguardia*. Barcelona. 13-4-1995.
- Vicerectorat de Professorat. UAB. (1996). Resposta a la «Resolució sobre l'avaluació docent feta per l'Oficina de Suport a la Docència, aprovada per l'assemblea de professors de la Unitat Docent de Bàsiques (Bellaterra) de la Facultat de Medicina».