

La construcció de tipologies: procés i tècniques d'anàlisi de dades

Màrius Domínguez i Amorós¹

Pedro López Roldán²

1. Universitat de Barcelona. Departament de Sociologia i Metodologia de les Ciències Socials.
08034 Barcelona. Spain

2. Universitat Autònoma de Barcelona. Departament de Sociologia.
08193 Bellaterra (Barcelona). Spain

Resum

Aquest article recull un mateix procés metodològic en el qual s'emmarquen diverses investigacions presentades en aquest volum de la revista, amb un mateix objectiu: la construcció de tipologies en diferents àmbits temàtics.

L'article especifica el marc general del disseny, descriu el seu procés metodològic i d'anàlisi de dades que pot caracteritzar-se per: (1) La font de dades prové d'una gran enquesta sobre hàbits i condicions de vida de la població, l'Enquesta Metropolitana de Barcelona 1990. (2) El plantejament d'un objecte d'estudi concret dins d'una realitat multivariable. (3) La utilització de tècniques d'anàlisi multivariables, en concret, l'Anàlisi de Correspondències Múltiples i les Tècniques de Classificació Automàtica.

Paraules clau: metodologia, tècniques d'anàlisi de dades, anàlisi de correspondències, anàlisi de classificació, tipologia, Enquesta Metropolitana de Barcelona.

Abstract. The construction of typologies: process and techniques of data analysis

In this article, we show the common methodological process that is shared by several researches that are presented in this journal, which follow an identical objective: the construction of typologies in the sociological investigation.

The investigations may be characterized by: (1) they work with a great survey of habits and conditions of life, the 1990 Barcelona Conurbation survey; (2) the analysis of a particular object of study within a multivariate reality, whose characteristics are correlated, and (3) the use of multivariate data analysis, particularly the Multiple Correspondence and the Cluster Analysis.

Key words: methodology, techniques of data analysis, correspondence analysis, Cluster Analysis, typology, Metropolitan Survey of Barcelona.

Sumari

- | | |
|---|--------------------------------------|
| 1. Introducció | 3. Procediment i tècniques d'anàlisi |
| 2. L'Enquesta de la Regió Metropolitana de Barcelona (EMB'90). Condicions de vida i hàbits de la població | Bibliografia |

1. Introducció

Aquest article el podem anomenar «el pòrtic». Pòrtic a una sèrie d'articles que són aplicacions i exemplificacions de la construcció de tipologies en la investigació sociològica. Tipologies enteses en la seva vessant estructurada i articulada, tal com ha quedat palès en l'article teòric anterior.

S'ha considerat adient realitzar un article conjunt d'introducció a les diferents investigacions que es presenten, ja que:

1. Totes elles parteixen de la informació d'una gran enquesta sobre les condicions de vida i hàbits de la població. En concret, la informació de base de les investigacions és l'Enquesta de la Regió Metropolitana de Barcelona 1990 (EMB'90), la qual ofereix informacions sobre diversos aspectes de les condicions de vida i els hàbits de vida configurant una realitat complexa i multivariant.
2. Les diverses investigacions comparteixen un dels objectius generals: l'estudi concret d'una temàtica, superant l'anàlisi sectorial i bivariant, i tenint en compte la complexitat de la realitat multivariant. A més a més de la recerca d'una tipologia general per a cada objecte d'estudi que permeti emergir les caracteritzacions socials en què s'estructuren. Per estudiar en profunditat els grups socials, les zones territorials, la segmentació del mercat laboral i l'estudi del consum s'han de tenir en compte altres àmbits temàtics, com són la llengua, la família, els ingressos, l'educació, etc., que ajuden a tenir una visió global de les condicions de vida.
Aquests objectes d'estudi ens encaren a una realitat complexa i multivariant que ens duen a resumir la informació, estructurar-la i articular-la mitjançant el procés tipològic.
3. Comparteixen un mateix procés metodològic i de tècniques d'anàlisi de dades. Totes les investigacions que es presenten empen l'anàlisi de correspondències múltiple (ACM) dins de les tècniques d'anàlisi multivariant i les anàlisis de classificació o d'agrupament.

2. L'Enquesta de la Regió Metropolitana de Barcelona (EMB'90). Condicions de vida i hàbits de la població

La investigació sota el nom d'Enquesta Metropolitana sobre condicions de vida i hàbits de població va tenir els seus inicis l'any 1984 a partir d'un encàrrec realitzat per l'antiga Corporació Metropolitana de Barcelona i l'actual Mancomunitat de Municipis de l'Àrea Metropolitana de Barcelona, en la qual es plantejava com a objectiu general immediat disposar de dades sistemàtiques sobre les activitats i les formes de vida de la població de l'Àrea Metropolitana. D'aquesta manera es tractaria de convertir l'Enquesta Metropolitana en un instrument periòdic de recollida de dades de caràcter social que permetés la comparabilitat amb altres de similars realitzades en altres països.

Des de 1984 n'han tingut lloc dues edicions, cronològicament situades en 1985 i 1990. Des d'aquests treballs d'investigació s'ha pretès establir com a objectiu global la realització d'anàlisis sociològiques i demogràfiques sobre els múltiples aspectes i dinàmiques que conformen la vida social dels individus, atenent a com es manifesten aquestes condicions en termes de desigualtat i a com es distribueixen entre les persones i els grups socials. En funció d'aquest plantejament es busca determinar la magnitud de les desigualtats econòmiques i socioculturals en diversos àmbits (treball, consum i habitatge, família, nivells educatius, hàbits culturals i lingüístics, ús del temps, accés a equipaments i serveis col·lectius, participació en associacions i xarxes de relació, ús i percepció del territori), quines característiques tenen, com es viuen col·lectivament i intentant analitzar les tendències a l'homogeneïtzació o fragmentació de la població.

Els objectius, l'estudi i l'elaboració d'aquesta problemàtica va dur a l'elecció del mètode de l'enquesta per a la construcció, la recollida i l'anàlisi de les dades. El qüestionari de l'EMB constitueix un extens i variat conjunt de preguntes sobre múltiples aspectes de la vida social i familiar dels individus. Cada un d'aquests àmbits d'estudi de la realitat social ens proporciona una gran riquesa i una gran complexitat d'informació. Aquesta organització i anàlisi en àmbits específics i la seva publicació en informes generals separats, ens ha proporcionat una visió sectorialment i globalment molt complexa, i en certa mesura independent del conjunt d'àmbits temàtics de les condicions de vida.

En el disseny de la mostra per a l'any 1990 es va establir com a objectiu l'elecció d'una mostra representativa dels individus més grans de divuit anys de la Regió Metropolitana de Barcelona, univers poblacional amb aproximadament 3,2 milions de persones¹.

En base a criteris de tipus estadístic i amb el mètode de mostreig aleatori estratificat, es va contemplar un procés de construcció de la mostra en dues etapes: la construcció dels estrats homogenis a partir d'una anàlisi multivariable de dimensionalització i classificació automàtica a partir de les dades del padró d'habitants de 1986, en termes de secció censal, i l'afixació de la mostra segons els criteris de Neyman a partir de la determinació de la seva grandària. El nombre final d'entrevistes va quedar establert en 5.061, valor que suposa un error mostral relatiu de l'1,9% per un nivell de confiança del 95,5% i referit al total de la mostra.

3. Procediment i tècniques d'anàlisi

En aquest apartat especificarem els moments i les operacions, els mètodes i les tècniques d'anàlisi de dades, coherents i adients que s'han emprat en les inves-

1. Comarques del Barcelonès, el Maresme, el Vallès Oriental, el Vallès Occidental i el Baix Llobregat, extnsió geogràfica que concentra les dues tercers parts de la població de Catalunya, amb un total de 129 municipis.

tigacions que es presenten i que són eines bàsiques per a la construcció de la tipologia articulada i estructural, adequats en relació amb el model metodològic que s'ha detallat en l'article anterior.

Es tracta d'elaborar un disseny d'anàlisi que opera en un caràcter eminentment interactiu: a partir d'una problemàtica teòrica i d'un model més o menys elaborat o explícit, a partir de la informació disponible en l'EMB'90, en l'elaboració de la qual ja s'han operativitzat els conceptes i s'ha recollit la informació, s'han aplicat una sèrie de tècniques d'anàlisi de les dades, pròpies de la perspectiva distributiva, que operen amb les relacions entre les variables. Aquestes ens permetran veure les relacions significatives entre les variables i l'ordenació de les unitats.

El procés d'anàlisi en els treballs que es presenten no ha estat apriorístic, en el sentit d'establir models com a punt de partida com a forma de teoria establerta que s'ha de validar. De fet, la lògica ha estat més empírica, en consonància amb el mètode general de l'enquesta.

Això no vol dir que no es tinguin en ment models possibles², però almenys no s'han explicat de forma analítica. En canvi, els conceptes i la seva operativització, i els models teòrics si que han estat presents en l'elecció de les variables pertinents, en la formalització de les tipologies i en la interpretació dels resultats.

Tres són les fases analítiques fonamentals des del punt de vista metodològic:

- a) primera, després de l'anàlisi descriptiva simple prèvia, es realitza una anàlisi de dimensionalització per estructurar inicialment el concepte tipològic a partir de les variables originals;
- b) segona, i en funció de la precedent, una anàlisi de classificació de les unitats en un conjunt de tipus significatius que estructurin el fenomen estudiat;
- c) tercera, l'obtenció i la identificació de la tipologia i la validació d'aquests resultats.

Amb aquests moments de disseny metodològic, els mètodes que proposem com els més adients són l'anàlisi factorial i les tècniques de classificació automàtica. Seran els instruments principals d'anàlisi, de dimensionalització, de classificació i de validació de les tipologies. Les dues tècniques expliciten i permeten la realització de les operacions de combinació i de reducció pròpies de l'anàlisi tipològica. A continuació tractarem successivament cada una de les tècniques d'anàlisi de les dades mostrant la seva adequació a la construcció tipològica.

2. Provenents tant d'altres estudis com de l'experiència en investigacions prèvies.

3.1 Anàlisi de dimensionalització: *anàlisi de correspondències múltiples (ACM)*

L'anàlisi de dimensionalització busca determinar la forma reduïda, estructurada i significativa de l'espai d'atributs inicial sobre el qual es basa i es construeix la tipologia. L'objectiu consisteix a fer emergir, a partir de l'estructura d'interrelació existent entre les variables observades, les variables i les dimensions fonamentals, mitjançant la combinació i la reducció de les primeres. A través de diversos indicadors observables, manifestos, establim el camp més delimitat de significació per al concepte tipològic no necessàriament observable, més aviat latent i no mesurable directament.

Com és farà palès més endavant, la complexitat rau en el caràcter multidimensional de l'objecte d'estudi, fer-la visible i interpretar-la en termes d'estructura. Procedim així a l'estructuració multidimensional del concepte tipològic a través de la transformació de l'espai d'atributs original en un espai d'atributs estructurat. Aquestes consideracions ens justifiquen l'elecció de tècniques d'anàlisi factorial. I, més concretament, l'ACM, ja que ens facilita l'anàlisi d'interdependència i dimensionalització que estructurin el conjunt definit per les variables, i ens permet tractar variables qualitatives.

En aquesta etapa existeixen tres operacions fonamentals: selecció, combinació i reducció, totes les quals són referides a les propietats de les unitats i enteses en un procés analític on s'interrelacionen teoria i realitat. L'operació de selecció és pròpia del nivell teòric de construcció de l'objecte, mentre que les operacions de combinació i reducció de l'espai de les variables es justifiquen a través d'un procés d'anàlisi que tècnicament materialitza les relacions entre les variables, encara que siguin guiades pel model teòric explícitament elaborat.

L'ACM és una anàlisi factorial descriptiva, alhora que amb una vessant estructural, que té l'objectiu de fer paleses les associacions més importants existents entre les variables qualitatives. I la necessitat d'aprofundir en les relacions de dependència que s'estableixen entre dues variables categòriques observades en una mateixa població. El seu avantatge radica a aplicar-la com una tècnica de representació de l'estructura geomètrica simultània del multiespai de relació entre individus i variables, la qual cosa respon a una relació biunívoca entre dos conjunts de dades representats per les línies i les columnes de múltiples taules de contingència que creua variables qualitatives tractades com a nominals. Hem d'imaginar-nos que el punt de partida de l'anàlisi de correspondències múltiples són les diverses taules de contingència de les variables entre si.

En aquest sentit, i des d'un punt de vista metodològic, l'anàlisi de correspondències múltiples ofereix una formalització geomètrica i algebraica que permet el pas de l'aspecte quantitatiu a l'aspecte qualitatiu, de variables mesu-

3. Una altra manera de presentar aquestes dades és a través de la construcció d'una taula disjuntiva completa. En aquesta taula les línies representen els individus i les columnes representen les modalitats de les variables: a la intersecció de la línia de l'individu i de la columna k ens trobarem X_{ik} , que anirà de 1 a 0 segons si l'individu i posseeix la modalitat k o no.

rades o tractades a nivell nominal a un tractament quantitatiu que conserva la natura i les operacions pròpies de les mesures qualitatives, permet el pas de l'aspecte heterogeni a una construcció d'ordre estructural basat en la comparació de perfils, en les posicions relatives i en l'ordenació espacial (tipològica) de formes, més que no pas en els efectes de magnitud (J.P. Benzécri, 1973; F. Conde, 1987; J.M. Cornejo, 1988; J. Ibáñez, 1985)³. Per aquestes raons, l'ACM ofereix la possibilitat d'analitzar estructures de relació, on la lectura tipològica, els perfils relatius i les relacions d'ordre són privilegiades enfront del caràcter més quantitatiu i de magnitud del nombre i la mètrica contínua. Es configura així un context articulat i estructural.

Com a anàlisi factorial extreu del cúmul de dades l'estructura del multiespai a través de la projecció reduïda als eixos de màxima inèrcia, intentant resumir la massa de les dades amb la menor pèrdua d'informació possible a unes poques dimensions. I com a mètode ens garanteix la propietat desitjada per a la construcció tipològica d'organitzar la realitat multidimensional d'un fet social a partir de conceptes, encara que no necessàriament estructurats.

L'ACM no és una mera tècnica d'extracció de factors de qualsevol taula de números positius, sinó un conjunt metodològic ben organitzat que inclou:

- regles per a la preparació de les dades sotmeses a l'anàlisi, en particular la codificació disjuntiva completa;
- criteris de validesa dels resultats;
- aportació de procediments eficaços de discriminació i de regressió;
- conjugació harmoniosa amb la classificació automàtica.

Per aquest tipus de transformacions es realitza un salt important, i és el de considerar cada categoria de les variables pertinents com a modalitats de resposta.

Sota la transformació d'aquesta informació en una taula de Burt, la matriu de dades no s'ha de veure ja com a individus per variables, sinó que és el creuament entre categories d'una variable, per exemple: la posició en el mercat laboral, per les categories d'una segona, l'origen geogràfic familiar. El principi de la representació dual en un mateix espai d'individus i variables, constitueix la seva característica geomètrica més destacable.

Es sota aquestes consideracions que la tècnica d'ACM aconsegueix dos dels objectius principals: l'estudi de les variables i l'estudi dels individus.

D'altra banda, cal assenyalar com un avantatge pràctic la fórmula inicial de la distància distribuïdal⁴ que permet donar una estructura matemàtica a una taula de números positius, evitant en certa mesura l'arbitrarietat de tot particionament i codificació de la realitat.

4. Recerca de la representació dual òptima en què les variables i els individus juguen un paper simètric. Aquesta simetria s'aconsegueix a través de la mètrica de *CHI-2*, establint relacions de dependència simètrica entre files i columnes.

L'equivalència distribucional garanteix una certa invariància dels resultats en relació amb el nombre de categories en què se subdivideixen les variables, ja que agrupant categories amb perfils semblants no es perdria informació. Alhora que tampoc es guanyaria informació subdividint en més categories de resposta la variable objecte d'estudi si aquestes subdivisions no suposessin perfils diferenciats. Així doncs, les divisions en categories té sentit únicament si suposen perfils de repartició diferents en l'altra variable.

Un altre dels avantatges que ens permet l'ACM és el de poder atorgar a les variables estatus diferents segons la seva importància estructuradora: variables actives, que són les que han entrat en l'anàlisi, i variables il·lustratives, que ens poden permetre enriquir la interpretació en termes d'estructura⁵.

D'aquesta forma es garanteix una primera estructuració del fenomen estudiat i es disposa del criteri bàsic del procés de classificació posterior que aboca en la constitució dels tipus i la tipologia.

3.2 Anàlisi de classificació

El segon gran moment del procés de construcció i anàlisi de la tipologia consisteix a classificar, a partir de l'espai d'atributs reduït obtingut en l'anàlisi de dimensionalització, les unitats considerades per obtenir un conjunt de tipus que siguin l'expressió de grups d'unitats homogenis internament i significativament diferenciats entre si. L'objectiu final és el d'obtenir un nombre reduït de tipus que són les categories de la variable tipològica latent que estructuri i defineixi l'objecte d'estudi. Igual que en l'anàlisi de dimensionalització, el de classificació té l'objectiu d'estructurar. En aquest cas, estructurar les unitats a partir de l'entramat conceptual en què s'inscriuen.

S'estableixen mesures de similitud i de dissimilitud amb les quals es podran avaluar proximitats i així constituir grups o classes d'equivalència que ordenin el conjunt d'unitats.

Per l'operació de classificació obtindrem, per l'agrupació dels individus segons l'espai de noves variables dimensionalitzades (les coordenades factorials), una nova variable que identifiqui el concepte tipològic, que estructuri l'objecte d'estudi. Aquesta operació és complementària a la de dimensionalització, en el sentit que en l'anterior es prenen les variables com a operacionables. Ara es prenen com a camp de tractament els individus, i les variables són els instruments de mesura.

Des del punt de vista de procediment, l'objectiu de les tècniques de classificació és l'obtenció de classes d'unitats al més homogènies possible en l'interior de la classe i al més heterogènies possible entre elles, d'acord amb un conjunt de variables criteri, que en el nostre cas són les variables seleccionades com a pertinents per a l'objecte tipològic i que han estat transformades en unes noves i poques dimensions mitjançant l'anàlisi de dimensionalització.

5. La possibilitat de projectar variables suplementàries en els plans de projecció, encara que aquestes no hagin contribuït a la determinació dels eixos factorials.

Com a procés general de l'anàlisi de classificació es poden enumerar les etapes següents: selecció de les variables, elecció de la mesura de semblança i proximitat, construcció de la matriu de distàncies, elecció del mètode de classificació, assignació dels individus en classes que han de ser fixades i validació dels resultats.

En aquest procés es van intercalant decisions de tipus estadístic amb decisions de consistència, estabilitat, interpretació i justificació sociològica. Les decisions tant de tipus estadístic com sociològic poden conduir a diversos resultats o conclusions que poden ser acceptables. Per aquesta raó, és molt important l'explicitació de criteris coherents amb els objectius fixats, com també la validació successiva i global dels resultats. Com a tècnica estadística amb tanta variabilitat té, doncs, un cert grau d'indeterminació en l'obtenció de les classificacions finals i en la delimitació de les unitats constitutives de cada classe. Per això, en darrer terme, juntament amb els criteris tècnics i estadístics que ens ajuden a determinar els grups que emergeixen, són els criteris teòrics i de coherència els que orientaran la significació de l'anàlisi i de les classificacions.

La tendència a l'empirisme en l'aplicació d'aquesta tècnica pot conduir a conclusions precipitades, incoherents o rígides on no es respecti la natura més aviat exploratòria i heurística de l'anàlisi de classificació. Per aquests motius és indicada la realització d'altres tipus d'anàlisis prèvies que ens permetin la millor elecció de les variables emprades, com el de la dimensionalització.

Així doncs, en les investigacions que es presenten, juntament amb l'ACM s'ha emprat la classificació dels subjectes sobre la base de les coordenades factorialment a través d'un índex de distància. L'objectiu final que es persegueix és el d'agrupar individus similars⁶ segons la realitat multivariant objecte d'estudi en classes representatives. D'aquesta forma, a posteriori, podem operar a partir de la noció d'individu artificial mitjà de cada classe. Perdem així tota la riquesa original de les diferències individuals, però guanyem en capacitat de síntesi i, per tant, de comprensió.

En les investigacions que es detallen s'ha emprat la tècnica de classificació mixta, per partició directa i jeràrquica, especialment adaptada al tractament de grans quantitats de dades. Aquesta tècnica de classificació ha significat, d'una banda, fer una anàlisi detallada del dendrograma, de l'arbre de partició, per tal de veure com va establint successivament els nusos de partició en classes. D'altra banda, aspectes tals com la grandària dels grups de la partició, la inèrcia que ens estableix l'homogeneïtat o la dispersió dels grups i la descripció de les classes per les modalitats significatives han estat bàsics per fer l'elecció de les tipologies.

6. Els individus similars seran aquells que més concordança i semblances tinguin en les respostes a les modalitats.

Bibliografia

- ALVIRA MARTÍN, F. (1989). «Diseños de investigación social: criterios operativos». En García Ferrando; M., J. Ibáñez y F. Alvira. *El análisis de la realidad social*. Madrid: Alianza.
- BENZÉCRI, J.P. (1973). *L'Analyse des données*. París: Dunod.
- BISQUERRA ALZINA, R. (1989). *Introducción conceptual al análisis multivariable. Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD*. Barcelona: PPU.
- CICOUREL, A. (1982). *El método y la medida en sociología*. Madrid: Editora Nacional.
- CONDE, F. (1987). «Un ensayo de articulación de las perspectivas cuantitativa y cuasi-investigación social». *Revista Española de Investigaciones Sociológicas* 51, julio-septiembre, p. 91-117.
- CORNEJO ALVÁREZ, J.M. (1988). *Técnicas de investigación social: el análisis de correspondencias (teoría y práctica)*. Barcelona: PPU.
- FLAQUER, Ll.; MASSATS, M.; MENDIZÁBAL, E.; PUJADAS, I. (1992). *Enquesta Metropolitana de Barcelona 1990. Vol. 1. Aspectes demogràfics i característiques familiars i relacionals*. Institut d'Estudis Metropolitans de Barcelona.
- GRANDE ESTEBAN, I.; ABASCAL FERNÁNDEZ, E. (1989). *Métodos multivariantes para la investigación comercial*. Barcelona: Ariel.
- GREENACRE, M. (1984). *Theory and application of correspondence analysis*. Londres: Academic Press.
- IBÁÑEZ, J. (1985). «Las medidas de la sociedad». *Revista Española de Investigaciones Sociológicas* 29, p. 85-127.
- LEBART, L.; MORINEAU, A.; LAMBERT, T. (1987). *SPAD.N. Système portable pour l'analyse des données*. Sèvres: Cisia.
- LEFEBVRE, J. (1983). *Introduction aux analyses statistiques multidimensionnelles*. París: Masson.
- LÓPEZ, P. (1994). *La construcción de tipologías en sociología: propuesta metodológica de construcción, análisis y validación*. Tesis doctoral. Bellaterra: Universitat Autònoma de Barcelona.
- LOZARES, C.; LÓPEZ, P. (1991). «El análisis multivariado: definición, criterios y clasificación». *Papers, Revista de Sociologia* 37.
- LOZARES, C.; DOMÍNGUEZ, M. (1993). *Territori i realitat social. Les zones sociodemogràfiques de la Regió Metropolitana de Barcelona. Enquesta de la Regió Metropolitana de Barcelona*. Barcelona: Institut d'Estudis Metropolitans de Barcelona.
- MIGUÉLEZ, F.; TORNOS, T. (1992). *Enquesta de la Regió Metropolitana de Barcelona 1990. Treball, condicions econòmiques i formes de consum*. Barcelona: Institut d'Estudis Metropolitans de Barcelona.
- RIERA, P.; PALLARÉS, M.; FLORES, J.L. (1992). *Enquesta de la Regió Metropolitana de Barcelona 1990. Característiques de l'habitatge, la mobilitat i la percepció del territori*. Barcelona: Institut d'Estudis Metropolitans de Barcelona.
- SÁNCHEZ CARRIÓN, J.J. (1984). *Introducción a las técnicas de análisis multivariable aplicadas a las ciencias sociales*. Madrid: CIS.
- SÁNCHEZ, C. (1994). *La definició dels grups socials a la Regió Metropolitana de Barcelona. Un problema metodològic i teòric*. Tesis doctoral. Bellaterra: Universitat Autònoma de Barcelona.

- SUBIRATS, M.; SÁNCHEZ, C.; DOMÍNGUEZ, M. (1992). *Grups i classes socials a la Regió Metropolitana de Barcelona. Enquesta Metropolitana de Barcelona 1990*. Institut d'Estudis Metropolitans de Barcelona.
- SUBIRATS, M.; MASSATS, M.; SÁNCHEZ, C. (1992). *Enquesta de la Regió Metropolitana de Barcelona 1990. Educació, llengua i hàbits culturals*. Barcelona: Institut d'Estudis Metropolitans de Barcelona.