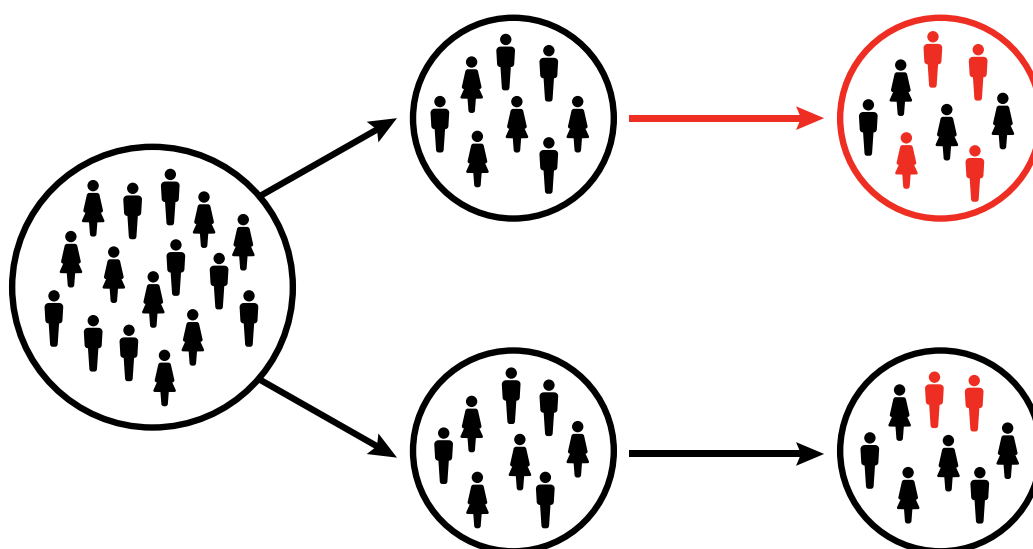


Guia pràctica 15

Guia pràctica d'avaluació de programes del tercer sector social

Col·lecció Ivàlua de guies pràctiques sobre avaluació de polítiques públiques



ivàlua  Institut Català d'Avaluació de Polítiques Públiques

 Taula d'entitats del Tercer Sector Social de Catalunya

Institucions membres d'Ivàlua:

 Generalitat de Catalunya
Departament de la Vicepresidència i d'Economia i Hisenda

 Diputació Barcelona

 upf. Universitat Pompeu Fabra Barcelona

 CTESC
Consell de Tercer Sector Social de Catalunya

 Generalitat de Catalunya
Consell Interuniversitari de Catalunya

©2017, Ivàlua

No es permet la reproducció total o parcial d'aquest document, ni el seu tractament informàtic, ni la seva transmissió en qualsevol forma o per qualsevol mitjà, ja sigui electrònic, mecànic, per fotocòpia, per registre o altres mètodes, sense el permís del titular del Copyright.

Equip de Treball:

Miquel Àngel Alegre

Marçal Farré

Ramon Sabes-Figuera

Federico Todeschini

Disseny: petitcomite.net

Impressió: Cachiman

Primera edició: abril 2017

Dipòsit legal: B. 16329-17

ÍNDEX

PRÒLEG

PÀG. 7

L'avaluació de programes socials.

Una aposta per un Tercer Sector Social més innovador.

pàg. 7

El perquè d'aquesta guia

pàg. 9

1. INTRODUCCIÓ

PÀG. 11

1.1. Presentació

pàg. 11

1.2. Preguntes, avaluacions i components dels programes

pàg. 12

1.3. Una guia per al tercer sector social

pàg. 15

2. AVALUACIÓ DE LES NECESSITATS

PÀG. 17

2.1. Punt de partida i motivació

pàg. 17

2.2. Què vol dir avaluar les necessitats?

pàg. 17

2.3. Preguntes i fases de l'avaluació de les necessitats

pàg. 19

FASE 1: DEFINICIÓ I DIMENSIÓ DE LA PROBLEMÀTICA

PÀG. 20

FASE 2: ANÀLISI DE LES NECESSITATS DE SERVEI

PÀG. 21

FASE 3: ANÀLISI DE LES NECESSITATS I UTILITZACIÓ DE SERVEIS

PÀG. 22

2.4. Mètodes per avaluar les necessitats

pàg. 24

2.5. Recursos

pàg. 29

3. AVALUACIÓ DEL DISSENY

PÀG. 31

3.1. Punt de partida i motivació

pàg. 31

3.2. Què vol dir avaluar el disseny d'un programa

pàg. 31

3.3. Preguntes i focus de l'avaluació del disseny

pàg. 32

3.4. Mètodes per avaluar el disseny

pàg. 41

3.5. Recursos

pàg. 46

3.6. Referències

pàg. 47

4. AVALUACIÓ DE LA IMPLEMENTACIÓ

PÀG. 48

4.1. Punt de partida i motivació

pàg. 48

4.2. Què vol dir avaluar la implementació d'un programa

pàg. 48

4.3. Preguntes i dimensions de l'avaluació de la implementació

pàg. 49

4.4. Mètodes per avaluar la implementació d'un programa

pàg. 51

4.5. Recursos

pàg. 58

4.6. Referències

pàg. 58

5. AVALUACIÓ DE L'IMPACTE	PÀG. 60
5.1. Punt de partida i motivació	pàg. 60
5.2. Què vol dir avaluar l'impacte d'una intervenció	pàg. 60
5.3. Mètodes per avaluar l'impacte	pàg. 62
5.4. Recursos	pàg. 73
5.5. Referències	pàg. 74
6. AVALUACIÓ ECONÒMICA O D'EFICIÈNCIA	PÀG. 75
6.1. Punt de partida i motivació	pàg. 75
6.2. Què vol dir avaluar econòmicament un programa	pàg. 75
6.3. Preguntes i elements clau d'una avaluació econòmica	pàg. 77
6.4. Mètodes per a l'avaluació econòmica	pàg. 79
6.5. Quina És la utilitat de les avaluacions econòmiques per a les entitats del tercer sector social?	pàg. 86
6.6. Recursos	pàg. 86
6.7. Referències	pàg. 87
7. CONCLUSIÓ	PÀG. 88
7.1. Recapitulació: preguntes, moments i mètodes d'avaluació	pàg. 88
7.2. El plantejament de l'avaluació: cinc passos bàsics	pàg. 93
PAS 1. ANÀLISI PRELIMINAR DE LA INTERVENCIÓ	PÀG. 93
PAS 2. IDENTIFICACIÓ DEL PROPÒSIT DE L'AVALUACIÓ	PÀG. 94
PAS 3. IDENTIFICACIÓ I PRIORITZACIÓ DE PREGUNTES D'AVALUACIÓ (TIPUS D'AVALUACIÓ)	PÀG. 94
PAS 4. IDENTIFICAR ELS RECURSOS DE L'AVALUACIÓ	PÀG. 96
PAS 5. PLA INICIAL DE L'AVALUACIÓ	PÀG. 96

FIGURES

Figura 1. Els components del programa: exemple “Feina amb Cor” (Caritas)	pàg. 13
Figura 2. Tipus d’avaluacions	pàg. 14
Figura 3. Exemple d’arbre de problemes	pàg. 21
Figura 4. Necessitat, demanda i oferta de serveis	pàg. 23
Figura 5. Teoria del canvi: hipòtesis del programa	pàg. 33
Figura 6. Exemple de model lògic	pàg. 35
Figura 7. Seqüència temporal dels objectius o estructura de l’impacte	pàg. 37
Figura 8. Exemples de decisions a diferents nivells de l’estructura organitzativa	pàg. 51
Figura 9. El disseny amb assignació aleatòria pura: tractats versus controls	pàg. 64
Figura 10. Alternatives per fer una avaluació d’impacte	pàg. 71
Figura 11. Esquema d’una avaluació econòmica	pàg. 76
Figura 12. Tècniques d’avaluació econòmica: exemple intervencions laborals	pàg. 84
Figura 13. Identificació de les preguntes d’avaluació, arbre de decisions	pàg. 95

TAULES

Taula 1. Treball de camp realitzat	pàg. 16
Taula 2. Fases de l’avaluació de necessitats	pàg. 19
Taula 3. Mètodes i tècniques segons la fase d’avaluació de les necessitats	pàg. 25
Taula 4. Procediments habituals de recollida i revisió d’evidències	pàg. 43
Taula 5. Bases de dades d’evidències i impactes, segons àmbits d’interès	pàg. 46
Taula 6. Preguntes rellevants segons els components als que fan referència	pàg. 50
Taula 7. Paràmetres per a l’estandardització de la recollida de dades	pàg. 54
Taula 8. Exemple d’una ACB (dades no reals)	pàg. 83
Taula 9. Trets principals dels diferents tipus d’avaluació	pàg. 90

GRÀFICS

Gràfic 1. Exemple de regressió discontinua.	pàg. 67
Gràfic 2. Exemple de dobles diferències	pàg. 69
Gràfic 3. Cost addicional (£) per millorar problemes de conducta en infants (de 3 a 8 anys) mitjançant programes de “parenting”	pàg. 89

PRÒLEG

L'AVALUACIÓ DE PROGRAMES SOCIALS. UNA APOSTA PER UN TERCER SECTOR SOCIAL MÉS INNOVADOR.

El Tercer Sector Social és el conjunt d'entitats sense afany de lucre que treballen per la inclusió i la cohesió social, amb especial atenció als col·lectius més vulnerables de la societat: infants, joves, persones immigrades, gent gran, discapacitats psíquics, físics i sensorials, drogodependents, persones amb dificultats per accedir a un lloc de treball, persones sense sostre, persones amb problemes de salut mental, entre d'altres. Segons dades del darrer "Baròmetre del Tercer Sector Social", les entitats socials que agrupa la Taula del Tercer Sector atenen a 1.550.000 persones a Catalunya, compten amb el suport i la col·laboració de 560.000 persones sòcies, 564.000 donants i 350.000 voluntàries, i ocupen 85.000 professionals. La seva activitat econòmica, propera als 3.000 milions d'euros, suposa l'1,42% del PIB de Catalunya.

Aquestes xifres ens donen una aproximació de la importància d'un sector que és clau en la provisió de serveis per part de les administracions públiques. El Tercer Sector Social aporta un valor afegit en la gestió dels programes que gestiona. Aquest valor es concreta, entre d'altres, en quatre aspectes que considero fonamentals: un arrelament incomparable en el territori; un compromís ferm per mantenir una ocupació de qualitat; una exigència cap a la transformació social i econòmica dels territoris on intervé; i un deure cap a la comunitat a partir de la implicació de la mateixa ciutadania que atenem. Malgrat aquests i d'altres aspectes que descriuen la vocació de servei públic amb la que actua el Tercer Sector Social no podem deixar de defensar, en igual mesura, la qualitat en el servei dels programes gestionats per les entitats socials. Aquesta qualitat ha estat i és reconeguda per les administracions amb les que ha cooperat el nostre sector.

Tanmateix, un dels principals reptes del Tercer Sector és, precisament, el fet de poder objectivar aquesta qualitat dels programes que gestiona. Aquest és un dels reptes als quals vol donar resposta el III Pla de Suport al Tercer Sector Social, que el 25 de juliol de 2016 vam signar amb el Govern de Catalunya, el qual en la seva mesura 3a estableix l'objectiu d' "aprofundir en l'avaluació de la gestió dels projectes i programes del tercer sector social, per millorar-ne l'efectivitat, l'eficiència i el rendiment de comptes envers la ciutadania". Es tracta, doncs, d'un objectiu estratègic per al Tercer Sector Social, que s'incardina dins d'una estratègia global que hem emprès en els darrers anys per tal d'assolir un major grau d'innovació i eficiència, i que engloba, entre d'altres, el foment de la internacionalització, de la innovació social, de la transparència o de la transformació digital de les entitats i dels seus programes socials. Poder disposar de dades i indicadors que permetin l'avaluació dels programes socials que gestiona el Tercer Sector Social, així com el seu impacte social, ha de permetre seguir fent avenços cap a un sector més innovador, més transparent, més competitiu i, en conseqüència, més fort.

No vull deixar d'agrair a Ivàlua l'haver acceptat el repte de definir aquest marc de referència sobre l'avaluació de programes del Tercer Sector Social. Aquest és un exercici nou i alhora necessari per al nostre sector. I, per descomptat, també vull manifestar el meu agraïment a les més de vint entitats socials que han fet possible l'elaboració d'aquesta guia a través de les seves reflexions sobre la necessitat d'avançar cap a una major política d'avaluació dels seus programes.

Oriol Illa

President de la Taula d'entitats del Tercer Sector Social de Catalunya

EL PERQUÈ D'AQUESTA GUIA

La missió d'Ivàlua és la de promoure l'avaluació de polítiques públiques i de programes del tercer sector social a Catalunya, per millorar-ne l'efectivitat, l'eficiència i la transparència envers la ciutadania; per tant, l'avaluació de programes del tercer sector social és també una missió fundacional d'Ivàlua. No podem obviar, en tot cas, que molts dels programes del tercer sector fan una tasca difícil de distingir de les polítiques públiques, sense entrar en valorar si això és bo o no. Aquests programes apleguen una elevada diversitat d'iniciatives i estan adreçats a gairebé tots els col·lectius que ens puguem imaginar. En tots els casos, es tracta de programes amb voluntat de transformació de la societat, una voluntat que fonamenta la lògica de les activitats que s'hi realitzen.

Avaluar significa també eixamplar la frontera del que se sap en relació a aquests programes i fer-ho amb certa sistemàtica. Només avaluant podrem generar coneixement sobre aquestes intervencions, sobre com s'han dissenyat, sobre com es posen en marxa, sobre quins efectes tenen i, al cap i a la fi, sobre si realment funcionen.

Només sabent tot això podran les entitats responsables dels programes corregir allò que és susceptible de ser corregit. És precisament per això que podem afirmar que aquesta és una guia amb una clara orientació pràctica, perquè sobretot té com a voluntat ajudar les entitats del tercer sector social a generar coneixement útil per a la millora de la presa de decisions; en altres paraules, generar coneixement per a millorar.

És important assenyalar que, per garantir aquest component pràctic de la guia, aquesta s'ha dissenyat de la mà de les entitats del tercer sector. S'han dut a terme entrevistes i grups focals amb responsables de diverses entitats, amb l'objectiu de copsar diferents visions i sensibilitats d'un sector que és molt ampli i heterogeni. Aquest treball conjunt ha permès identificar les diferents impressions del sector en relació als obstacles i les oportunitats que les entitats es troben a l'hora d'incorporar l'avaluació dins la seva cultura i pràctica rutinàries.

Així doncs, la guia que teniu a les vostres mans aporta un marc de referència comú, exhaustiu i pràctic per apropar l'avaluació de programes al tercer sector social.

En aquest sentit és important agrair la voluntat i l'empenta de la Taula d'Entitats del Tercer Sector Social de Catalunya que aposta per posar l'avaluació al centre.

Aquesta aposta compartida en posar l'avaluació de programes del tercer sector al centre és de ben segur la més efectiva si el que es vol és contribuir a que els programes, i per tant, també les entitats, puguin dur a terme el seu objectiu de transformar la realitat.

Frederic Udina

President de l'Institut Català d'Avaluació de Polítiques Públiques (Ivàlua)

1. INTRODUCCIÓ

1.1. PRESENTACIÓ

Com en altres països, a casa nostra és ja un fet comú la idea que totes aquelles intervencions (públiques, privades, concertades) que cerquin produir millores en les oportunitats i la qualitat de vida de les persones han de ser avaluades. Que l'avaluació, el coneixement que aporta, ha de contribuir a millorar aquestes intervencions, els seus processos, la seva equitat, efectivitat i eficiència. Que l'avaluació ha de respondre al requeriment de transparència i rendició de comptes davant dels usuaris o destinataris de la intervenció i davant la ciutadania en general. La força d'aquest missatge no és pas nova en el món del tercer sector social.

Ara bé, podem afirmar que aquestes idees s'estan fent efectives a la pràctica? En altres paraules, disposa el tercer sector a Catalunya d'un marc coherent i complet d'avaluació dels seus programes? Produïm prou evidències, sòlides i robustes, sobre el funcionament i els impactes d'aquests programes? Utilitzem el coneixement que proporciona l'avaluació per millorar el disseny i planificació de les línies d'acció de les entitats? Generem les condicions que han de permetre fer de l'avaluació un procés obert i transparent, susceptible d'oferir informació rellevant a les entitats, els seus usuaris i la ciutadania en general?

És innegable que les entitats del tercer sector han anat incorporant diferents accions d'avaluació dins les seves rutines de funcionament. Aquesta incorporació és, a dia d'avui, força desigual, en intensitat i en interessos d'avaluació. Mentre que algunes entitats han fet esforços rellevants en aplicar sistemes de seguiment de processos, qualitat i *benchmarking*, d'altres han intentat elaborar anàlisis força exhaustives de les necessitats dels usuaris o estimacions més o menys reeixides del retorn social de les seves accions. Altres entitats han romàs més al marge d'aquests avenços.

Sigui com sigui, el camí a recórrer per fer de l'avaluació (de l'avaluació més rigorosa) un element intrínsec a l'ADN del tercer sector és encara llarg i es troba carregat de reptes. I val a dir que en aquest camí es troba no només el tercer sector social, sinó també la resta d'àmbits d'acció social (inclosa l'Administració pública) a casa nostra, com en tants altres països del nostre entorn.

La guia que aquí presentem no espera convertir-se en cap remei als dèficits d'avaluació del tercer sector a Catalunya. Tanmateix, sí aspira a proveir un marc conceptual i metodològic comú i bàsic que contribueixi a motivar i orientar la pràctica avaluativa de les entitats, una aportació que comença per fer explícites quines són les preguntes clau que els diferents tipus d'avaluacions poden ajudar a respondre: preguntes sobre la problemàtica que justifica la intervenció, sobre el disseny del programa, sobre la seva implementació i sobre el seu

impacte social i econòmic. Perquè avaluar passa, en primer lloc, per conèixer què volem saber dels programes, això és, per saber quines preguntes volem respondre. La guia s'ocupa d'identificar aquests interrogants i d'aportar orientacions i eines metodològiques per donar-hi resposta.

1.2. PREGUNTES, AVALUACIONS I COMPONENTS DELS PROGRAMES

L'estructura de continguts d'aquesta guia progressa en el mateix sentit com ho hauria de fer el funcionament dels programes, des d'allò que justifica la seva raó de ser (el "problema") i sobre la base dels que podrien considerar-se els seus components principals. Val la pena, doncs, començar identificant aquest conjunt d'elements, la consideració dels quals s'emmarca en la concepció de les polítiques i els programes en tant que "transformació intencionada d'uns recursos (inputs) en activitats (processos) a fi de produir uns impactes desitjats (outcomes), en el marc d'un context específic" (McLaughlin i Jordan, 2004).

- **Problema (el que motiva el programa).** Remet a la raó de ser del programa, el que justifica el seu disseny i implementació. De forma general, es fa referència a l'existència d'un "problema" o d'una "situació social insatisfactòria" que s'espera pugui ser corregida mitjançant la intervenció en qüestió. Podem entendre que aquesta situació pot ser insatisfactòria ja en el present (problema manifest) o que pot acabar-ho sent en el futur si no s'hi actua en el present (problema latent).
Són nombroses les problemàtiques que motiven l'acció del tercer sector social, entre d'altres: la pobresa infantil, el fracàs escolar, la violència domèstica, les dificultats d'inserció laboral dels col·lectius amb trastorns mentals, l'aïllament i exclusió de la gent gran, la pobresa energètica, etc.
- **Recursos o inputs (el que el programa necessita).** Fa referència al conjunt de recursos (humans, monetaris, materials, cognitius, etc.) que es preveu mobilitzar per afrontar la problemàtica en qüestió. Per exemple, en el cas d'un programa d'inserció laboral per a col·lectius vulnerables, es podrien requerir treballadors socials, tutors, insertors i mediadors laborals, i/o personal voluntari; unes aules, uns materials formatius i uns centres de treball col·laboradors; un finançament públic i/o filantròpic per cobrir el cost de les activitats.
- **Activitats (el que el programa fa).** Recull el conjunt de línies d'acció de la intervenció, per tant, els seus serveis i continguts concrets. Representa la traducció programàtica dels recursos que es preveu mobilitzar. Per exemple, en el cas d'un programa terapèutic per a malalties neurològiques, probablement s'inclourien activitats com tallers d'estimulació cognitiva, seguiments i visites individualitzades, sessions de teràpia ocupacional, musicoteràpia, dinàmiques amb grups de familiars, etc.

- **Productes o outputs (el que el programa assoleix)**¹. Són els assoliments o realitzacions del programa, allò que acaba fent i abastant per mitjà de les seves activitats. A diferència dels impactes de la intervenció (que fa referència als canvis en la realitat del problema atribuïbles a l'acció del programa), la idea de productes o outputs remet al conjunt de béns i serveis que el programa aconsegueix proveir. Per exemple: nombre i perfil d'usuaris atesos o persones formades, cobertura sobre població diana, seguiments individuals realitzats, cursos programats, etc.
- **Impactes (el que el programa transforma)**. Són els canvis en les característiques i la magnitud del problema que són conseqüència de les activitats i productes de la intervenció. Els impactes, per tant, fan referència al grau d'assoliment dels objectius estratègics del programa, mesurat en termes de canvi en els **outcomes** (indicadors de resultat) d'interès. Si pensem, per exemple, en una intervenció centrada en la reinserció social de joves extutelats, podríem fixar-nos en outcomes personals (millora de l'autonomia i la resiliència), educatius (retorn als sistema educatiu formal i no formal) i laborals (millora de l'ocupabilitat), entre d'altres.

Figura 1. Els components del programa: exemple “Feina amb Cor” (Caritas)

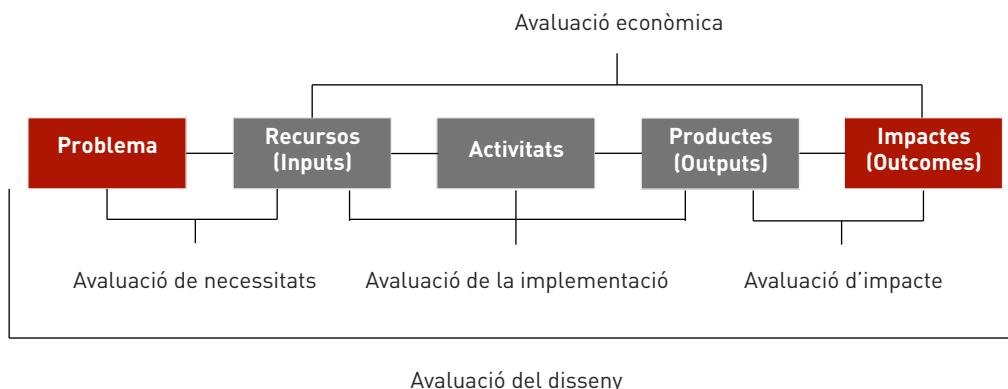


Font: Elaboració pròpia a partir de Blasco (2009: 7).

¹ La traducció dels conceptes *output* i *outcome* al català pot portar a equívocs. Ambdós termes es tradueixen de vegades per resultat, per la qual cosa hem preferit emprar els termes *productes* i *impactes* per referir-nos-hi. En qualsevol cas, és important distingir que *output/producte* fa referència als resultats immediats de les activitats de la intervenció, mentre que *outcome/impacte* es refereix als canvis, com conseqüència de la intervenció, en el problema o situació social que dona raó de ser a la intervenció.

La consideració d'aquests components ofereix un bon punt de partida per situar i abordar els **diferents tipus de preguntes que poden ser objecte dels diferents exercicis d'avaluació** (figura 2).

Figura 2. Tipus d'avaluacions



Font: Adaptació de Blasco (2009: 8).

Així, ens podem preguntar sobre les característiques, l'abast i la distribució de la problemàtica que vol redreçar el programa. D'afinar aquestes preguntes i de donar-hi resposta s'encarrega el que coneixem com a **avaluació de necessitats**. El capítol que segueix a aquesta introducció s'ocupa de descriure els punts de partida, orientacions i tècniques pròpies d'aquest tipus d'avaluació.

Ens podem preguntar també sobre els fonaments teòrics del disseny del programa (el que anomenem la seva teoria del canvi) i sobre la seva avaluabletat. En primer lloc, ens plantejaríem qüestions com: la naturalesa i abast del problema justifiquen els recursos mobilitzats? Els recursos són suficients per dur a terme les activitats previstes? Les activitats estan ben plantejades i permetran assolir els productes desitjats? Els productes esperats induiran els impactes perseguits? Com podem estar segurs de tot plegat? En segon lloc, ens preguntàrem: contempla el disseny del programa aquells indicadors i sistemes d'informació que permetran fer-ne una bona avaluació un cop es comenci a implementar? Aquests interrogants són objecte d'atenció de l'**avaluació del disseny**.

Un cop el programa comença a rutllar, ens pot interessar conèixer com funciona i s'executa a la pràctica; això és, com es produeix l'accés al programa, quin acaba essent el perfil de la població usuària, i quina cobertura s'aconsegueix, com es mobilitzen els recursos i tenen lloc les activitats, quantes hores, sessions o seguiments s'acaben realitzant, com de satisfets es mostren usuaris i gestors, etc. I pot ser interessant preguntar-se sobre el grau de fidelitat entre allò que es produeix a la pràctica i allò originalment previst en el disseny. Aquestes anàlisis són pròpies de l'**avaluació de la implementació**.

Finalitzat el programa (o si més no, una part de les seves intervencions), és moment de preguntar-se sobre els seus impactes, això és, sobre aquells canvis en la realitat del problema atès que són netament atribuïbles a la seva acció. Desentrellar els impactes d'una intervenció és una tasca que pot ser complexa i que passa per donar resposta a un interrogant bàsic (la pregunta pel "contrafactual"): com els hagués anat als participants d'un programa en cas que no hi haguessin participat? Òbviament aquest interrogant no és directament contestable; calen grups de control comparables als individus participants i diferents dades i mètodes per acabar estimant l'efecte de les intervencions. Aquest és el focus d'atenció de l'**avaluació de l'impacte**.

Finalment, una vegada coneguts els impactes del programa probablement ens interressi posar-los en relació amb els costos de dur-lo a terme i/o amb la capacitat d'impacte d'altres programes amb els mateixos objectius de canvi. Ens preguntariem, per tant, si la intervenció en qüestió val la pena o compensa econòmicament. Aquestes són les qüestions que es planteja l'**avaluació econòmica** dels programes, mitjançant diferents enfocaments, entre els més coneguts, el cost-efectivitat i el cost-benefici.

1.3. UNA GUIA PER AL TERCER SECTOR SOCIAL

Els continguts i enfocaments d'aquesta guia tracten de ser sensibles als interessos i necessitats d'avaluació de les entitats del tercer sector social. Si bé és cert que una part considerable dels conceptes i metodologies d'avaluació que aquí es presenten són aplicables a qualsevol política o programa d'intervenció social (independentment del sector que provingui), hem tractat de connectar amb les particularitats del tercer sector mitjançant els èmfasis posats en uns i altres tipus d'avaluació, els mètodes escollits i els exemples seleccionats.

Per tal de garantir aquesta connexió, els autors de la guia han dut a terme **sis entrevistes** en profunditat amb responsables d'entitats del sector amb un interès especial en l'avaluació dels seus programes, així com **dos focus groups** amb entitats de mida gran i mitjana (veure taula 1). L'objectiu tant de les entrevistes com dels grups de discussió va ser conèixer i debatre sobre els obstacles i oportunitats que confronten les entitats a l'hora d'incorporar els diferents tipus d'avaluació dins la seva cultura i pràctica rutinàries.

Aquesta guia vol contribuir a superar part d'aquests obstacles i a aprofitar part d'aquestes oportunitats tot oferint un marc de referència comú i, dins el seu caràcter introductor, el més exhaustiu possible a l'àmbit de l'avaluació de programes del tercer sector social.

Taula 1. Treball de camp realitzat

ENTREVISTES	GRUPS DE DISCUSSIÓ	
	ENTITATS GRANS	ENTITATS MITJANES
ATRA	Casal dels infants	ATRA
AMPANS	ABD	Fundació Esclerosi Múltiple
Plataforma Educativa	Fundació Pere Tarrés	Ser.Gi
Creu Roja	Plataforma Educativa	ACIDH
Fundació Joia	Fundesplai	El Redós
Intress	AMPANS	Dianova
	Suara	Càritas Girona
	Càritas Barcelona	Fundació Esclat
	Intress	Fundació Joia

Volem agrair la col·laboració de totes aquelles persones que han contribuït a l'elaboració d'aquesta guia aportant les seves idees, coneixements i punts de vista sobre l'avaluació i el tercer sector: Toni Codina, Roger Civit, Agnès Pàmies, Anna Ayneto, Anna Suñer, Enric Arqués, Gisela Carrés, Imma Lluva, Jordi González, Javier Martínez, Jordi Mir, Jordi Pascual, Lluís Moner, Lluís Puigdemont, Maite Balduque, Maria Bruno, Mercè Darnell, Míriam Feu, Mònica Canela, Mònica Plana, Montserrat Cervera, Núria Comas, Núria Vergé, Pilar Bosch, Raquel Carrión, Rosa Masriera, Salvador Cots i Susanna Roig.

2. AVALUACIÓ DE LES NECESSITATS

2.1. PUNT DE PARTIDA I MOTIVACIÓ

La raó de ser de qualsevol programa social, ja sigui del tercer sector o del sector públic, és donar resposta a una determinada problemàtica o necessitat social. Tanmateix, aquesta necessitat no sempre està formulada prèviament al disseny i execució del programa, sovint es parteix d'una idea difusa de quin és el problema al que es vol donar resposta o aquesta es desenvolupa de forma paral·lela. En alguns casos, fins i tot podríem pensar que la solució precedeix l'anàlisi del problema, de manera que disposem d'uns recursos i d'un pla d'acció sense estar segurs de si el problema realment existeix tal com ens l'imaginem.

REQUADRE 1: PENSAR EN LES NECESSITATS

Una entitat del tercer sector obté uns recursos públics per formar a dones que vulguin emprendre projectes empresarials. L'equip de l'entitat dissenya el programa i fa tots els preparatius per a la seva implementació, però es troben amb un petit problema: no troben demanda per aquest tipus de servei. Després de parlar amb diferents potencials usuàries formulen la hipòtesi de que no hi ha una necessitat de formació sinó de temps disponible, per la qual cosa seria més efectiu disposar d'escoles infantils on les famílies poguessin deixar els fills i així tenir temps per dedicar al projecte empresarial.

Font: (Exemple extret i adaptat del curs "Democràcia y decisiones públicas. Introducción al análisis de políticas públicas" de l'IGOP-UAB a la plataforma Coursera)

Si parlem específicament del tercer sector, hi ha algunes limitacions a tenir en compte. En primer lloc, és habitual que la definició de les problemàtiques socials i el disseny dels programes vinguin establerts o condicionats pels seus finançadors, donant poc marge per a la seva redefinició. Per altre banda, la disponibilitat i accés a dades no sempre permet una anàlisi acotat a l'àmbit territorial d'interès ni per als perfils d'usuaris concrets, i els esforços avaluatius se solen centrar en l'anàlisi d'allò sobre el que es té un major control: recursos, activitats i resultats immediats dels programes.

En qualsevol cas, i tenint en compte les limitacions mencionades, en aquest capítol abordem com orientar millor l'activitat de les entitats del tercer sector cap a les diferents problemàtiques socials, aportant elements d'objectivitat que permetin entendre i dimensionar les necessitats socials i preparar respostes adequades i proporcionals.

2.2. QUÈ VOL DIR AVALUAR LES NECESSITATS?

Una problemàtica social es pot entendre com una discrepància entre l'estat actual d'una qüestió (allò que és), i el seu estat desitjat (allò que hauria de ser). Així doncs, l'objectiu d'un programa social consisteix en modificar l'estat actual de les coses per aproximar-nos a la situació desitjada. L'avaluació de les necessitats aborda dues qüestions fonamentals per

afrontar aquesta missió: (1) **la comprensió de la problemàtica social i l'estimació del seu abast poblacional** en un territori concret, i (2) **la determinació del tipus i volum de serveis necessaris** per a donar-hi resposta.

Cal tenir en compte que no totes les avaluacions de les necessitats aborden les dues qüestions: mentre que algunes es focalitzen només en la primera, d'altres intenten respondre a totes dues. Aquesta diferència, però, no suposa dos enfocaments alternatius, sinó dos estadis del desenvolupament de l'avaluació de les necessitats.

2.2.1. Propòsits de l'avaluació de les necessitats

Tenint en compte les dues qüestions que aborda l'avaluació de les necessitats, podem distingir diferents propòsits que ens portarien a impulsar-la:

- **Planificar els serveis d'un nou programa:** En el cas de programes de nova creació, l'avaluació de les necessitats pot ser útil per dimensionar el volum de recursos i la provisió de serveis, incloent el grau o intensitat de prestació per perfils.
- **Focalitzar els serveis ja existents a les necessitats i usos reals:** Es pot esperar que hi hagi grups d'usuaris que malgrat no necessitar els serveis els estiguin utilitzant, mentre que d'altres, malgrat necessitar-los poden no saber de la seva existència. La detecció d'aquestes situacions ha d'ajudar a prendre decisions i introduir ajustos que permetin una millor focalització.
- **Millorar l'assignació territorial de recursos:** Fent una anàlisi territorialitzada es poden detectar àrees amb un nivell de necessitats observades superiors als d'utilització de serveis i viceversa. A partir d'aquí es poden prendre decisions relatives a la distribució dels recursos.
- **Anticipar les necessitats futures de serveis:** L'avaluació de les necessitats pot preveure l'evolució futura de la problemàtica analitzada, ajudant així a anticipar les necessitats futures de servei, i permetent pensar en estratègies d'intervenció a mig i llarg termini.
- **Mesurar la realitat abans i després del programa:** Quan es disposa d'una avaluació de necessitats realitzada fa uns anys, una de les utilitats de realitzar-ne una de nova és la de poder comparar ambdós diagnòstics. Tanmateix, com veurem en el capítol dedicat a l'avaluació d'impacte, cal tenir en compte que els canvis observats poden haver estat provocats per molts altres factors que no tenen a veure amb el programa.

2.2.2. En quin moment cal avaluar les necessitats?

S'acostuma a pensar que l'avaluació de les necessitats es fa prèviament al disseny del programa, tanmateix sovint es un procés paral·lel. Cal tenir en compte que la realitat social canvia al llarg del temps, com ho fan els programes i serveis que hi donen resposta, per la qual cosa cal avaluar les necessitats constantment i fer els canvis oportuns.

2.3. PREGUNTES I FASES DE L'AVALUACIÓ DE LES NECESSITATS

L'avaluació de les necessitats consta de diverses fases, cadascuna de les quals contesta unes preguntes diferents. Tanmateix, és recomanable que per respondre les preguntes d'una fase s'hagin respost abans les de les anteriors. Segons les preguntes en que estiguem interessats, podem planificar una avaluació més o menys ambiciosa:

Taula 2. Fases de l'avaluació de necessitats

FASE 1. DEFINICIÓ I ANÀLISI DE LA PROBLEMÀTICA

- Quantes persones de la població de referència pateixen la problemàtica social?
- Quins factors distingeixen aquest col·lectiu de la resta d'individus que conformen la població de referència?
- Què provoca la problemàtica social?

FASE 2. ESTIMACIÓ DEL TIPUS I VOLUM DE SERVEIS NECESSARIS

- Quins tipus de serveis haurien d'utilitzar les persones que pateixen la problemàtica? Amb quina intensitat i/o freqüència?
- Han de ser aquests serveis diferents en funció de la gravetat de la problemàtica i/o d'altres característiques dels individus?

FASE 3. ANÀLISI DE LA RELACIÓ ENTRE NECESSITATS, DEMANDA I UTILITZACIÓ DELS SERVEIS

Polítiques en funcionament:

- Quantes persones que necessiten els serveis no els utilitzen?
- Quins factors determinen l'existència de necessitats no ateses? Barreres econòmiques?
- Desconeixement dels serveis? Altres factors?
- Hi ha llistes d'espera? Quina és la gravetat de la problemàtica dels que hi són?

Noves polítiques:

- Hi ha serveis privats que ja s'estan utilitzant per fer front a la problemàtica? I mecanismes de caràcter familiar, com ara transferències de renda o provisió informal d'atenció?
- Quantes persones utilitzen aquest tipus de serveis o mecanismes?
- Quins factors sociodemogràfics o d'altres tipus determinen la seva utilització?

Font: Elaboració pròpia.

FASE 1: DEFINICIÓ I DIMENSIÓ DE LA PROBLEMÀTICA

Definició de la problemàtica

El primer pas per a l'avaluació de les necessitats és definir quina és la problemàtica a la que volem donar resposta. En aquest sentit resulta útil esbrinar com els diferents actors han definit el problema. Per això ens hem de preguntar què diuen els documents elaborats per l'Administració o altres entitats del tercer sector, quines són les percepcions dels diferents actors implicats, què n'opinen les persones afectades o com s'ha definit el problema en contextos similars al nostre.

La definició de la problemàtica no és una tasca senzilla, la majoria de problemes socials afecten a múltiples dimensions de la vida de les persones, i per tant, es poden explicar des de diferents perspectives. Així mateix, una necessitat social es pot concebre de forma aïllada o com a part d'una problemàtica major.

REQUADRE 2: EL PROBLEMA DEL RISC DE POBRESA

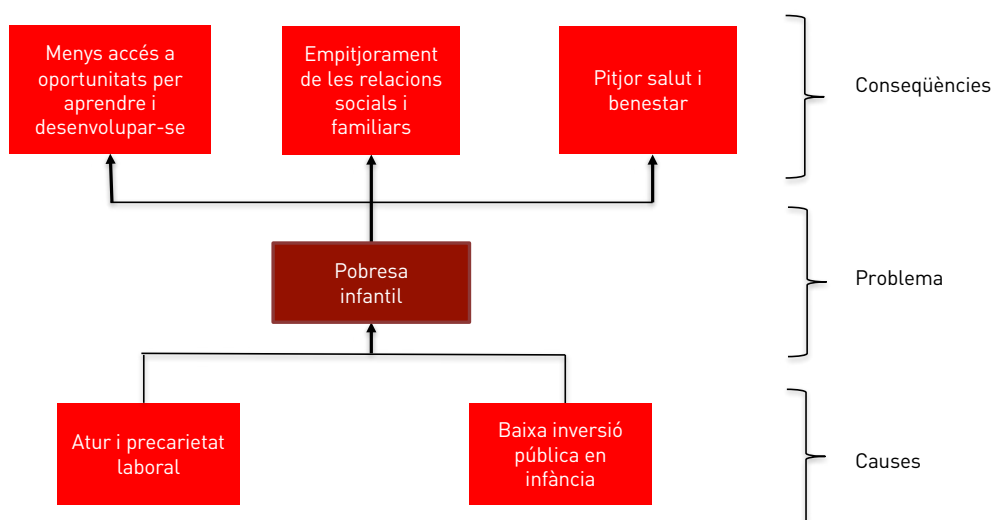
Segons l'Enquesta de Condicions de Vida (ECV) de l'INE, a l'any 2014 la taxa de risc de pobresa infantil a Catalunya es situa en el 29,9%, una de les més altes de la Unió Europea. En base a aquesta situació diferents actors han posat sobre l'agenda el problema de la pobresa infantil, proposant diferents línies de resposta que posen el focus en els infants, oferint serveis específics que garanteixin una major igualtat d'oportunitats entre ells. Tanmateix, una conceptualització alternativa del problema que moguéss el focus de la pobresa infantil a la pobresa familiar o laboral, previsiblement ens conduiria a buscar solucions que posen el seu focus en les famílies o les condicions laborals més que en els infants.

Font: http://www.tscat.cat/download/web/juliol/infancia/presentacio_pau_mari_klose.pdf

Un cop hem definit la problemàtica social, la següent tasca és l'anàlisi de les causes que la provoquen i les conseqüències que se'n deriven. Aquesta és una tasca que sovint no es té en compte, però quan es fa dona claus de com es pot mitigar o revertir el problema. El principal recurs per identificar les causes d'una problemàtica és la recerca social i l'evidència científica que aquesta pugui oferir. Si no es disposa d'aquest tipus d'informació, es recomana intentar entendre les causes del problema consultant informes de l'administració i del tercer sector i recollint l'opinió d'experts, professionals i dels propis afectats.

Una eina pràctica per recollir aquesta informació de forma visual és l'**arbre de problemes**. Per elaborar un arbre de problemes cal plasmar les relacions entre el problema, les causes i les conseqüències de forma gràfica, traçant línies de causa-efecte (veure figura 3).

Figura 3. Exemple d'arbre de problemes



Font: Elaboració pròpia a partir de Unicef-Catalunya (2015): <https://old.unicef.es/cat/actualitat-documentacio/publicacions/infografia-pobresa-infantil-catalunya>

Dimensió de la problemàtica

Un cop s'ha definit conceptualment el problema, cal establir l'àmbit **territorial d'anàlisi** així com, dins d'aquest territori, la **població de referència** sobre la qual pretenem detectar la problemàtica social. El subgrup de persones en els que identifiquem la problemàtica social constituïran **població diana**, i serà, en última instància, a qui es dirigirà el programa. Per exemple, si la problemàtica en que estem interessats és la pobresa infantil al barri del Raval de Barcelona, considerarem que el Raval és l'àmbit territorial, que els infants residents al Raval són la població de referència, i que el subconjunt d'infants que es troben sota el llindar de la pobresa són la població diana.

La majoria de problemes socials no es distribueixen a l'atzar entre la població, sinó que hi ha característiques individuals, territorials i altres factors que augmenten la probabilitat de patir-los. Malgrat que en l'apartat anterior ja hem definit el problema, les seves causes i conseqüències, es recomana fer una anàlisi acotada de la població diana, que permeti saber com són i així poder dissenyar un programa que els sigui adequat. Com veurem posteriorment, són diversos els mètodes que existeixen per estimar la dimensió del problema

FASE 2: ANÀLISI DE LES NECESSITATS DE SERVEI

Un cop hem definit el problema, les seves causes i conseqüències, i coneixem a la població diana, cal pensar en els serveis o programes que volem oferir i el volum en que ho volem fer.

Idealment, podrem identificar diversos tipus de servei i **comparar la seva efectivitat**. L'efectivitat d'un programa és la seva capacitat d'aconseguir uns resultats determinats, mentre que el seu cost-efectivitat és la relació entre la consecució d'aquests resultats i els costos del programa, expressats en una ràtio que mostra el cost per a cada unitat de resultat. La disponibilitat d'aquest tipus d'informació permet saber quins programes són més efectius i eficients, i ens dóna un criteri per escollir de forma més encertada.²

Tanmateix, en absència d'evidència empírica sobre l'efectivitat d'un programa o servei, un recurs habitual és la **constitució d'un grup d'experts**, ja siguin investigadors, gestors, professionals, etc. El grup d'experts té per objectiu aprofitar la seva experiència i determinar quins i quants són els serveis necessaris per atendre la problemàtica social identificada. Tanmateix, es tracta d'un mètode que pot comportar biaixos, ja que fa servir l'experiència personal, la intuïció i l'evidència anecdòtica, més que no pas criteris de base científica. Finalment, un darrer recurs és la **definició de les necessitats de serveis de forma comparativa**. Es tracta de determinar les necessitats de serveis de la població d'un territori a partir de la utilització real de serveis en un altre territori. Tanmateix, la credibilitat d'aquest exercici depèn de dues premisses: 1) l'abast de la problemàtica i llurs factors determinants són similars en els dos territoris, i 2) es considera que el nivell d'utilització de serveis en el territori de comparació és el més adequat per donar resposta a la problemàtica que hem definit.

En qualsevol cas, independentment de quins siguin els mètodes emprats, el resultat d'aquesta fase de l'avaluació ha de ser el coneixement sobre quins serveis i en quina quantitat respondrien de forma satisfactòria a les necessitats definides.

FASE 3: ANÀLISI DE LES NECESSITATS I UTILITZACIÓ DE SERVEIS

Arribats a aquest punt, hem quantificat les necessitats socials i hem planificat l'abast dels serveis. Tanmateix, el fet de patir una problemàtica social i tenir un programa a disposició, no en garanteix l'ús, de la mateixa manera no patir la problemàtica no implica que no s'acabi fent ús del programa. És per això que cal complementar l'anàlisi amb dos elements addicionals: la demanda i l'oferta de serveis. Mentre que la primera fa referència al desig per part dels individus de rebre el serveis del programa, l'oferta té a veure amb el volum i abast dels serveis que es posen a la seva disposició, i junts determinen la utilització efectiva del servei.

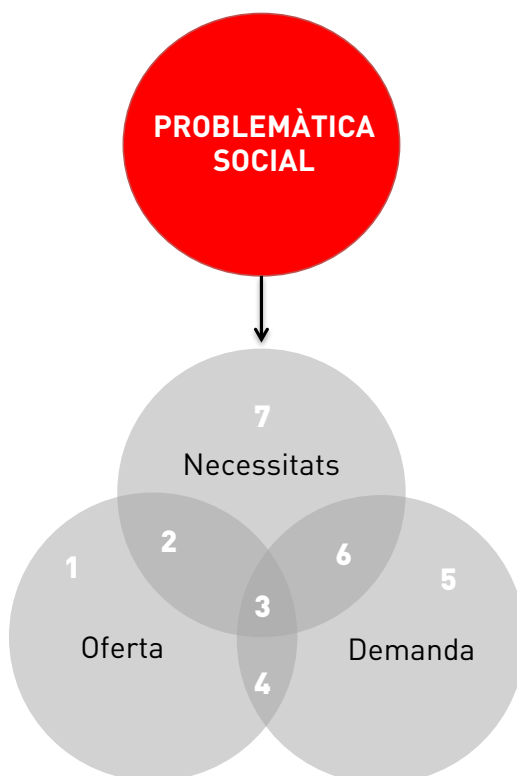
² El capítol de la guia sobre l'avaluació del disseny dedica un apartat a l'enfocament "què funciona" i a la revisió d'evidències com a estratègies per incrementar les probabilitats d'èxit dels programes.

Cal puntualitzar, que aquesta anàlisi generalment la fem per a programes que ja estan en funcionament, ja que no podem valorar l'oferta i demanda d'un programa que està en fase de disseny i encara no existeix. Tanmateix, quan dissenyem un programa nou, algunes de les persones a les que es dirigeix probablement estaran utilitzant serveis similars, ja siguin de caràcter públic, privat o informal. En aquests casos, doncs, estaríem analitzant l'oferta i demanda de serveis similars als del programa que estem dissenyant.

REQUADRE 3: POSSIBLE UTILITZACIÓ D'UN RECURS DE LLEURE

Una entitat de l'àmbit del lleure educatiu detecta que hi ha una manca d'oferta i oportunitats educatives en un determinat barri, per la qual cosa es planteja impulsar un esplai de caps de setmana amb una mirada comunitària. L'entitat ha definit la seva població diana, infants de 4 a 12 anys residents al barri, i ha quantificat el seu abast poblacional així com les seves característiques sociodemogràfiques. Tanmateix, desconeix si la població diana podria estar realitzant altres activitats similars fora del barri, i quina seria la seva predisposició a participar en el nou programa. És per això que l'entitat decideix parlar amb els agents clau del territori que tenen contacte amb la població diana, especialment les escoles, a través de les quals pensa poder entrevistar o fins i tot passar un petit qüestionari a les famílies.

Figura 4. Necessitat, demanda i oferta de serveis



Font: Adaptat de Williams i Wriugh (1998).

Tal com es mostra en la figura 4, la interacció entre els elements identificats configura diferents escenaris a tenir en compte:

- **Oferta de serveis innecessaris sense demanda (àrea 1).** Es tracta d'una situació poc freqüent, sovint relacionat amb un intent d'induir la demanda des de l'oferta. Un exemple seria un curs de formació ocupacional que no ha rebut prou sol·licituds d'inscripció com per cobrir les places disponibles, i en què les poques persones inscrites responen a un perfil diferent a la població diana. En aquestes situacions, la millor solució sol ser redimensionar el servei, o en alguns casos com seria el mencionat, eliminar-lo, i destinar els recursos a altres necessitats no cobertes.
- **Utilització de serveis per part de persones que el necessiten (àrees 2 i 3).** Aquesta és la situació que desitjaríem veure amb més freqüència, ja que indica que el programa està ben focalitzat en la població diana. Tanmateix, cal tenir en compte la diferència entre les àrees 2 i 3, i esbrinar perquè hi ha una part de la població diana que no demana els serveis (àrea 2).
- **Demanda no necessària atesa (àrea 4).** Es tracta d'una situació preocupant, sobretot en situacions en que al mateix temps hi ha demanda necessària no atesa, ja que impedeix l'ús del recurs a persones que realment ho necessiten.
- **Demanda no atesa, necessària (àrea 6) i no necessària (àrea 5).** Serien persones que demanen el servei però per algun motiu no hi tenen accés. Podríem estar parlant de llistes d'espera o d'altres tipus de barreres d'entrada. En aquests casos és important valorar si es tracta de persones que caurien dins la definició de la població diana (àrea 6) o no (àrea 5). Una estratègia habitual és valorar el grau de necessitat, prioritzant els casos en què aquest sigui major.
- **Necessitats no cobertes sense demanda expressada (àrea 7).** Es tractaria d'un col·lectiu de persones que tot i patir la problemàtica definida no tenen serveis a l'abast ni els demanen.

El mètode més fiable per estimar les àrees descrites és l'enquesta i els registres administratius. A continuació repassem aquests mètodes i d'altres que han de permetre implementar les tres fases descrites de l'avaluació de les necessitats.

2.4. MÈTODES PER AVALUAR LES NECESSITATS

Al llarg de les tres fases descrites en l'apartat anterior, ens podem servir de diferents tècniques d'investigació, aquestes es divideixen entre les quantitatives i qualitatives. Els mètodes quantitius, com l'explotació d'enquestes o registres administratius, són centrals per quantificar la població diana, definir les seves característiques o estimar el volum de serveis per fer-hi front. Per la seva banda, els mètodes qualitius, com els grups de discussió i les entrevistes, ens permeten obtenir informació en profunditat sobre percepcions i opinions que ens poden ajudar a conceptualitzar el problema així com entendre i explicar els fets observats.

Taula 3. Mètodes i tècniques segons la fase d'avaluació de les necessitats

FASE 1. DEFINICIÓ I ANÀLISI DE LA PROBLEMÀTICA

- Entrevistes en profunditat i grups de discussió
- Enquestes
- Registres administratiu
- Combinació de fonts secundàries

FASE 2. ESTIMACIÓ DEL TIPUS I VOLUM DE SERVEIS NECESSARIS

- Revisió d'estudis sobre efectivitat i cost-efectivitat
- Grups de discussió amb experts
- Combinació de fonts secundàries

FASE 3. ANÀLISI DE LA RELACIÓ ENTRE NECESSITATS, DEMANDA I UTILITZACIÓ DELS SERVEIS

- Entrevistes i grups de discussió
- Enquestes
- Registres administratius

Font: Elaboració pròpia.

Tècniques qualitatives

- **Entrevistes en profunditat i grups de discussió.** Els grups de discussió i entrevistes en profunditat són un recurs que pot ser molt útil en les tres fases de l'avaluació, és a dir, ja sigui per conceptualitzar el problema o necessitat social a la que volem donar resposta, per identificar diferents alternatives pel que fa al tipus i volum de serveis que podem oferir o per entendre la utilització de serveis que estem observant per part dels usuaris. Ens permetrà obtenir informació qualitativa sobre percepcions i opinions, així com coneixement expert provinent de potencials beneficiaris, ciutadans, professionals, experts o altres actors que es considerin rellevants.

REQUADRE 4: GRUPS DE DISCUSSIÓ AMB BENEFICIARIS

Després de detectar un descens de l'assistència dels infants i les seves famílies al centre obert, decidim organitzar un grup de discussió amb les famílies per entendre les seves percepcions i identificar possibles explicacions que permetin pensar en solucions pràctiques.

Tècniques quantitatives

- **Enquestes.** Les enquestes són una font d'informació molt útil per a l'avaluació de les necessitats, especialment pel que fa a les fases 1 i 3, és a dir, per estimar l'abast poblacional de la problemàtica social per analitzar la utilització del servei.

Si optem per **realitzar una enquesta pròpia**, podrem assegurar-nos que aportarà prou informació per respondre les preguntes que considerem d'interès i que els resultats seran generalitzables, és a dir, que seran estimacions acurades per al conjunt d'un territori. Per contra, els seus inconvenients són l'elevat cost i el fet d'exigir un nivell d'expertesa notable.

Una alternativa menys costosa és optar per l'**explotació d'una enquesta ja existent**. En molts casos, és possible aconseguir les microdades d'aquestes enquestes, és a dir, la informació detallada de tots i cadascun dels individus que varen ser entrevistats. Aquesta opció, tanmateix, té dues limitacions: per una banda, algunes d'elles poden no ser representatives per certes unitats territorials, i per l'altre, les variables ja venen definides

- **Registres administratius.** Alternativament a l'explotació d'enquestes, els registres administratius poden ser un recurs molt útil, sobretot pel que fa a les preguntes de la fase 3 de l'avaluació de les necessitats, és a dir, la relació entre necessitats i utilització dels serveis. El seu anàlisi permetria delimitar el nombre de persones que usen un servei, les seves característiques sociodemogràfiques i el seu grau de necessitat en relació a la problemàtica social definida. La principal limitació d'aquesta font d'informació és que no conté dades sobre les persones que no utilitzen el servei

Explotació quantitativa d'enquestes i registres

Independentment de quina sigui la font de les dades, el primer pas és calcular les necessitats per a cadascun dels individus sobre els quals disposem d'informació. Per fer-ho, podem optar per construir **indicadors dicotòmics**, que indiquin l'existència o no del problema en un individu (p.ex. situació de pobresa, d'atur, d'abandonament educatiu prematur, etc.), o bé indicadors de **categories de necessitat**, que contenen diferents valors per distingir graus de necessitat (nivells de dependència o d'ocupabilitat).

REQUADRE 5: INDICADOR DE RISC DE POBRESA

El risc de pobresa és un indicador dicotòmic que es calcula a partir de la renda total disponible de l'individu. Si aquesta és inferior al 60% de la mediana nacional es considerarà que la persona es troba en situació de risc de pobresa. En el cas de Catalunya, el llindar es fixa com el 60% de la mediana dels ingressos disponibles anuals equivalents de la població catalana (després de transferències socials). D'aquesta manera, amb dades del 2015, per una llar amb dos adults i un infant considerarem que els seus membres es troben en situació de risc de pobresa si en el seu conjunt disposen d'una renda anual inferior als 17.401,1 euros anuals, mentre que per un individu que viu sol considerarem que es troba en situació de risc de pobresa si té una renda disponible de menys de 9.667,3 euros anuals.

Un cop sabem de l'existència o no de la problemàtica social en cada individu, podem procedir a l'**anàlisi**. A continuació expliquem breument diferents tipologies d'anàlisi.

- **Anàlisi univariant.** Una part important de les preguntes de l'avaluació de les necessitats es pot abordar fent un anàlisi univariant, és a dir, centrant-nos en una sola variable. Dins d'aquest tipus d'anàlisi podem pensar en multitud de mesures estadístiques, com les freqüències, la mitjana, la mediana, la moda, etc. Probablement, la mesura més popular en l'avaluació de les necessitats és la prevalença. Es tracta d'un concepte important del camp de l'epidemiologia, i es defineix com el percentatge de persones que pateixen un determinat problema o condició. Es recomana començar l'anàlisi descrivint les principals variables de les que disposem d'informació.
- **Anàlisi bivariant.** Un cop caracteritzada la població de referència i diana, l'anàlisi bivariant ens pot aportar informació sobre la relació que existeix entre les característiques sociodemogràfiques dels individus, el seu grau de necessitat i ús de serveis. D'aquesta manera podem obtenir una fotografia més acurada sobre com són les persones afectades per la problemàtica social que estiguem analitzant. L'objectiu d'aquesta aproximació és esbrinar si existeix alguna relació entre dues variables. La **taula de contingència** és una de les eines d'anàlisi bivariant més populars.

REQUADRE 6: RISC DE POBRESA I IMMIGRACIÓ (TAULA DE CONTINGÈNCIA)

L'Enquesta de Condicions de Vida del 2014 ens mostra com el risc de pobresa és considerablement més elevat en el grup d'infants amb pares d'origen estranger:

	RISC DE POBRESA	
	SÍ	NO
Infants amb pares d'origen estranger	65,10%	34,9%
Infants amb pares d'origen autòcton	17,3%	82,7%

Font: http://www.tscat.cat/download/web/juliol/infancia/presentacio_pau_mari_klose.pdf

En aquest tipus d'anàlisis, per verificar si les diferències detectades són reals o si simplement són producte de l'atzar (biaix de la mostra), cal realitzar un test de significació estadística.³ Així mateix, l'anàlisi bivariant no ens permet saber si la relació entre les dues variables d'interès es troba influenciada per altres factors intervinents.

³ Per verificar si la informació que ofereixen aquestes taules són representatives, existeixen diferents tests de significació estadística. Per a més informació sobre aquestes operacions, consulteu el manual introductori d'estadística de Freedman et al. (1993).

- **Anàlisi multivariant.** L'anàlisi multivariant ens permet estudiar la relació entre una variable d'interès (variable dependent) i un conjunt de factors (variables independents). El mètode multivariant més utilitzat és l'anàlisi de **regressió**, que ens permet aïllar la relació entre la variable d'interès (p.e. el risc de pobresa) i un determinat factor sociodemogràfic (p.e. el nivell d'estudis), tenint en compte les relacions existents entre aquest factor i d'altres que també poden estar relacionats amb la variable dependent (p.e. l'edat i el sexe). Aquest mètode és útil per comparar la influència que tenen diferents factors en la variable d'interès.

Combinació de fonts secundàries: mètodes d'estimació indirecta

Per últim, quan volem estimar la persistència d'un problema o la utilització d'un servei en una unitat territorial sobre la qual no tenim prou dades, podem recórrer a l'estimació indirecta mitjançant la combinació de diferents fonts secundàries.⁴ Aquestes metodologies combinen informació específica de la unitat territorial analitzada i informació corresponent a una unitat superior de la qual la primera en forma part.

Un exemple seria una entitat que vol estimar l'abast dels problemes de dependència i el volum de serveis necessaris en un municipi concret. Podria consultar, a nivell de província, el percentatge de persones que pateixen dependència segons les variables que es consideren rellevants: En aquest cas serien grup d'edat, sexe i nivell educatiu. Al mateix temps, amb el Cens d'Habitants o el Padró Continu del municipi, podria saber la distribució de la població del municipi per a aquestes tres variables. Un cop disposa d'aquesta informació, només li caldria aplicar a cada grup la taxa de prevalença que li correspondria a nivell provincial.

REQUADRE 7: PROJECCIÓ DE POBLACIÓ I ÚS DE RECURSOS

Suposem que el 50% dels homes sense estudis de més de 85 anys de la província de Barcelona són dependents. Si el municipi en qüestió té 100 homes sense estudis i majors de 85 anys, estimarem que 50 d'ells es trobaran en situació de dependència. Ara, imaginem que coneixem el rati d'utilització de serveis residencials per part del perfil esmentat en cas de dependència, i és del 50% a la província de Barcelona. En aquest cas, considerarem que dels 50 homes sense estudis de més de 85 anys i dependents, 25 faran ús dels serveis residencials que estem planificant.

En qualsevol cas, hem de tenir en compte que la fiabilitat i viabilitat d'aquest tipus d'anàlisis depèn de la disponibilitat d'una enquesta que permeti estimar la prevalença a un nivell territorial superior, que aquesta contingui indicadors relacionats amb la problemàtica també disponibles al Cens, i que les dues fonts d'informació siguin temporalment pròximes.

⁴ S'anomena font d'informació secundària qualsevol que no hagi estat directament generada per qui l'està utilitzant.

2.5. RECURSOS

Manuels de referència:

- Casado, D. Avaluació de necessitats socials. Barcelona: Ivàlua, 2009. (Col.lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 2).
- Freedman, D., Pisani, R., Purves, R., i Adhikari, A. Estadística. Barcelona: Antoni Bosch editor, 1993.

Enquestes:

Els instituts oficials d'estadística ofereixen informació de tota mena provinent d'enquestes i estadística oficial que ens pot ser útil per a una avaluació de necessitats. Recomanem navegar per les seves pàgines web i explorar tot el que ofereixen, tot i així a continuació destaquem algunes de les principals fonts d'informació que hi podem trobar:

INE:

- Cens de la població (2011):
http://www.ine.es/censos2011_datos/cen11_datos_inicio.htm
- Encuesta nacional de salud (2011-12):
http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176783&menu=resultados&idp=1254735573175
- Encuesta europea de Salud en España (2014)
http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176784&menu=resultados&idp=1254735573175
- Encuesta de condiciones de vida (2015)
http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=ultiDatos&idp=1254735976608
- Encuesta de presupuestos familiares (2015)
http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176806&menu=ultiDatos&idp=1254735976608
- Encuesta sobre discapacidades, autonomía personal y situaciones de dependencia (2008)
http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176782&menu=resultados&idp=1254735573175
- Encuesta sobre las personas sin hogar (2012)
http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176817&menu=resultados&secc=1254736194829&idp=1254735976608

- Encuesta sobre equipamiento y uso de tecnologías de información y comunicación en los hogares (2016)

http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=resultados&secc=1254736194579&idp=1254735976608

Idescat:

- Padró municipal d'habitants (2016)
- Enquesta de condicions de vida i hàbits de la població (2011)

<http://www.idescat.cat/pub/?id=pmh>

<http://www.idescat.cat/pub/?id=ecvhp>

Altres enquestes:

- Enquesta de Salut de Catalunya (2015):

http://salutweb.gencat.cat/ca/el_departament/estadistiques_sanitaries/enquestes/esca//

Nota:

Aquestes URL estan actives en el moment de publicació de la guia, però poden patir modificacions posteriors.

2.6. REFERÈNCIES

Blasco, J. *Com iniciar una avaluació: oportunitat, viabilitat i preguntes d'avaluació*. Barcelona: Ivàlua, 2009. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 1)

Casado, D. *Avaluació de necessitats socials*. Barcelona: Ivàlua, 2009. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 2)

Freedman, D., Pisani, R., Purves, R., i Adhikari, A. *Estadística*. Barcelona: Antoni Bosch editor, 1993.

Williams, R.; Wright, J. *Health needs assessment: epidemiological issues in health needs assessment*. *British Medical Journal* (1998), núm. 316, p. 1379-1382.

3. AVALUACIÓ DEL DISSENY

3.1. PUNT DE PARTIDA I MOTIVACIÓ

Dins el tercer sector social (igual com succeeix en altres sectors i àmbits d'intervenció), no és infreqüent que els programes es dissenyin sense una reflexió acurada sobre els seus objectius i la forma com s'espera assolir-los i avaluar-los. Les presses amb què acostumen a gestar-se les intervencions, el fet que una part important dels programes que proveeixen les entitats es financin a través de licitacions i convocatòries que pre-defineixen els termes bàsics del seu disseny (objectius, població diana, activitats, indicadors, etc.), juntament amb altres obstacles com la poca atenció prestada a l'evidència empírica disponible o la manca d'expertesa metodològica..., tot plegat fa que ben sovint el disseny dels programes respongui més a tradicions i intuïcions que a una anàlisi de necessitats acurada i a una consideració atenta sobre les diferents alternatives d'intervenció.

Així, no és habitual que els programes (socials, laborals, de salut, formatius...) s'acompanyin d'un document en el qual es detallin els fonaments teòrics que es troben rere les seves activitats i objectius de canvi. D'igual forma, no és infreqüent que les intervencions comencin a implementar-se sense tenir ben precisats els indicadors, les dades i les fonts d'informació que permetran documentar-ne el procés i, eventualment, els seus impactes.

L'avaluació del disseny s'adreça justament a estudiar i afinar la formulació dels programes, els seus objectius, components, fonaments teòrics i sistemes d'avaluació, amb l'objectiu final de millorar la implementació i capacitat d'impacte de les intervencions. Des d'aquest punt de vista, conèixer els mecanismes i aplicacions d'aquest tipus d'avaluació és un primer pas per contrapesar l'abast dels obstacles abans esmentats.

3.2. QUÈ VOL DIR AVALUAR EL DISSENY D'UN PROGRAMA

Hauríem de començar fent explícit que no existeix un corpus teòric ni metodològic consolidat sobre què és i com s'ha de dur a terme l'avaluació del disseny del programes (Blasco, 2009). Aquesta circumstància afegeix dificultat a l'elaboració d'aquest capítol, al temps que el fa especialment rellevant en el marc d'un context d'actuació, el tercer sector social, que intenta cimentar les seves apostes i pràctiques d'avaluació.

A efectes de claredat expositiva, direm que l'avaluació del disseny s'ocupa de dos aspectes nuclears de la definició dels programes: els fonaments lògics de la intervenció (el que anomenem la **"teoria del canvi"** del programa), d'una banda, i les seves **condicions d'avaluabilitat** (formulació d'objectius, indicadors i sistemes d'informació), de l'altra.

Avaluar el disseny d'un programa és quelcom que podem fer ex-ante, això és, abans de la seva implementació, o ex-post, per tant, com a exercici d'avaluació en sí mateix. En el primer cas (enfocament **ex-ante**), l'avaluació del disseny ens permetrà anticipar el potencial d'una proposta d'intervenció per resoldre el problema que la motiva i suggerir millores en la seva formulació per fer-la més efectiva i avaluable. En el segon cas, l'avaluació **ex-post** del disseny ajuda a interpretar les raons dels possibles resultats d'implementació i impacte del programa, siguin aquests positius o negatius.

En tot cas estem davant d'una **aproximació no empírica**: es tracta de valorar la formulació teòrica del programa en tant que estratègia que pretén solucionar el problema i generar coneixement sobre els seus resultats. L'objectiu, doncs, no és avaluar l'execució pràctica del programa, ni els seus impactes sobre la realitat.

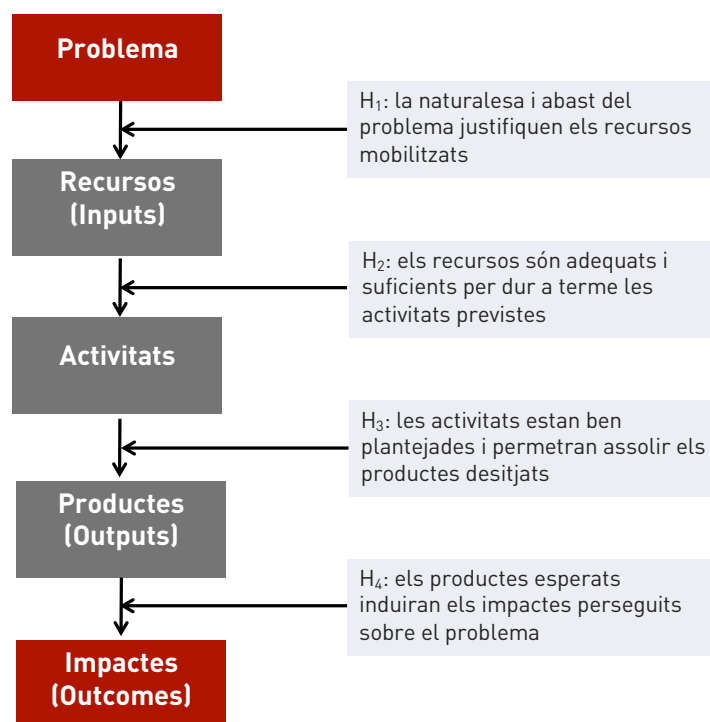
3.3. PREGUNTES I FOCUS DE L'AVALUACIÓ DEL DISSENY

3.3.1 Els fonaments lògics de la intervenció: la teoria del canvi

De forma general, entenem la teoria del canvi d'un programa com aquell conjunt d'hipòtesis o assumpcions que fan explícit com s'espera que el programa funcioni fins a assolir els seus objectius. Altrament dit, la teoria del canvi fa explícites les hipòtesis sobre com se suposa que ha de funcionar el programa i per què s'espera que es generin els impactes desitjats. Com veurem, concebre el programa des d'una teoria del canvi, com un conjunt d'hipòtesis, no només ajuda a dissenyar-lo, sinó que orienta les preguntes d'avaluació que serà pertinent formular-se ex-post.

Seguint la definició abans establerta, podem arribar a desagregar la teoria del canvi d'un programa en cadascuna de les hipòtesis que connecten els seus diferents components, tal i com es mostra en la figura 5. D'aquesta manera establiríem: 1) el tipus de problema que es vol resoldre justifica mobilitzar determinats recursos; 2) si mobilitzem aquests recursos, podem dur a terme les activitats planificades; 3) si realitzem aquestes activitats, generarem determinats productes; 4) si generem aquests productes, assolirem els impactes desitjats en els outcomes d'interès i mitigarem el problema de partida.

Figura 5. Teoria del canvi: hipòtesis del programa



Font: Elaboració pròpia, adaptat de Blasco (2009: 11).

Aquest conjunt d'hipòtesis parcials es poden agrupar en dos conjunts d'assumpcions teòriques més generals, de tal manera que podríem entendre la teoria del canvi d'un programa com la suma de dues "teories" (Rossi et al. 2004):

- **Teoria de la implementació:** recull el conjunt d'hipòtesis relatives a la posada en marxa i execució de la intervenció; hipòtesis 1, 2 i 3 de l'esquema anterior.
- **Teoria de l'impacte:** recull la hipòtesi causal general de la intervenció (si la duem a terme mitigarem el problema); hipòtesi 4.

REQUADRE 8: TEORIA DEL CANVI – EN TÀNDEM (AFEV)

[Teoria del canvi = Teoria de la implementació + Teoria de l'impacte]

La teoria de la implementació sosté la hipòtesi de què amb uns determinats recursos i formes organitzatives, es proveiran uns serveis adreçats a una població concreta que en farà un determinat ús, mentre que la teoria de l'impacte sosté que aquests serveis produiran un canvi en relació al problema al que inicialment es volia donar resposta.

En Tàndem és un programa de mentoria social i educativa desenvolupat per l'entitat AFEV, que disposaria de la següent teoria del canvi:

- **Teoria de la implementació:** Durant dues hores a la setmana al llarg de tot un curs escolar, un/a estudiant universitari (mentor) acompanya individualment un jove en els moments claus del seu recorregut educatiu per estimular les capacitats, l'autonomia i la confiança personals, l'autoestima i motivar l'interès per a estudiar d'aquests joves i infants. Els acompanyaments es fan amb el seguiment i suport d'un educador social. Prèviament els mentors reben una formació.
- **Teoria de l'impacte:** Els infants milloren en diferents aspectes, com l'autoestima, la motivació i les habilitats de relació.

Font: <http://entandem.cat/>

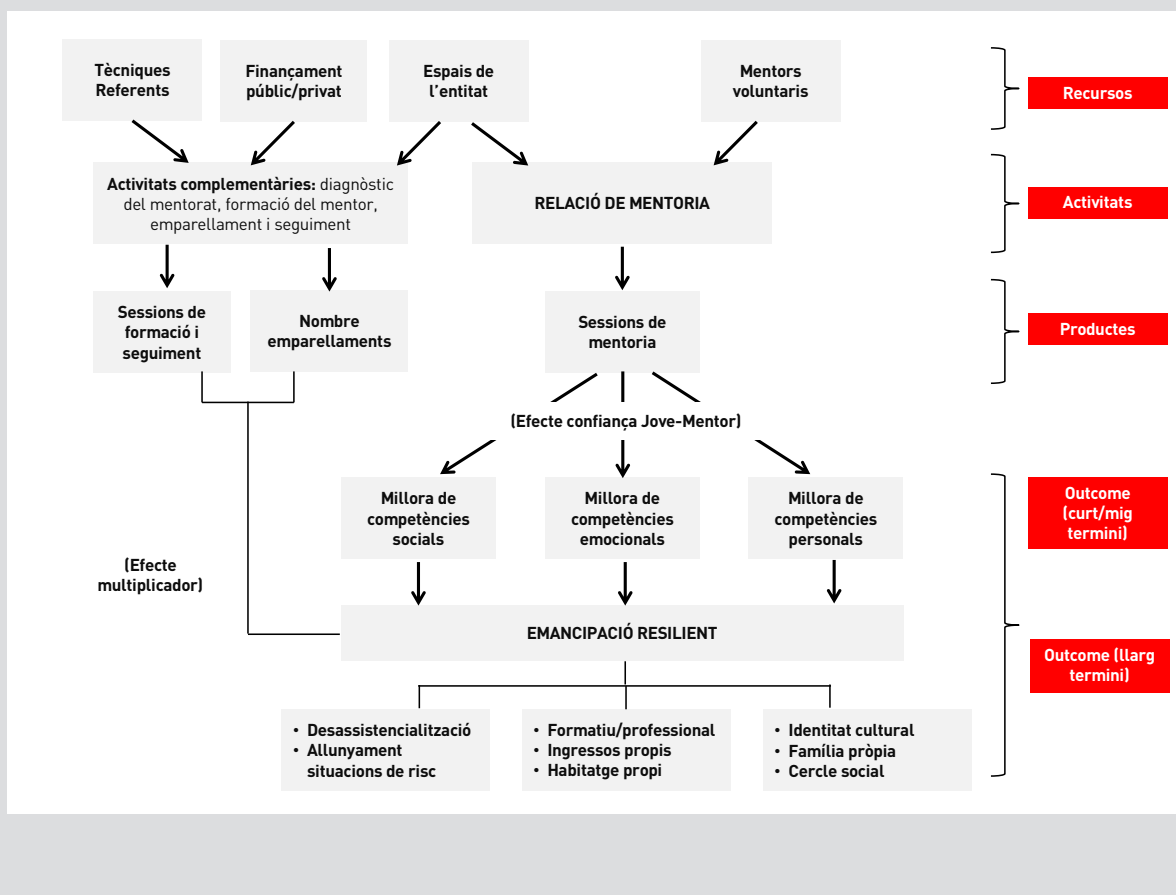
Probablement, la manera més habitual d'expressar la teoria del canvi és de forma gràfica, mitjançant un diagrama de fluxos conegut com a **model lògic**. Aquesta representació gràfica pot ser útil a efectes d'assegurar la coherència interna de les seqüències lògiques (i comprovar que no han quedat recursos sense activitat o activitats sense recursos, per posar un exemple) i per comunicar de forma més efectiva els esborranys de la teoria del canvi als diferents actors que hi aporten el seu criteri i informació.

El model lògic d'un mateix programa pot tenir diferents graus de complexitat i sofisticació. Una regla important per triar el nivell de sofisticació en la representació del model lògic és que el gràfic resultant sigui visualment atractiu i fàcilment comprensible i comunicable (veure l'exemple il·lustrat en la figura 6).

REQUADRE 9: EL MODEL LÒGIC DEL PROGRAMA REFERENTS, DE L'ENTITAT PUNT DE REFERÈNCIA

El programa Referents s'adreça a jovent extutelat i en situacions de risc d'exclusió social, i té com a objectiu facilitar el seu procés d'emancipació mitjançant la millora de les seves competències socials, emocionals, formatives i laboral. L'activitat bàsica del programa són les sessions de mentoria, que corren a càrrec d'adults voluntaris (<http://www.puntdereferencia.org/>).

Figura 6. Exemple de model lògic



Font: Elaboració pròpia.

L'avaluació del disseny es pregunta, per tant, sobre el grau de plausibilitat de cadascuna de les hipòtesis que componen la teoria del canvi del programa. Si l'assoliment dels objectius previstos per una intervenció social depèn del compliment d'una cadena d'hipòtesis que lliga els recursos amb les activitats, els productes i, finalment, els impactes desitjats, aleshores cadascuna d'aquestes hipòtesis ha de ser plausible per tal que el conjunt de la teoria també ho sigui. Recordem que es tracta d'una plausibilitat "teòrica", no confrontada encara amb el funcionament real del programa. Per tant, la conclusió sobre la validesa de la teoria serà més robusta si el judici sobre la plausibilitat de cada baula de la cadena no només se sustenta en percepcions i el sentit comú dels qui la valoren, sinó també en algun tipus d'evidència provinent de les ciències socials o d'altres avaluacions (veure apartat 3.4.1.).

3.3.2 Preguntes sobre l'avaluabilitat: objectius, indicadors i sistemes d'informació

Parlem aquí del segon centre d'interès nuclear de l'avaluació del disseny: les condicions d'avaluabilitat dels programes. Aquest focus d'atenció ens du a parlar de com es dissenyen els objectius de la intervenció, els seus indicadors i els seus sistemes d'informació. En altres paraules, es tracta d'analitzar **fins a quin punt el disseny del programa reuneix els requisits que n'hauran de permetre l'avaluació, tant de la implementació com del seu impacte.**⁵

Definir el que el programa persegueix: objectius estratègics i operatius

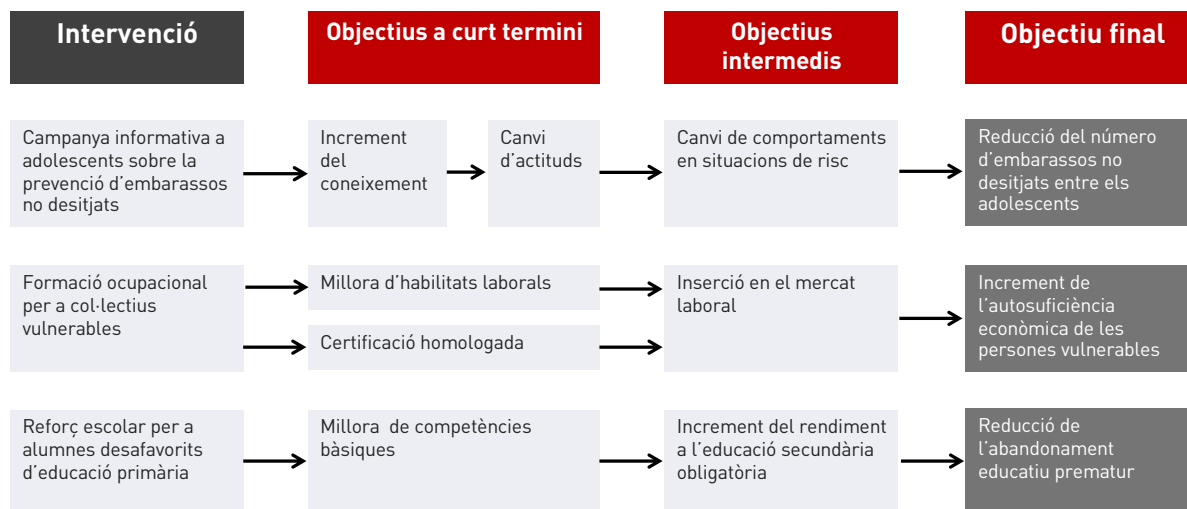
De forma general, els objectius del programa especifiquen les seves finalitats, allò que vol aconseguir. Els objectius poden ser de naturalesa estratègica o operativa:

- **Objectius estratègics:** fan referència a l'impacte esperat de la intervenció, això és, els canvis que es pretén que es produeixin en el problema o situació social que ha motivat la intervenció. Es vinculen, per tant, als **outcomes** del programa: reduir el nivell d'assistencialització de la gent gran, incrementar l'autonomia personal i laboral dels joves amb trastorns mentals, millorar les oportunitats educatives dels alumnes immigrants, etc.

Els objectius estratègics acostumen a diferenciar-se en funció del termini en què es preveuen assolir. Es pot parlar així d'una seqüència temporal d'objectius o "estructural d'impactes" de la intervenció (McLaughlin i Jordan, 2004), la qual sol incloure: objectius o impactes a **curt termini** (directament causats per la intervenció), **intermedis** (que resulten dels anteriors) i **finals** (fruit de l'acumulació dels impactes intermedis) (veure exemple en la figura 7). A mida que, en aquesta seqüència temporal, ens acostem cap a l'assoliment del objectius finals, acostumen a incrementar-se els beneficis esperats per al conjunt de la societat (no només per als participants en el programa) (Ruegg i Feller, 2003).

⁵ En capítols posteriors es detallen les pautes principals de l'avaluació de la implementació i de l'avaluació de l'impacte dels programes.

Figura 7. Seqüència temporal dels objectius o estructura de l'impacte



Font: Elaboració pròpia, adaptat de Blasco (2009: 15)

Una formulació poc precisa dels objectius estratègics d'una intervenció fa perdre sentit i orientació pràctica a les accions del programa i obliga a reconstruir ex-post la seva definició per tal de fer-los avaluable. Per aquest motiu, és recomanable, ja en el mateix moment de dissenyar (ex-ante) el programa, formalitzar correctament els seus objectius de canvi. Una manera d'afrontar aquesta formalització és mitjançant la **caracterització SMART** dels objectius, això és, objectius específics, mesurables, assolibles, rellevants i temps-específics (requadre 10).

REQUADRE 10: ELS OBJECTIUS S.M.A.R.T.

- **(E)Specífics:** concrets i precisos, de manera que no hi hagi marge per a dobles interpretacions.
- **Mesurables:** quantificables, de manera que se'n pugui acabar verificant el grau d'acompliment.
- **Assolibles:** abastables, de manera que puguin servir d'incentius a l'acció.
- **Rellevants:** des del punt de vista del programa i l'organització, de manera que es generi consens i voluntat de canvi.
- **Temps-específics:** emmarcats en el temps, de manera que orientin la planificació d'execució i rendiment de comptes, i ajustin expectatives.

Exemples:

- "Reduir el nombre de víctimes de violència domèstica en, almenys, un 20% en els propers tres anys".
- "Reduir el consum habitual d'alcohol entre els menors de 18 anys en més d'un 40% durant els propers tres anys".
- "Incrementar en un 15% el nombre de contractacions indefinides entre les persones amb discapacitats psíquiques durant els propers cinc anys".
- "Incrementar en 20 punts el nombre de joves amb trastorns mentals que aconsegueixen finalitzar els estudis secundaris postobligatoris durant els propers dos anys".

Font: Elaboració pròpia, adaptat de Blasco (2009: 15)

- **Objectius operatius** Cada objectiu estratègic ha de tenir assignats uns objectius operatius, això és, fites específiques i mesurables adreçades a aconseguir els objectius estratègics i cal seleccionar els que hi contribueixin més directament. Es poden fixar objectius operatius tant pel que fa als recursos, a les activitats, com als productes. De forma general, doncs, entenem que els objectius operatius fan referència a **fites pròpies de la implementació**. Per exemple: aconseguir 20 nous voluntaris involucrats en la provisió del servei, introduir un servei de mentoria en un programa d'inserció laboral, incrementar en un 10% el nombre d'usuaris que completen un itinerari formatiu, serien objectius operatius.

Com en el cas dels objectius estratègics, de cara a poder monitoritzar el grau d'ajust i acompliment dels objectius operatius és necessari que aquests parteixin d'una definició clara i precisa, d'uns indicadors.

Definir mesures i indicadors: allò que es vol mesurar

Els indicadors de seguiment dels programes són instruments que permeten mesurar (quantificar) dades relatives als seus diferents components. Genèricament s'entén que un **indicador hauria de tenir els següents atributs:**

- **Vàlids** fan referència a les dimensions que es volen mesurar i no a d'altres factors.
- **Mesurables i factibles:** observables a partir de la realitat i de les dades disponibles. Aquesta característica obliga a tenir present els sistemes i fonts d'informació amb què es podrà comptar.
- **Sensibles:** susceptibles de capturar diferències i variacions reals. És important definir categories i/o unitats de mesura que ofereixin variabilitat associada a les accions del programa.
- **Comprensibles i comunicables:** la definició de l'indicador ha de ser clara i detallada, fàcil d'entendre i d'interpretar.
- **Contrastats:** preferentment, han de disposar d'algun grau de validació metodològica en el marc de processos de seguiment i avaluació com els que ens plantejem.
- **Rellevants:** idealment, haurien de referir-se a dimensions fonamentals dels diferents components del programa. Prioritzar el disseny d'uns indicadors i no uns altres, ajuda a prioritzar la planificació d'objectius, al temps que simplifica el sistema d'informació que n'ha de permetre el seguiment i avaluació.

Una manera simple de diferenciar els indicadors de seguiment és en funció del seu objecte d'atenció. A la pràctica, això vol dir distingir-los en base als objectius estratègics (**indicadors d'outcome**) o operatius (**indicadors d'implementació**) als quals s'associarien.

- **Indicadors d'outcome** (impactes). El mateix exercici consistent a caracteritzar els objectius

estratègics mitjançant el model SMART obliga a pensar en termes d'aquells indicadors o mesures que caldrà associar a cadascun d'ells. Al capdavant, un objectiu mesurable (característica SMART) és un objectiu traduïble a indicadors.

Reprement els exemples exposats en el requadre 1, podríem mesurar (en el temps i en l'espai indicats): el nombre de denúncies efectuades per violència domèstica, el nombre de menors de 18 anys que declaren consumir alcohol habitualment, el percentatge de contractacions indefinides entre les persones amb discapacitat psíquica, el nombre de joves que finalitzen els estudis secundaris postobligatoris.

- **Indicadors d'implementació** (operatius). Situàriem aquí aquells indicadors que s'adrecen a mesurar qüestions relatives als recursos i als productes del programa⁶, com per exemple:
 - Indicadors de *recursos inputs*: nombre de professionals implicats en el programa segons tipologia; nombre d'hores destinades pels professionals implicats; cost monetari de les accions, etc.
 - Indicadors de *producte* (output): nombre d'intervencions realitzades segons tipologia; nombre d'usuaris/beneficiaris que accedeixen i/o finalitzen el programa segons perfil sociodemogràfic; cobertura segons població diana; nivell de satisfacció dels agents implicats, etc.

Obtenció i gestió de les dades: fonts i sistemes d'informació

Paral·lelament al disseny d'objectius i indicadors, cal pensar en les fonts i instruments que permetran proveir el programa de la informació requerida per al seu seguiment i avaluació. Així mateix caldrà també definir el sistema que permetrà la gestió de tot aquest flux d'informació.

Pel que fa a les **fonts i instruments d'informació**, aquests poden ser d'índole diversa:

- a) **Registres administratius** S'hi recull informació no directament vinculada al programa i a la seva avaluació, i poden trobar-se en bases de dades diferents:
 - *Bases de dades censals*: per exemple, registres de la seguretat social, d'escolarització, de serveis socials, padró municipal, etc.
 - *Bases de dades mostrals* (enquestes d'estadística oficial): per exemple, enquestes de

⁶ Els indicadors d'implementació referits al component d'activitats queden subsumits en les mesures de recursos i productes.

l'INE (Enquesta de Població Activa, Enquesta de Condicions de Vida o Enquesta de Pressupostos Familiars), l'Idescat (Enquesta d'usos del temps, i d'usos lingüístics, o Enquesta de consum i pràctiques culturals) o d'altres administracions (Enquesta de Salut o Enquesta de Serveis, de l'Ajuntament de Barcelona).

- *Bases de dades construïdes per les pròpies organitzacions* i entitats als efectes de monitoritzar les seves activitats i els seus usuaris.

El principal avantatge de l'ús de registres administratius és que pot proveir informació bàsica (laboral, educativa, dependència, etc.) sobre un gran nombre d'individus al llarg d'un ampli període de temps. El fet que es tracti de dades existents "al marge" del programa, agilitza i abarateix el procés de recollida d'informació i, per tant, qualsevol exercici d'avaluació que se'n vulgui fer. Com a contrapartida, les dades administratives no sempre s'adapten als requeriments d'informació més específics dels programes.

b) Dades ad-hoc per a l'avaluació. Parlem, en aquest cas, de dades creades expressament per al seguiment o avaluació del programa. Aquestes poden ser:

- Dades quantitatives: obtingudes i processades a partir, bàsicament, d'enquestes.
- Dades qualitatives: generades a partir d'entrevistes, grups de discussió, observació o anàlisi documental.

A diferència dels registres administratius, les dades ad-hoc tenen l'avantatge que permeten recollir aquella informació que resulti més rellevant per al programa. Per contra, la generació ex-novo de dades (quantitatives o qualitatives) pot tenir un cost econòmic i de temps força elevat.

Pel que fa als **sistemes de gestió de la informació**, s'espera que aquests s'ajustin a les necessitats de seguiment dels programes i dels seus usuaris. El capítol sobre avaluació de la implementació aprofundeix en aquesta qüestió.

Des del punt de vista de l'avaluabilitat dels programes (això és, de cara a facilitar que aquests siguin avaluable en termes d'implementació i d'impacte), és condició necessària disposar d'un sistema de registres susceptible de permetre completar i integrar el conjunt d'informacions relatives a la participació en el programa en base a l'usuari o beneficiari individual com a nexa central. D'aquesta manera es podran combinar i analitzar amb els nivells d'agregació/desagregació aquelles dades que es considerin d'interès per a l'avaluació (característiques sociodemogràfiques, diagnòstics individuals, activitats realitzades, informes de seguiment, indicadors de resultat, etc.).

Finalment, de cara a fer possible una avaluació retrospectiva de l'impacte, caldrà que el disseny del programa compti amb un marge per a la **identificació de grups de control vàlids**. Existeixen diversos mètodes per avaluar l'impacte dels programes; la possibilitat i conveniència d'utilitzar un o altre dependrà de com el programa hagi definit els canals i requisits d'accés a la participació. Aquesta és una qüestió complexa, a la qual dediquem plena atenció en un capítol posterior d'aquesta guia (avaluació d'impacte).

Recapitulant: les preguntes sobre l'avaluabilitat dels programes

Resumint, pel que fa a les condicions d'avaluabilitat dels programes, l'avaluació del disseny es formula tres interrogants bàsics:

- Quins és el grau de precisió, mesurabilitat i rellevància dels objectius que defineix el programa, tant dels objectius estratègics com dels operatius?
- Els indicadors d'impacte i d'implementació dels programes, compleixen amb els requeriments de validesa, mesurabilitat, homologació i rellevància propis de tot indicador?
- En quina mesura les fonts d'informació i els sistemes de gestió i monitorització previstos permeten proveir el programa d'unes dades i un coneixement útil i rellevant, i alhora faciliten que aquest pugui ser avaluat en termes d'implementació i d'impacte?

Val la pena aclarir que l'avaluació del disseny no esgota la valoració global de l'avaluabilitat que es pot fer d'un programa. En efecte, analitzar l'avaluabilitat d'una intervenció passa també per ocupar-se d'altres focus d'interès com la capacitat d'incidència que pot acabar tenint l'avaluació a l'hora de provocar canvis reals en la orientació i continguts del programa. Per a una noció àmplia del que pot implicar valorar, de forma general, l'avaluabilitat d'un programa o servei, vegeu el **Protocol d'avaluabilitat de polítiques públiques** (Farré, 2016).

3.4. MÈTODES PER AVALUAR EL DISSENY

Com hem expressat a l'inici d'aquest capítol, l'avaluació del disseny no és un exercici "empíric", en el sentit que no té com a objectiu analitzar l'execució pràctica del programa, ni els seus impactes sobre la realitat. El seu propòsit és valorar els fonaments teòrics del programa, així com l'instrumental que el seu disseny preveu de cara a afavorir-ne la posterior avaluació.

Per tal d'acomplir amb aquest propòsit, en especial quan és en qüestió la teoria del canvi del programa, és de vital importància recórrer a l'evidència científica acumulada sobre el funcionament d'altres programes amb dissenys semblants als que ens puguem estar plantejant. A aquest enfocament dediquem el següent subapartat.

3.4.1. Fonamentar la teoria del canvi: l'aproximació "què funciona"

En efecte, una acció clau a l'hora de fonamentar la teoria del canvi d'un programa, això és, cadascuna de les seves hipòtesis (d'implementació i d'impacte), passa per aprofitar el coneixement existent sobre com funcionen o han funcionat altres intervencions amb plantejaments i objectius semblants. L'objectiu últim d'aquesta estratègia (en direm aproximació "què funciona") és aportar evidència empírica a la presa de decisions sobre el disseny dels diferents components dels programes. Per exemple, els dissenyadors d'un programa de prevenció del consum de drogues entre la població adolescent poden plantejar-se diferents estratègies: oferir uns tallers grupals i/o unes sessions de mentoria;

treballar dins els instituts o fer-ho en espais d'oci o lleure; comptar amb professionals dels serveis socials o juvenils i/o amb joves adults voluntaris amb una determinada experiència; concentrar les activitats en determinats perfils d'usuaris (segons edat, sexe, nivell de risc...) o obrir-les a un target ampli d'adolescents, etc.

Són múltiples i diversos els factors a tenir en compte a l'hora de decidir quines opcions prioritzar; factors relacionats amb la disponibilitat de recursos, però també de viabilitat tècnica (quina és la xarxa d'atenció juvenil en el territori, quina és la seva expertesa, quin perfil tenen els seus professionals, quina col·laboració esperem tenir dels instituts i els centres de lleure, etc.) i, fins i tot, política (quins compromisos polítics existeixen amb relació a aquest programa, quant a la seva prioritització i formes d'execució). La tesi que defensem en aquesta guia és que, dins d'aquests factors, hauria també de comptar-se **l'evidència sobre el "què funciona"**, com a element que fonamenta la teoria del canvi del programa i, per tant, pot incrementar-ne la probabilitat d'èxit. Èxit en la seva implementació i en el seu impacte.

Reprent l'exemple anterior, podríem llavors preguntar-nos:

- **En termes d'impacte:** existeixen avaluacions, a casa nostra o en altres països, que ens informin sobre el funcionament de diferents programes de prevenció del consum de drogues entre adolescents? Què ens diuen aquests estudis sobre l'efectivitat d'accions basades en tallers grupals? I en mentories individuals? Alguna d'aquestes estratègia ha mostrat tenir més èxit? Els beneficis d'aquests programes, depenen del perfil dels tutors que els implementen, o de l'espai on es desenvolupen? S'ha observat si hi ha grups d'adolescents més sensibles que d'altres als efectes del programa?
- **En termes d'implementació:** quin mecanisme d'accés al programa afavoreix una major adequació entre els usuaris reals i la població destinatària de la intervenció? És més fàcil acomplir amb els continguts de les activitats quan aquestes les duen a terme professionals dels serveis juvenils? En quina mesura la coordinació del programa es veu afectada quan es treballa amb els instituts o amb personal voluntari? De cara a afavorir que els participants completin el programa, és millor concentrar-lo en poques setmanes o espaiar-lo al llarg d'un període relativament llarg de temps? Quina incidència té cadascuna d'aquestes alternatives en el pla organitzatiu?

Cal tenir clar, però, que la qualitat i el grau d'utilitat de l'"evidència" que puguem obtenir dependrà de com s'estableixin els procediments emprats per cercar-la i revisar-la. Són diversos els mètodes que es poden utilitzar quan del que es tracta és d'elaborar un cert balanç del coneixement acumulat al voltant d'un determinat tema o problemàtica. I cal tenir en compte que el seu nivell d'exhaustivitat i sistematicitat pot ser força dispar; en conseqüència, igualment dispar serà la validesa del coneixement que se n'acabi obtenint.

La taula 4 sintetitza les característiques dels principals mètodes reconeguts de recol·lecció i revisió d'evidències: del mètode menys sistemàtic i menys protocol·litzat (revisió "ad-hoc" de la literatura), al mètode més exhaustiu (revisió sistemàtica múltiple).

Taula 4. Procediments habituals de recollida i revisió d'evidències

PROCEDIMENT DE RECOLLIDA/ REVISIÓ	NATURALES	DEFINICIÓ	LIMITACIONS	DURADA
<u>Revisió de la literatura “ad-hoc” (literature review)</u>	No sistemàtica	<ul style="list-style-type: none"> Recull dels estudis més rellevants i/o il·lustratius No expliciten el procés de cerca bibliogràfica. 	<ul style="list-style-type: none"> Existeix un risc alt de biaix de selecció: estudis disponibles o que tenen resultats positius Adequat en contextos de recursos limitats. 	1 setmana a 2 mesos
<u>Revisió de curt abast (quick scoping review)</u>	No sistemàtica	<ul style="list-style-type: none"> Elabora un ‘mapa’ de la literatura disponible sobre un tema 	<ul style="list-style-type: none"> Respon a preguntes molt descriptives Fa recerca en poques bases de dades Cerca poques paraules clau 	1 setmana a 2 mesos
<u>Revisió de revisions (review of reviews)</u>	Sistemàtica	<ul style="list-style-type: none"> Elabora síntesis de les revisions realitzades, més que no pas de materials primaris 	<ul style="list-style-type: none"> Es pot utilitzar només quan hi ha prou evidència acumulada i revisada La qualitat de les revisions seleccionades pot ser diversa 	2 a 4 mesos
<u>Revisió ràpida de l'evidència (rapid evidence assessment)</u>	Sistemàtica	<ul style="list-style-type: none"> Recerca restringida sobre l'efectivitat d'una política Compta amb un protocol de cerca Es pot traçar i replicar 	<ul style="list-style-type: none"> Cal delimitar una pregunta sobre efectivitat molt concreta La finestra temporal és limitada i recent 	2 a 6 mesos
<u>Revisió sistemàtica completa (full systematic review)</u>	Sistemàtica	<ul style="list-style-type: none"> Revisió àmplia de la literatura sobre l'efectivitat d'una política Compta amb un protocol sistemàtic Cerca bases de dades acadèmiques com literatura gris 	<ul style="list-style-type: none"> El procés de recollida documental pot ser força artesanal Requereix força temps Pot requerir de recursos humans addicionals – interns o externs 	8 a 12 mesos
<u>Revisió sistemàtica múltiple (multi-arm systematic review)</u>	Sistemàtica	<ul style="list-style-type: none"> Revisió àmplia de la literatura atenent a diferents preguntes d'efectivitat d'una política Ídem “revisió sistemàtica completa” 	<ul style="list-style-type: none"> Incrementa les limitacions de la “revisió sistemàtica completa” 	Més de 12 mesos

Font: Elaboració pròpia a partir de goo.gl/JKHf4M

La realització d'exercicis de cerca i revisió d'evidències que tinguin un mínim de rigor planteja tot un seguit de reptes per a qualsevol institució, també per al tercer sector social. Cal disposar de temps i expertesa tècnica (més o menys temps i expertesa segons el procediment de cerca que es vulgui emprar). No és infreqüent que des del sector es recorri a recercadors o grups d'investigació de l'àmbit universitari per dur a terme revisions mínimament sistemàtiques de la literatura.

I cal també tenir present que l'evidència empírica més rigorosa prové sovint d'avaluacions de programes implementats en altres països (països amb més tradició avaluativa i/o entorns on es testen diferents programes experimentals). Per tant, els resultats d'aquestes avaluacions no són directament traslladables al context d'aplicació dels "nostres" programes.

Sigui com sigui, convé remarcar que, en els darrers anys, hi ha hagut un creixement molt ràpid de la recerca aplicada en l'àmbit de les polítiques i els programes socials i de la seva difusió. Augmenta el nombre d'articles publicats en revistes especialitzades, però també apareixen noves revistes per a nous àmbits d'intervenció. Paral·lelament, cal tenir en compte que existeix molta literatura grisa (notes breus, informes, 'policy briefs'...) sobre determinades intervencions socials. El resultat de tot plegat és que, per a algunes problemàtiques socials, s'ha generat prou coneixement científic per poder respondre en millors condicions a preguntes clau sobre la naturalesa del problema, què funciona per resoldre'l, el per què i el com. En paral·lel, acompanyant aquest coneixement, s'ha incrementat també el nombre de recursos (bancs de dades, repositoris, plataformes on-line, etc.) que faciliten la seva difusió.

Per a més informació sobre les diferents eines de revisió de l'evidència i sobre els recursos de síntesi i difusió existents, podeu consultar la Guia pràctica d'avalua **El paper de les revisions sistemàtiques en el disseny i elaboració de polítiques públiques** (Sanz, 2017).

4.2. Altres focus d'interès: la visió dels actors del programa

L'avaluador o el grup de treball que tingui la missió d'avaluar el disseny d'un programa social necessita recollir, de forma relativament sistemàtica, la informació que li permeti anar més enllà de la seva impressió inicial o coneixement parcial del programa o política, i integrar les versions sobre el programa dels diferents actors que prenen part en el seu disseny i implementació. Els principals mètodes de recollida d'informació rellevant per a aquest fi es descriuen en el requadre 11.

REQUADRE 11: TÈCNiques DE RECOLLIDA D'INFORMACIÓ PER A L'AVALUACIÓ DEL DISSENY

- **Entrevistes** a les persones que tinguin una relació més directa amb el disseny i implementació del programa a avaluar, per recollir les seves percepcions sobre el problema que la intervenció vol adreçar i la manera com se suposa que ho ha de fer.
- **Revisió de fonts documentals** que poden contenir la versió oficial de la teoria del canvi o bé algunes assumpcions o expectatives que ajudin a construir-la, com ara la normativa que regula el programa, comunicats de premsa, memòries de programes pressupostaris que incloguin la intervenció a avaluar, plans estratègics o operatius, presentacions públiques, i altres informes i documents interns.
- **Observació directa** del programa, en què l'avaluador presencia directament processos i activitats (com ara reunions de coordinació, l'atenció als beneficiaris, el procés de revisió de projectes a ser subvencionats, etc.) i en registra la informació rellevant.
- **Revisió de la literatura acadèmica** (Veure punt 4.2.)

Font: Elaboració pròpia

En qualsevol cas, el procés d'anàlisi de la teoria del canvi i l'avaluabilitat del programa **no hauria de ser una feina solitària**. Si la intervenció és prou complexa, mal definida o no hi ha consens sobre el disseny, es pot crear un petit grup de treball amb persones que tinguin algun rol rellevant en relació amb la implementació del programa o els seus resultats (gestors, operadors del programa, beneficiaris, etc.) per fer l'anàlisi inicial i elaborar un primer esborrany.

En els casos en què no es consideri necessari crear aquest grup, és igualment important concebre el desenvolupament del model lògic i operatiu de la intervenció com un **procés iteratiu**, en què la informació facilitada pels actors rellevants s'integra en un esborrany, que es contrasta de nou amb els actors i es va modificant fins arribar a desenvolupar una visió compartida sobre com se suposa que el programa ha de funcionar, quins objectius ha d'assolir i com els ha d'avaluar. Aquestes iteracions han de permetre comprovar que el disseny realment representa els fonaments lògics del programa, i que haurà de permetre conèixer i documentar aquells aspectes de la intervenció considerats més rellevants.

3.5. RECURSOS

La taula 5 esmenta els principals repositoris d'evidència i avaluacions d'impacte existents a nivell internacional. Com es pot comprovar, la majoria d'ells tenen base als Estats Units i sintetitzen evidències d'avaluacions de programes principalment anglosaxons.

Taula 5. Bases de dades d'evidències i impactes, segons àmbits d'interès

EDUCACIÓ	
Best Evidence Encyclopedia	http://goo.gl/qmxuO
What Works Clearing House	http://goo.gl/qdj6Q
Education Endowment Foundation	https://goo.gl/0KGCQQ
'Què funciona en educació?'	http://goo.gl/DjVEk0
INFANTS I ADOLESCENTS	
Promising Practices Network	http://goo.gl/QVzYr
Blueprints for healthy youth development	http://goo.gl/HHsBo7
Early Intervention Foundation	http://goo.gl/jmCd1b
POLÍTICA SOCIAL	
Coalition for evidence	http://goo.gl/MgHqw
MDRC	http://goo.gl/1VYGhs
Mathematika	http://goo.gl/ef46c
Campbell	http://goo.gl/bP0no3
Social Care Online	http://goo.gl/GY9EoT
Social services Knowledge Scotland	http://goo.gl/0yZOh
Self-sufficiency research clearinghouse	https://goo.gl/2Cw5W
The EPPI-Centre	http://goo.gl/Pki7s
JOVES	
Youth.gov	http://goo.gl/Sr0j9r
MERCAT LABORAL	
Workforce System strategies	http://goo.gl/GcTHdg
IZA World of Labour	http://goo.gl/c9Cr4L
JOVES I OCUPACIÓ	
Youth employment inventory	http://goo.gl/2rRGS
POLÍTIQUES FAMILIARS	
Home Visiting Evidence of Effectiveness	http://goo.gl/32c51A

SALUT	
Substance Abuse and Mental Health Services Administration	http://goo.gl/WV4SXS
The Cochrane	http://goo.gl/2Adwel
National Care Institute for Excellence	http://goo.gl/yDqkmY
Health Systems evidence	https://goo.gl/i6VeCG
JUSTÍCIA	
Office of justice justice and delinquency prevention	http://goo.gl/AATf1L
Crime solutions.gov	http://goo.gl/tLxg5
COOPERACIÓ I DESENVOLUPAMENT	
International Initiative for Impact evaluation	http://goo.gl/pmozs

Font: Sanz (2017: 41)

3.6. REFERÈNCIES

Blasco, J. *Avaluació del disseny*. Barcelona: Ivàlua, 2009. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 3).

Farré, M. *Protocol d'avaluabilitat de polítiques públiques*. Barcelona: Ivàlua, 2016 (Mimeo)

McLaughlin, J.A.; Jordan, G.B. *Using Logic Models*. A: Wholey, J.S.; Hatry, H.P.; Newcomer, K.E. *Handbook of practical program evaluation*. 2nd ed. San Francisco: Jossey-Bass, 2004.

Rossi, P.H.; Lipsey, M.W.; Freeman, H.E. *Evaluation: a systematic approach*. 7th ed. London: SAGE, 2004.

Ruegg, R.; Feller, I.A. *A Toolkit for evaluating public R&D investment: models, methods, and findings from ATP's first decade*. (NIST GCR 03-857). Washington, DC: National Institute of Science and Technology, 2003.

Sanz, J. *El paper de les revisions sistemàtiques en el disseny i elaboració de polítiques públiques*. Barcelona: Ivàlua, 2017 (Mimeo)

4. AVALUACIÓ DE LA IMPLEMENTACIÓ

4.1. PUNT DE PARTIDA I MOTIVACIÓ

D'entre els tipus d'avaluació que presentem en aquesta guia, l'avaluació de la implementació és probablement la més present dins del tercer sector. Efectivament, la gran majoria d'entitats reporten anualment unes memòries econòmiques i d'accions que recullen dades relacionades amb els recursos, activitats i productes dels seus programes, i és habitual que comptin amb metodologies i eines tals com mapes de processos interns, sistemes de gestió de la qualitat o quadres de comandament integral. Aquests exercicis poden encabir-se dins d'allò que aquí entendrem com a avaluació de la implementació o de procés.

Tanmateix, cal tenir en compte que el nombre d'indicadors que es registren, sovint és molt superior als que després s'analitzaran o seran rellevants, i no és estrany que la seva recollida es percebi com una càrrega de feina que respon a una voluntat externa i que està desvinculada de la missió de l'entitat. Així mateix, el tractament que es fa de la informació sol ser descriptiu, i la seva transformació en anàlisis rellevants sobre el funcionament del programa i sobre com podria millorar continua suposant un repte important.

Tenint en compte aquestes consideracions, en els següents apartats d'aquesta guia plantegem com abordar l'avaluació de la implementació dins del tercer sector.

4.2. QUÈ VOL DIR AVALUAR LA IMPLEMENTACIÓ D'UN PROGRAMA

Un cop hem identificat unes necessitats socials i hem dissenyat una intervenció o programa per donar-hi resposta, cal implementar-lo, és a dir, dur-lo a la pràctica. Es tracta d'una fase en què sovint s'acaben produint desviacions significatives del disseny originalment previst, les quals poden ser fruit de factors diversos (canvis en la realitat sobre la qual s'actua, canvis organitzatius en l'entitat o en la disponibilitat de recursos, etc.).

Així doncs, l'avaluació de la implementació valora la consistència i coherència de l'execució (pràctica) d'un programa en relació amb el disseny inicial (teoria) (Lázaro i Obregón, 2009), permetent respondre preguntes sobre la relació entre els diferents elements de la implementació: els recursos emprats, les activitats desenvolupades i els productes resultants.

Quan avaluem aquestes relacions ho podem considerar diferents perspectives, a continuació esposem les més rellevants (Humphrey, 2016):

1. **Adherència.** Grau en el que la implementació s'adhereix a allò planificat.
2. **Dosificació.** Quantitat de la intervenció que s'ha implementat i l'efectivament rebuda.
3. **Abast.** Grau en el que s'aconsegueix arribar a la població diana del programa.
4. **Qualitat.** Mesura en que les activitats compleixen uns requisits de qualitat.
5. **Receptivitat dels participants.** Grau en que els participants s'impliquen en el programa.
6. **Adaptació.** Canvis que s'hagin fet respecte al disseny inicial del programa.

L'avaluació de la implementació pot tenir lloc en diferents moments de la vida d'un programa. En els seus estadis inicials, pot ajudar a entendre millor el seu funcionament, identificar i anticipar elements clau en la seva gestió, suggerir mesures correctives i informar decisions i canvis en el seu disseny. Pel que fa a programes més madurs, ja consolidats, l'avaluació de la implementació, combinada amb l'avaluació de l'impacte, pot contribuir a una millor comprensió dels efectes i dels mecanismes causals en joc. Més encara, quan es demostra que un programa té impactes positius, l'avaluació pot servir per documentar i sistematitzar la seva implementació, assegurant així la seva replicabilitat i escalabilitat consistent.

Els motius per a dur a terme una avaluació de la implementació es poden dividir entre **sumatius** o **conclusius**, que són aquells que busquen fer una valoració global del programa i la seva utilitat, i **formatius**, aquells que busquen obtenir coneixements per introduir les modificacions que s'estimin oportunes (Ballart, 1992). En resum, els principals propòsits de l'avaluació de la implementació serien:

- Contribuir a la millora del disseny del programa.
- Validar el model d'intervenció del programa.
- Obtenir *feedback* sobre si el programa s'està duent a terme segons allò previst.

En els següents apartats descrivim els passos per a dur a terme una avaluació de la implementació, començant per la definició de les preguntes d'avaluació, i seguint pels mètodes i enfocaments, l'anàlisi de la informació i la comunicació i millora del programa.

4.3. PREGUNTES I DIMENSIONS DE L'AVALUACIÓ DE LA IMPLEMENTACIÓ

El punt de partida per a la definició de les preguntes d'implementació és el disseny del programa, és a dir, la identificació dels seus components i dimensions clau i el conjunt d'hipòtesis que expliquen com s'espera que aquests elements es relacionin entre sí: com uns determinats recursos han de permetre dur a terme unes activitats, generar uns productes concrets, i aconseguir uns impactes desitjats. Així doncs, en base al disseny original, podem definir una sèrie de preguntes classificables segons a quina dimensió de la implementació facin referència:

Taula 6. Preguntes rellevants segons els components als que fan referència

DIMENSÍO	PREGUNTA D'AVALUACIÓ
Adherència	<ul style="list-style-type: none"> • S'ha disposat dels recursos planificats i aquests han estat suficients? • S'han seguit els procediments i protocols establerts? (p.e. metodologia de la formació) • Els processos clau s'han executat tal i com es preveïa? • Els professionals tenen les capacitats i coneixements necessaris en el seu àmbit d'intervenció? • Les activitats s'han dut a terme al ritme planificat? • L'assignació de tasques i responsabilitats i la coordinació interna han estat les previstes? • S'han dut a terme les accions de coordinació externa previstes?
Dosificació	<ul style="list-style-type: none"> • S'han proveït tots els serveis previstos en la quantitat prevista? • Quin ha estat el nivell d'ús per part dels usuaris? (p.e. assistència)
Abast	<ul style="list-style-type: none"> • El perfil dels usuaris es correspon amb el de la població diana? • Quantes persones han fet ús dels serveis? • Quina proporció representen sobre la població diana?
Qualitat	<ul style="list-style-type: none"> • S'han produït llistes d'espera? De quina durada? • Quin és el nivell de satisfacció dels beneficiaris?
Receptivitat	<ul style="list-style-type: none"> • Els participants desenvolupen les conductes esperades després de rebre el servei?
Adaptació	<ul style="list-style-type: none"> • Quins canvis s'han realitzat al llarg de la implementació respecte al disseny original?

Font: Elaboració pròpia a partir de Humphrey (2016) i Rossi et al. (2004).

Amb l'objectiu de maximitzar la utilitat de l'avaluació, pot ajudar vincular cadascun d'aquests interrogants als àmbits de decisió i millora que més directament apel·len. Més concretament, per tal d'assegurar que els resultats de l'avaluació d'implementació s'incorporen en la millora del programa, es recomana identificar diferents cursos d'acció segons els resultats de l'avaluació, anticipant els passos a seguir en cadascun d'ells, tenint en compte els recursos necessaris, els actors clau que hauran de fer possible la gestió del canvi i estimant l'horitzó temporal en el que aquest canvi seria desitjable.

Figura 8. Exemples de decisions a diferents nivells de l'estructura organitzativa



Font: Elaboració pròpia

REQUADRE 12: COM TREBALLAR PER INCREMENTAR LA PERMANÈNCIA ALS PROGRAMES

Una entitat proveeix un servei d'assessorament i formació laboral per a joves a l'atur. El programa s'inicia amb una sessió informativa general i una entrevista personal amb els diferents joves. Aquells que es considera que no compleixen les característiques de la població diana són derivats a altres serveis, mentre que a la resta se'ls ofereixen 12 sessions de formació destinades a millorar la seva ocupabilitat. El feedback informal que els professionals reben per part dels participants que finalitzen el curs és positiu, tanmateix, un volum important de participants abandonen després de les primeres sessions. L'entitat decideix introduir canvis en la metodologia per fer la formació més amena i orientada a la pràctica, i decideix avaluar com s'implementen els canvis i l'evolució dels nivells d'assistència i satisfacció per part dels nous participants. Les persones responsables i els docents, amb el coneixement i suport del finançador extern, es comprometen, en base als resultats de l'avaluació, a prendre noves decisions sobre els continguts i la metodologia de la formació de cara a la següent edició del curs.

4.4. MÈTODES PER AVALUAR LA IMPLEMENTACIÓ D'UN PROGRAMA

Un cop definides les preguntes d'implementació rellevants, cal identificar o construir aquells indicadors i fonts d'informació que ens permetran donar-hi resposta, així com establir uns valors de referència del que considerariem una implementació satisfactòria.

Els **indicadors** seleccionats han de ser aproximacions al concepte que volem mesurar, i els hauríem de tenir definits ja en el disseny del programa (veure capítol anterior), per a

cada element de la teoria del programa. Pel que fa als **valors de referència**, es recomana fer servir els objectius concrets fixats en el disseny i planificació del programa, tanmateix, si no s'han fixat, es poden fer servir altres punts de comparació com estàndards definits en les convocatòries i licitacions, valors de referència en la literatura acadèmica, resultats d'avaluacions similars o valors de referència provinents del benchmarking amb altres entitats. En els casos en què no s'aconsegueixi trobar un valor de referència, al final de l'avaluació s'haurà de valorar si els resultats observats es consideren acceptables.

4.4.1 Mètodes d'obtenció d'informació

Un cop sabem què volem mesurar i hem establert uns valors de referència que permetin valorar els resultats, cal assegurar l'accés a la informació. Idealment les dades ja existiran (és el cas dels registres administratius), però si no és així, caldrà recollir-les. A continuació repassem els principals mètodes disponibles per a l'obtenció d'informació:

- **Revisió de documents.** Pot incloure documents com les bases de la convocatòria d'un servei, memòries d'activitats, actes de reunions, protocols d'actuació, mapes de processos, anàlisi funcional realitzada prèviament al desenvolupament de nous sistemes d'informació, documents de descripció de llocs de treball...
- **Registres del programa.** Poden ser una font de feedback que permeti fer un seguiment de la implementació de les intervencions. És clau que els indicadors d'interès estiguin ben definits i registrats. Cal comptar amb sistemes d'informació capaços d'agregar dades i facilitar-ne l'anàlisi de forma àgil i poc costosa. Així, és recomanable tendir cap a eines ben integrades en la pràctica professional, que recullin només allò que és indispensable i que pot ser rellevant des del punt de vista del seguiment i avaluació.
- **Informes derivats de sistemes de seguiment.** Com a resultat de l'explotació de les bases de dades d'un programa, sovint s'elaboren de forma periòdica informes breus, que inclouen valors actualitzats, generalment d'activitat, d'una bateria d'indicadors preestablerts. Un exemple seria les memòries anuals d'un programa o els informes de justificació d'un projecte adreçats al finançador.
- **Enquestes.** Permeten obtenir informació sobre els beneficiaris, professionals o altres actors rellevants en el marc del programa. Es poden fer a través de diferents canals, ja sigui en línia, en persona o per telèfon. Cal tenir en compte que cada modalitat (online, telefònica i presencial) presenta costos i taxes de resposta diferents. Poden ser molt útils per mesurar les expectatives i satisfacció dels participants en relació al programa o a diferents components d'aquest.
- **Entrevistes individuals i grups de discussió amb informants** clau. Són un recurs per obtenir informació qualitativa sobre percepcions i opinions en relació a diferents aspectes del programa per part de professionals, gestors, beneficiaris, voluntaris. Poden servir per identificar factors rellevants en la implementació d'un programa, ja siguin obstacles, problemes o oportunitats de millora etc. També pot servir per explorar possibles canvis i millores.

- **Observació directa.** Consisteix en que sigui l'avaluador qui registri les dades, presenciant directament les activitats del programa. Es pot fer amb eines com els *check lists* o bé prenent notes qualitatives.
- **Estudis de cas.** Es tracta de la recollida d'informació en profunditat concentrada en un grup d'usuaris concrets o una part determinada del programa. Combina diferents de les altres tècniques d'obtenció d'informació enumerades, quantitatives i qualitatives (registres, entrevistes individuals, grups de discussió, observació directa...). Serveix per entendre en profunditat la implementació en determinades circumstàncies.
- **Opinió d'experts.** Aquesta es pot obtenir aplicant altres tècniques, com les entrevistes o grups de discussió, amb persones expertes. Es fa servir quan es requereix un coneixement especialitzat, ja sigui del propi programa que s'avalua, de les necessitats socials i impactes esperats o del context.

4.4.2. Procediments de recollida de dades i sistemes de seguiment

De forma general, l'avaluació de la implementació es planteja com un exercici específic, de caràcter més o menys puntual, que pretén donar una resposta detallada a preguntes destacades sobre el procés d'execució dels programes. Tanmateix, bona part de la informació que es requereix per dur a terme aquest exercici es recull de forma continuada en el que s'anomenen sistemes de seguiment.

En el cas d'aquelles dades relatives a la implementació que l'entitat vulgui recollir i generar de forma rutinària i periòdica, caldrà garantir que els procediments per a la recollida d'aquesta informació estiguin ben **planificats i estandarditzats**. Serà necessari establir **qui** recollirà la informació i qui se'n veurà implicat o afectat, **quan** es recollirà, amb quin **instrument** i on quedarà **registrat**. Pel que fa als instruments, es recomana **testejar-los** abans de generalitzar-ne l'ús, és a dir, provar-los a petita escala, amb un grup de persones representatives dels qui han de ser inclosos en la recollida de dades, i assegurar que són comprensibles, fàcils d'usar i culturalment apropiats, en el sentit de que no xoquen amb les normes i valors de les persones implicades.

Pel que fa a la **frequència en la recollida de dades**, caldrà que aquesta s'ajusti als diferents moments i processos de la implementació dels programes, evitant al mateix temps esquemes de periodicitat que suposin una càrrega excessiva de recursos i de feina per als professionals (Durlak & DuPre, 2008).

Els instruments de mesura i seguiment de la implementació tendeixen a ser específics de cada programa, tanmateix hi ha alguns repositoris d'indicadors que poden resultar útils com a referència⁷.

⁷ Vegeu, per exemple, el sistema IRIS (<https://iris.thegiin.org/>).

Taula 7. Paràmetres per a l'estandardització de la recollida de dades

INTERVENCIÓ: PROGRAMA D'ASSESSORAMENT I FORMACIÓ LABORAL PER A JOVES A L'ATUR				
Preguntes d'avaluació: <ul style="list-style-type: none"> • La nova formació laboral s'implementa tal com s'ha dissenyat? • Quina valoració en fan els assistents? 				
TIPUS	VARIABLE	INSTRUMENT	MOMENT DE RECOLLIDA	PERSONA ENCARREGADA
Input / Recurs	Nombre de persones inscrites	Formulari	Abans de l'inici	Personal de gestió
Activitat	Nombre de classes realitzades	Registre	Durant	Professor/a
Activitat	Correspondència entre el contingut i metodologia de la classe i el disseny	Checklist	Durant dues classes aleatòries.	Avaluador
Producte	Nombre d'assistents	Registre	Durant	Professor/a
Producte	Valoració dels assistents	Qüestionari	Al final del curs	Professor/a

Font: Elaboració pròpia

Finalment, es recomana comptar amb **controls de qualitat** periòdics, com podria ser la revisió de mostres aleatòries de dades per comprovar si hi ha camps en blanc o valors que no es corresponen amb el que es demanava, assegurant que les respostes tinguin sentit. En cas de discrepàncies, cal observar si hi ha algun patró comú (dades introduïdes per la mateixa persona, en un moment determinat etc.) i parlar-ho amb les persones involucrades en la seva recollida o pensar en introduir en els elements mencionats.

REQUADRE 13: QUALITAT DE LES DADES: VALIDESA I FIABILITAT

L'objectiu tot procediment de recollida d'informació i sistema de seguiment és assegurar la qualitat de les dades, entesa en base a dos conceptes:

- **Validesa:** Precisió de la informació, expressa el grau en que un instrument mesura allò que es proposa mesurar.
- **Fiabilitat:** Consistència o grau en que les dades són reproduïbles. Si recollim un indicador en repetides ocasions hauríem d'obtenir el mateix valor.

Font: Elaboració pròpia a partir de Casado i Blasco (2009: 17).

4.4.3. Anàlisi de la informació

El seguiment de les dades operatives d'un programa sovint compta amb una mala reputació per part dels professionals, en part pel gran volum d'indicadors que es demana recollir i perquè es pot percebre més com una demanda externa, sovint de control, més que com

una tasca que contribueixi a la missió de l'entitat. L'objectiu, doncs, ha de ser donar-li un valor a les dades de les que es disposa, de forma que suposi un aprenentatge sobre què està passant i com es podria millorar el programa. Cal tenir en compte que la funció de l'avaluació no és únicament descriure, sinó donar compte de l'adequació o no de l'execució del programa en relació a allò esperat, així com detectar espais i oportunitats de millora. En aquest sentit, no pretenem una descripció detallada de totes les dimensions de la implementació, sinó una anàlisi orientada a la millora i presa de decisions.

Tenint en compte aquestes premisses, es tracta de comparar els indicadors que haguem decidit prioritzar i que haurem recollit durant la implementació del programa, amb els valors de referència que haguem fixat prèviament. Aquest exercici ens permetrà avaluar la correspondència entre allò que esperàvem del programa i allò que ha acabat sent.

Durant la valoració de les desviacions hi pot haver una tensió entre la **fidelitat** al disseny original i la **necessitat d'adaptació** al context i a les situacions externes que estan fora del control del programa (Mowbray et al., 2003). En aquest sentit, una desviació del disseny original pot ser positiva si ha permès donar resposta a canvis en les preferències dels usuaris o aprofitar oportunitats imprevistes. Tanmateix, quan el programa implementat difereix significativament d'allò que s'havia dissenyat, cal plantejar-se si estem parlant del mateix programa o no, i per tant, si l'avaluació té sentit tal com està plantejada o val la pena fer-la agafant com a referència un disseny diferent a l'original.

Anàlisi de dades quantitatives

L'anàlisi de les dades variarà en funció de les preguntes d'avaluació plantejades, les dades disponibles, així com de l'expertesa i habilitats de l'equip encarregat de l'avaluació. Podem dividir els tipus d'anàlisi que es poden realitzar en: descriptiu, comparatiu i relacional o causal:

- **Descriptiu.** Aquest tipus d'anàlisi consisteix en quantificar una o diverses variables de forma separada, com podria ser el total d'estudiants inscrits, el nombre de cursos oferts, el total d'assistents o la valoració mitjana feta pels assistents al final de cada curs. Aquest tipus d'anàlisi pot ser **puntual**, és a dir, la fotografia d'un determinat moment, o continuat, és a dir, l'**evolució temporal** d'un indicador. Quan disposem d'observacions de forma repetida en el temps, resulta fàcil representar les dades en forma de sèries temporals que ens permetin observar tendències. Aquest tipus d'anàlisi ens permet comparar els diferents valors obtinguts amb allò que havíem previst en el disseny del programa. Quan no haguem previst objectius clars i concrets, podem adoptar altres referències, ja siguin estàndards de qualitat generalment acceptats o valors provinents de la literatura acadèmica.
- **Comparatiu.** Mentre que els finançadors d'un programa poden estar més interessats en l'evolució general d'un programa, l'entitat pot estar interessada en poder comparar diferents formes de treballar o diferents contextos i situacions en què es dona la

intervenció. L'anàlisi comparativa consisteix en observar com es comporta una determinada variable entre diferents grups d'individus. Un exemple seria comparar els nivells de satisfacció dels assistents a un curs entre diferents grups d'edat, o analitzar si un determinat perfil de persones té més dificultat per completar la formació.

- **Causal.** Es tracta d'un tipus d'anàlisi generalment associat a l'avaluació de l'impacte, tractada en el següent capítol de la guia, però que també es pot aplicar a elements relacionats amb la implementació del programa. Es tracta, no només d'analitzar la relació entre diferents variables, sinó també de determinar si hi ha una relació de causalitat entre elles. En aquest sentit podríem comparar si un canvi en els horaris d'atenció influeix significativament en el nombre de persones ateses o si una nova metodologia formativa fa augmentar la satisfacció per part dels estudiants.

Anàlisi de dades qualitatives

L'anàlisi qualitativa consisteix a interpretar el significat de la informació sobre el programa obtinguda mitjançant mètodes com entrevistes, grups de discussió, observacions directes o revisions de documents, i a presentar de forma responsable el que aquestes dades revelen (Wholey, 2004). Aquest exercici es pot dividir en dues parts: (1) la reducció de les dades i la identificació de patrons, i (2) la producció i comunicació d'unes conclusions.

Durant el procés d'obtenció i anàlisi de dades, es poden produir diferents amenaces a l'objectivitat i la validesa, com el fet que la informació cobreixi només una part del programa analitzat (p.ex. si només entrevistem aquells usuaris que han completat el servei, mentre que no tenim informació dels que l'abandonen abans d'hora) o la reinterpretació inconscient d'aquestes per acomodar-les a creences prèvies. Per evitar aquests problemes es recomana:

- Preguntar-se si les dades són prou completes o pot haver-hi informació que ha quedat exclosa de l'avaluació.
- Assegurar que la informació rellevant provingui de diferents persones i no només d'una mostra filtrada d'informants.
- Combinar diferents tècniques de recollida de dades (com les anteriorment esmentades).
- Buscar activament diferents interpretacions que encaixin amb les dades de les que es disposa.

La incorporació de tècniques qualitatives ens ha de permetre desenvolupar una millor comprensió del funcionament del programa, que incorpori els coneixements i visió dels diferents actors d'alguna manera implicats o afectats pel programa, així com els factors contextuais que puguin estar jugant un paper important en la implementació.

La combinació de diferents mètodes: els mètodes mixtos

És habitual que les avaluacions de la implementació combinin diferents mètodes d'obtenció d'informació i tècniques d'anàlisi, la qual cosa permet incorporar un major nombre de perspectives, afavoreix una major profunditat en el coneixement del funcionament del programa i pot ajudar a detectar resultats no esperats. Tanmateix, això planteja la necessitat d'integrar bé els diferents tipus de dades i la conveniència d'una planificació prèvia.

Aquesta integració depèn en bona mesura de si els diferents mètodes emprats responen o no a la mateixa pregunta o cadascun respon a preguntes diferents, així com del tipus de complementarietat que es preveu es produeixi entre els resultats d'un i altres. A continuació exposem les formes més comunes de fer servir mètodes mixtes:

- **En paral·lel.** Els mètodes s'utilitzen i s'integren un cop s'han analitzat les dades, fent una síntesis o el que es coneix com a triangulació de la informació.
- **Seqüencialment.** Els diferents mètodes es fan servir un rere l'altre, i els resultats obtinguts per un influeixen en el disseny i ús del següent.
- **Combinacions multi-nivell.** Algunes avaluacions poden tenir diferents nivells d'anàlisi (p.e.: classe, curs, escola), per la qual cosa es podrien considerar apropiats diferents mètodes per a cada nivell. Aquest tipus de combinació de mètodes és compatible amb els dos anteriors.

Quan s'opta per la forma paral·lela, es pot donar la situació que els diferents resultats obtinguts no siguin consistents entre ells. Això obligaria a reexaminar les dades qualitatives, per veure si hi ha factors que s'han ignorat en algun moment del seu processament, i a analitzar les dades quantitatives de forma desagregada, per veure si hi ha algun patró que ens hagi passat inadvertit i que pugui explicar els resultats discrepants. Al mateix temps, és recomanable reunir als encarregats de la recollida de les dades per debatre i explorar possibles explicacions a la divergència de resultats, tals com possibles biaixos. Finalment, si es considera que hi ha algun tipus d'informació que podria ajudar a entendre millor els fets, es pot plantejar una nova fase de treball de camp que permeti complementar les dades disponibles.

4.5. RECURSOS

Manuais de referència

- Carreras, I. Iglesias, M. Sureda, M. Liderazgo orientado a resultados en las ONG. Estrategia, sistemas de medición y cuadros de mando. Barcelona: Instituto de Innovación Social de ESADE, 2011.
- Lázaro, B. Obregon, I. (2009). Avaluació de la implementació. Col. Guies pràctiques sobre avaluació de polítiques públiques N°4. Barcelona: Ivàlua, 2009.

Anàlisi de dades i estadística bàsica

En l'actualitat existeix una multiplicitat de recursos disponibles en forma gratuïta a la web, a partir dels quals aprofundir sobre l'anàlisi de dades. Fem aquí esment d'aquells recursos que es poden trobar a les plataformes Coursera o EdX, ordenats per nivell de dificultat:

- Técnicas cuantitativas y cualitativas para la investigación
<https://www.edx.org/course/tecnicas-cuantitativas-y-cualitativas-upvalenciav-tcci201-x>
- Analyzing and Visualizing Data with Excel
<https://courses.edx.org/courses/course-v1:Microsoft+DAT206x+4T2016/info>
- Statistical Inference
<https://www.coursera.org/learn/statistical-inference>
- Master statistics with R
<https://www.coursera.org/specializations/statistics>

4.6. REFERÈNCIES

Ballart, X. *¿Cómo evaluar programas y servicios públicos?: Aproximación sistemática y estudios de caso*. Madrid: Ministerio de Administraciones Públicas, 1992.

Blasco, J. I Casado, D. *Avaluació de l'impacte*. Barcelona: Ivàlua, 2009. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 5).

Durlak, J. A. DuPre, E. P. "Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation". *American Journal of Community Psychology* (2008), núm. 41, p.327-50.

Humphrey, N. et al. *Implementation and process evaluation (IPE) for interventions in education settings: A synthesis of the literature*. Manchester: University of Manchester, 2016.

Humphrey, N. et al. *Implementation and process evaluation (IPE) for interventions in education settings: An introductory handbook*. Manchester: University of Manchester, 2016

Lázaro, B. Obregon, I. *Avaluació de la implementació*. Barcelona: Ivàlua, 2009. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 4)

Mowbray, A. Et al. "Fidelity criteria: Development, measurement, and validation". *American Journal of Evaluation* (2003), Vol. 24, Núm. 3, p. 315-340.

Nilsen, P. *Making sense of implementation theories, models and framework*. 2015.

Ramírez, R. Brodhead, D. *Utilization Focused Evaluation. A primer for evaluators*. Southbound Sdn, 2013.

Rossi, P.H.; Lipsey, M.W.; Freeman, H.E. *Evaluation: a systematic approach*. 7th ed. London: SAGE, 2004.

5. AVALUACIÓ DE L'IMPACTE

5.1. PUNT DE PARTIDA I MOTIVACIÓ

Fa temps que les entitats del tercer sector involucrades en intervencions per a persones en situació de vulnerabilitat han començat a preguntar-se sobre la capacitat d'impacte dels seus programes i serveis. Un cop realitzada la cerca de finançament, establerta l'organització i metodologia de la intervenció i realitzades les activitats previstes, l'interès a poder saber en quina mesura les intervencions en qüestió han aconseguit modificar o mitigat la problemàtica que pretenien corregir. Addicionalment, donants, socis, o també els gestors del sector públic, comencen també a insistir en la necessitat que les entitats del tercer sector aportin evidència empírica sobre el seu impacte, per tal de continuar col·laborant o prioritzar el finançament d'unes o altres iniciatives.

El primer pas per avaluar l'impacte d'un programa consisteix a concretar la pregunta (o les preguntes) objecte d'estudi. Aquest és un punt molt rellevant, ja que darrera de l'interès d'avaluar el programa que puguin compartir els seus responsables no sempre existeix un acord clar sobre quina mena d'impactes haurien de mesurar-se, o, ni tan sols, sobre el sentit últim d'intentar avaluar-los.

5.2. QUÈ VOL DIR AVALUAR L'IMPACTE D'UNA INTERVENCIÓ

Les preguntes que es planteja l'avaluació d'impacte són de naturalesa causal. Per exemple: ha millorat l'ocupabilitat d'un grup d'aturats que han realitzat un programa de formació ocupacional? Ha disminuït la reincidència penal d'un grup d'interns que van participar en un servei de suport psicològic dins de la presó? Aquestes preguntes porten implícita la frase "**com a conseqüència d'haver participat en el programa**", és a dir, l'interès pel que els hagués passat als participants o usuaris del programa en cas que no hi haguessin participat i, per tant, haguessin fet allò que habitualment fan (situació que denominarem "status quo" o "*business as usual*"). Anticipem ja ara que les preguntes d'impacte no tenen perquè adreçar-se sempre a la totalitat del programa; en canvi, poden orientar-se a determinar si petites innovacions dins d'un programa generen un canvi en la direcció que s'havia pretès de forma tal que la intervenció esdevingui més efectiva (Imbens i Rubin, 2015).

El concepte de **contrafactual**, noció nuclear de l'avaluació d'impacte, remet precisament a aquest escenari "hipotètic", no observat: el que hagués passat als participants en cas de no haver passat pel programa o servei. En el pla teòric, direm que l'impacte d'un programa és la diferència entre els resultats (outcomes) dels participants i els resultats de l'escenari contrafactual un cop finalitzat el programa; entre el nivell d'ocupabilitat dels participants en un programa de formació ocupacional i el nivell que aquests mostrarien si no hi haguessin

participat; entre el grau de reincidència dels presos que reben suport psicològic i el que aquests demostrarien en cas que no l'haguessin rebut. És així com podrem anar més enllà de la correlació estadística (entre participar en un programa i els outcomes observats) i destil·lar en quina mesura la participació és causa d'un o altres outcomes.

El problema fonamental de l'avaluació d'impacte és que l'escenari contrafactual no és, doncs, observable, no és real. Això fa que aquest escenari hipotètic s'hagi d'inferir de forma indirecta, mitjançant la identificació d'un **grup de control o comparació** idèntic a tots els efectes al grup d'individus que participen en el programa.

Quan comparem individus (o famílies, o escoles, o territoris) diferents, participants i no participants, correm el risc que uns i altres siguin diferents en certes característiques més enllà de la participació (o no) en el programa, característiques que poden ser rellevants de cara a la consecució dels outcomes. Aquestes diferències es coneixen a la literatura com el "**biaix de selecció**" en l'accés al programa: o bé els participants són seleccionats activament en base a determinats requisits objectius i/o valoracions subjectives; o bé són els mateixos participants els que s'autoseleccionen – aquests podrien ser els qui estan més motivats per recuperar-se, o els qui creguin que tenen més opcions de trobar feina si participen en el programa.

Així doncs, tota avaluació d'impacte que aspiri a poder respondre correctament si els canvis en els outcomes d'interès són conseqüència del programa, necessita identificar un grup de comparació molt similar al grup que efectivament ha participat del programa. En aquest sentit, l'estimació de l'impacte del programa serà tan bona com vàlid sigui el grup de comparació. D'aquesta manera, participants i controls haurien de ser idèntics en totes aquelles característiques relacionades amb la probabilitat d'èxit (o fracàs) del programa i en la decisió de participar del programa. En definitiva, es tracta d'eliminar, o de limitar tant com sigui possible, el biaix de selecció. Si això fos possible, podrem afirmar que les diferències que arribem a observar una vegada finalitzada la intervenció entre un i altre grup representen el seu impacte.

REQUADRE 14: PROGRAMA CONTRA LA DROGODEPENDÈNCIA - PLANTEJAMENT

Imaginem que una entitat que dóna tractament i suport a persones amb drogodependència a Catalunya vol estimar l'impacte d'un programa, en particular vol saber si els pacients que hi han participat en el passat han millorat en termes de no reincidència en el consum d'estupefaents, millora en el seu estat d'ànim i motivació, millora en el seu comportament social, inserció laboral, i allunyament de causes penals vinculades a delictes.

El programa consisteix en un treball individualitzat per a cadascun dels pacients, dut a terme durant nou mesos, amb components sanitaris i psicològics, i amb un fort èmfasi en les relacions i habilitats socials de la persona. L'entitat que l'ofereix té places suficients per poder atendre un màxim de 200 persones a l'any, en diferents centres del territori. El tractament és totalment voluntari i no té cap cost per al pacient (està subvencionat en la totalitat pel govern de la Generalitat de Catalunya). L'entitat fa un diagnòstic inicial dels pacients per tal de detectar comportaments antisocials que facin inviable la participació, incloent qüestions relacionades amb l'estat d'ànim i depressió. Tot i aquest descart inicial, es produeix un excés de demanda, i no totes les persones interessades a participar-hi acaben rebent la intervenció

5.3. MÈTODES PER AVALUAR L'IMPACTE

5.3.1. Com no estimar l'impacte d'un programa: estratègies que sovint generen resultats esbiaixats

Aproximació "abans versus després"

Un possible intent per determinar l'impacte de la intervenció contra la drogodependència passa per comparar com es trobaven els individus participants abans de rebre el servei i com es troben després. En aquest cas, el contrafactual assumeix que els outcomes dels participants s'haurien mantingut en el mateix nivell sense la intervenció. El problema amb aquest supòsit és que els canvis que podrien atribuir-se a la intervenció es confonen amb els canvis que haguessin experimentat els beneficiaris per qualsevol altra circumstància. En efecte, les dimensions sobre les quals el programa pretén generar un impacte estan sotmeses a la influència d'altres factors (altres polítiques públiques, l'evolució del mercat de treball, situacions familiars, reversió a la mitjana), que ocorren al mateix temps que el programa en qüestió i que podrien tenir una incidència sobre els resultats dels participants.

En el cas que ens serveix d'exemple, podria succeir que, al mateix temps que es desplega el programa, hagi empitjorat el context econòmic i social, especialment en l'entorn dels participants. Un estudi abans-després podria detectar que els participants, un cop passat el programa, no han millorat les seves perspectives laborals ni el seu benestar psicològic. Podríem concloure, en aquest cas, que el programa no ha tingut èxit? La resposta és que no podem determinar amb certesa si el programa ha tingut impacte o no. Podria, fins i tot, donar-se el cas que hagi estat gràcies al programa que la situació laboral i psicològica dels participants no ha empitjorat, cosa que hagués succeït, conseqüència del context econòmic general, si el programa no hagués existit.

Únicament quan podem estar segurs que no existeix cap factor causal que, tret del programa, pugui haver produït els canvis observats en l'outcome, podem llavors considerar que la comparació abans-després ens informa de l'impacte de la intervenció.

En ocasions, aquesta serà l'única aproximació possible per aproximar-nos a l'impacte. Per exemple, hi ha situacions en què totes les unitats que integren la població diana, en cas que ho sol·licitin, poden rebre el programa durant el mateix període temporal; és a dir, l'accés al programa és universal i per tant no cal establir cap tipus de prioritització. En aquests casos, és probable que no tinguem més opció disponible que la comparació de l'outcome abans i després del programa per a les unitats de la població diana "tractades", controlant per tots aquells atributs i variables de l'entorn que hagin estat susceptibles de modificar-se durant l'interval d'anàlisi.

Participants versus no participants

Un altre intent d'avaluació de l'impacte consisteix en comparar els resultats dels individus que van rebre la intervenció amb els d'aquells que no la van rebre tot i satisfer els requeriments bàsics d'accés al programa. En aquest cas, el supòsit és que les persones que no participen del programa constitueixin un contrafactual vàlid dels tractats. El problema és que acostuma a succeir que participants i no participants difereixen en tot un seguit de característiques algunes de les quals ni tan sol són observables, el que pot representar un fort biaix en la selecció de les mostres i, per tant, donar lloc a estimacions incorrectes de l'impacte (Angrist i Pischke, 2016).

REQUADRE 15: PROGRAMA CONTRA LA DROGODEPENDÈNCIA – BIAIX DE SELECCIÓ

Si els gestors d'un programa contra la drogodependència decideixen prioritzar l'accés d'aquells individus més motivats i aparentment disposats a no reincidir, la comparació dels outcomes finals entre participants i no participants donarà una visió sobredimensionada de l'efecte de la intervenció: com de millor els vagi als participants podrà ser tan o més conseqüència d'aquesta diferència inicial motivacional que de la participació en el programa. En cas contrari, si es decideix prioritzar l'accés dels col·lectius amb menys implicació emocional, aleshores acabarem subestimant l'impacte de la intervenció.

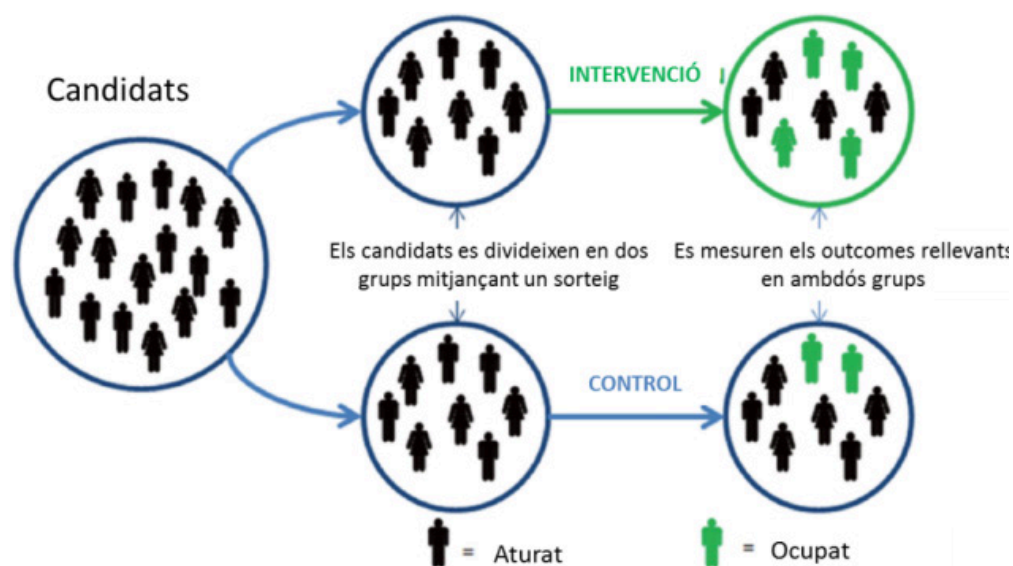
D'aquesta manera, hi ha dues situacions en les quals la comparació entre persones tractades i no tractades podrà donar mesures no esbiaixades de l'impacte: en primer lloc, quan la selecció estigui basada únicament en variables observables, que a més a més podem mesurar-les; en segon lloc, quan en la selecció dels participants ha existit un component d'atzar.

5.3.2. Estratègies d'identificació, contrafactual i estimació de l'impacte d'un programa

Experiments aleatoritzats

Quan un programa o servei presenta una situació d'**excés de demanda** (té més sol·licituds d'accés que places disponibles), i quan aquesta situació es resol mitjançant l'atzar (un sorteig), aleshores es genera una oportunitat interessant per avaluar-ne l'impacte de forma no esbiaixada. L'assignació aleatòria funciona exactament com la extracció de dues mostres representatives, on partint d'una població diana una part queda en el grup de tractament i la resta en el grup de control, generant d'aquesta manera dos grups de comparació estadísticament idèntics en tots els atributs, eliminant així el biaix de selecció.

Figura 9. El disseny amb assignació aleatòria pura: tractats versus controls



Font: Adaptat de Haynes et al. (2012: 9).

Si bé és cert que el plantejament d'un experiment aleatoritzat no sempre és viable com a mètode d'avaluació de l'impacte d'una política (per exemple, quan es tracta d'avaluar els efectes de canvis institucionals de gran abast), determinades **circumstàncies poden afavorir la seva factibilitat** més enllà d'un excés de demanda (Duflo, Gennerster i Kremer, 2007): quan un programa s'implementa gradualment per territoris o entre grups de persones (s'aleatoritza aleshores el moment d'accés al programa); quan el programa vol testar una innovació (s'aleatoritza qui participa en el programa "convencional" i qui ho fa en el programa "innovat"); quan es vol testar algun mecanisme d'estímul o crida a la participació en una determinada acció (en el cas que la participació en el programa sigui més baixa de l'esperada, es poden enviar en forma aleatòria recordatoris a la població general), etc.

Un tret característic de les avaluacions amb un disseny aleatori és que el mecanisme de selecció s'ha de resoldre abans que el programa comenci, la qual cosa requereix un esforç de planificació i anticipació organitzativa que altres esquemes d'avaluació no requereixen.

REQUADRE 16: PROGRAMA CONTRA LA DROGODEPENDÈNCIA – ALEATORITZACIÓ

L'accés al tractament podria venir condicionat per una valoració inicial que es faria, per mitjà d'una entrevista personal, a tota persona interessada a participar-hi i que reunís als requisits objectius bàsics. Aquesta valoració distingirà tres categories: 1) persones que es creu respondran bé al programa; 2) persones que es creu necessiten una intervenció diferent (i que per tant no respondran al programa); 3) persones que no es té clar si es beneficiaran o no del programa. Els individus del grup 1 passarien a rebre el tractament, els del grup 2 són derivats a d'altres recursos, i és entre les persones del grup 3 que s'opta per assignar aleatòriament la participació (comptant, per exemple, que el 70% han d'acabar participant i el 30% restant no). D'aquesta manera, estimar l'impacte de la intervenció es reduiria a comparar els outcomes entre les persones tractades i els no tractats dins del grup 3.

L'esquema d'aleatorització més adient estarà determinat per la pregunta que es vulgui respondre. Si volem saber si rebre la intervenció de suport a la drogodependència aconsegueix un canvi significatiu en els outcomes dels participants, l'esquema plantejat en l'exemple anterior és correcte.

Tanmateix, es podria interessar-nos conèixer si incorporar la figura d'un tutor genera una millora substancial amb relació al programa "convencional". Atès que partim d'una situació d'incertesa, conèixer l'impacte d'aquest canvi és rellevant abans de escalar-ho. Així, es podria plantejar que durant un temps, entre els participants que rebin la intervenció, uns tinguin el complement del tutor i altres no, a partir d'una assignació aleatòria. L'impacte d'aquest recurs es calcularia comparant els outcomes d'interès entre els qui l'han rebut i els qui no.

Una altra alternativa d'aleatorització és l'**assignació per fases**. Imaginem que el nostre programa té normalment una demanda molt similar al nombre de places existents, de forma que rarament es produeixen situacions de sobredemanda. No obstant això, en un determinat moment, la demanda de persones elegibles creix substancialment per sobre de l'oferta de places. Una forma de prioritzar de forma justa i que alhora ens pot permetre avaluar l'efecte del programa en el curt termini és aleatoritzar quins participants rebran primer la intervenció i mesurar el canvi amb relació als que la rebran més endavant.

Finalment, un cas també interessant és l'**ús d'estímuls**, aplicable en aquelles circumstàncies en què no és possible negar el tractament a un grup de persones. Suposem que el programa que ens serveix d'exemple està considerant demanar i avaluar l'efecte de la implicació activa dels familiars en l'acompanyament dels participants. Òbviament, fóra inoportú impedir que els familiars que s'hi volen involucrar deixin de fer-ho de resultes d'un sorteig.

La valoració de l'impacte d'aquesta modificació es podria resoldre per mitjà de l'enviament aleatori de correus electrònics, missatges a telèfons mòbils o visites cara a cara incentivant la implicació familiar en el tractament. L'estimació de l'impacte d'aquest component d'implicació seria, en aquest esquema, estadísticament més complex que les estimacions anteriors.⁸ Tanmateix, aconseguiria, gracies al mecanisme de l'aleatorització, eliminar el biaix de selecció.

Regressió discontinua

La regressió discontinua es pot aplicar quan l'assignació al tractament depèn del fet que una variable prengui un conjunt de valors determinats; això és, quan la possibilitat de rebre'l s'estableix a partir d'un tall en una variable contínua (l'edat, la renda, una puntuació acadèmica, etc.). Quan això succeeix, al voltant d'aquest punt de tall s'observarà una **discontinuitat en la probabilitat de rebre el programa** que podem aprofitar a efectes d'identificar un possible grup de comparació (Skovron i Titiunik, 2015). Més en particular, la variable en qüestió (anomenada "rating") hauria d'estar vinculada als outcomes d'interès.

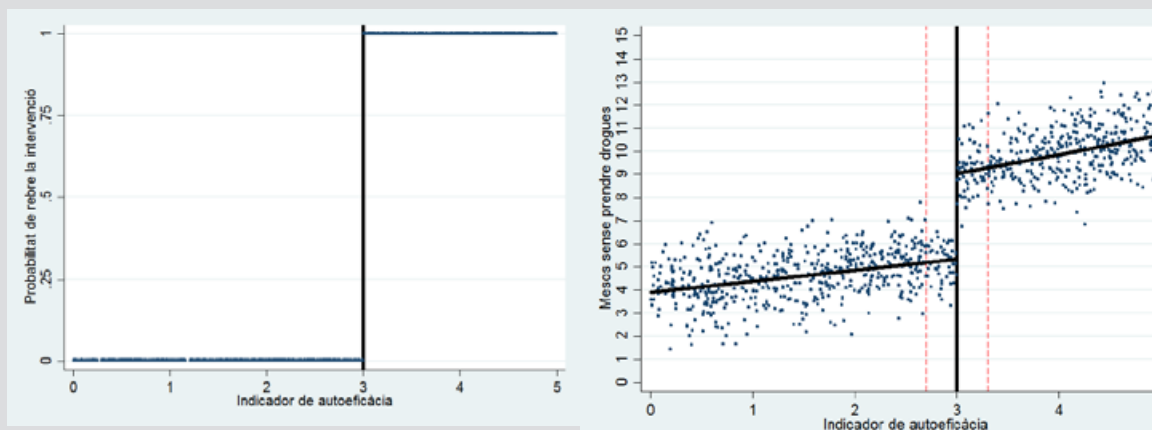
Per exemple, en el cas d'una beca per als estudis universitaris, podria establir-se que els estudiants que obtenen una nota mitjana de 8 punts o més en la prova d'ingrés a la universitat podran optar a la beca, mentre que els que obtinguin una puntuació inferior a 8 podran accedir a la universitat, però no tindran dret a l'ajut. Si estiguéssim interessats en conèixer l'impacte de rebre la beca en el sou de les persones cinc anys després d'acabats els estudis, una comparació entre alumnes que van rebre l'ajut i alumnes que no el van rebre estaria esbiaixada pel fet que els alumnes amb beca parteixen d'un nivell competencial i, probablement, motivacional superior. Per tant, no podrem distingir si un sou més alt en el grup becat és conseqüència de l'ajut o d'unes millors condicions de partida. En canvi, es podria argumentar que els alumnes que van obtenir entre 8 i 8,2 no són gaire diferents dels que van obtenir 7,8 a 7,99. Llavors, si volguéssim analitzar l'impacte d'una beca escolar sobre les oportunitats laborals posteriors, una comparació entre aquestes persones ens donaria una idea de l'impacte no esbiaixada.

⁸ En aquest cas caldrà fer una regressió de variables instrumentals. Atesa la seva complexitat tècnica, aquest mètode ha quedat exclòs d'aquest capítol. Per saber-ne els detalls, podeu consultar Angrist i Pischke (2009).

REQUADRE 17: PROGRAMA CONTRA LA DROGODEPENDÈNCIA – REGRESSIÓ DISCONTÍNUA

Suposem que per a què una persona sigui admesa al programa ha de tenir una puntuació per sobre de 3 en un test d'autoeficàcia generalitzat⁹. Els que queden per sobre d'aquesta puntuació són admesos, mentre que la resta són descartats per participar en la intervenció.

Gràfic 1. Exemple de regressió discontinua.



Font: Elaboració pròpia.

Com podem observar en el gràfic de l'esquerra, les persones que queden a la esquerra del llindar tenen una probabilitat zero de rebre el tractament, mentre les que estan a la dreta tenen una probabilitat igual a 1, això és, del 100%. Tot i què l'algorisme per fer la estimació d'impacte en aquest tipus de regressió és més complex que en el cas de l'aleatorització, la idea darrera la regressió discontinua no deixa de ser similar: al capdavall és un factor exogen, fora del control dels participants potencials, el qui determina l'assignació. La principal diferència és que en el cas de la regressió discontinua, la comparació la fem posant molt pes en les observacions (individus) que estan al voltant del llindar de tall (línies vermelles), i per tant no ens dona una idea global de l'impacte, sinó una estimació "local", centrada en aquells participants que tenen atributs similars i es troben a prop de la nota de tall. La discontinuïtat que s'observa en la mesura de l'outcome (en aquest cas, "mesos sense prendre drogues") al voltant del punt de tall entre participants i no participants (gràfic dreta) és l'impacte del programa.

⁹ <http://userpage.fu-berlin.de/%7Ehealth/spanscal.htm>

Per tal que aquesta estratègia d'identificació del contrafactual es pugui utilitzar cal que s'acompleixin certes condicions:

1. La variable de "rating" no pot ser influenciada o causada pel tractament. És a dir, aquesta variable s'ha de mesurar abans de l'inici del programa o altrament ha de ser un factor constant, no subjecte a canvis.
2. El llindar ha de ser exogen, és a dir, no pot ser objecte d'aplicació discrecional ni existir marge per a què els individus triïn en quin costat del punt de tall s'ubiquen.
3. Cap altra variable rellevant, més enllà de la probabilitat de participar o no en el programa, pot presentar discontinuïtats entorn del mateix punt de tall que decideix l'assignació.

Dobles diferències

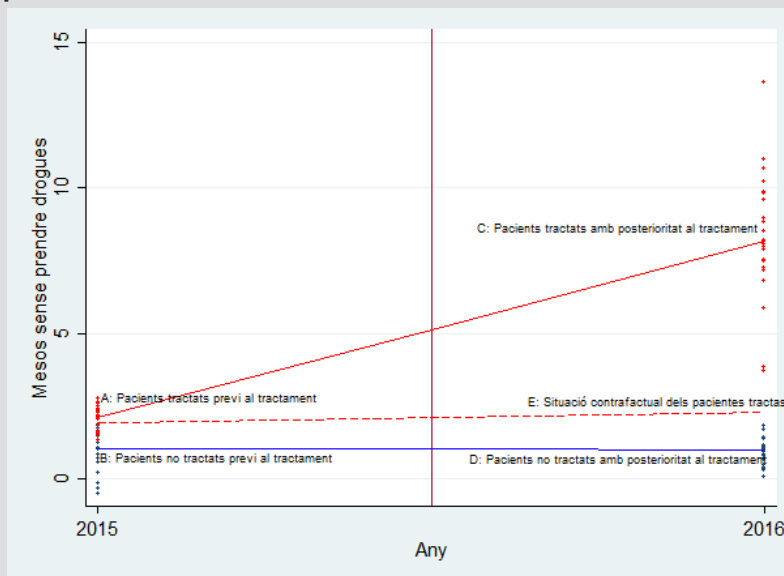
Els models de dobles diferències representen un pas més enllà del disseny d'anàlisi "abans-després" o de comparació un cop finalitzada la intervenció entre tractats i no tractats. A diferència d'aquests dissenys, la tècnica de les dobles diferències (o diferències-en-diferències) tracta de controlar els efectes dels factors contemporanis al programa comparant allò que els succeeix als individus o unitats que hi participen amb allò que els succeeix a aquells que no hi participen en un mateix període d'observació (Blasco i Casado, 2009).

L'aplicació més habitual dels models de dobles diferències és aquella que estableix la comparació entre dos grups d'individus al llarg del temps: en un primer moment, cap dels dos grups rep la intervenció; en un segon moment, en canvi, un grup la rep (tractament) i l'altre no (control). L'estratègia d'estimació de l'impacte que proposen aquests models es basa en la **comparació de les diferències en els outcomes de cada grup abans i després de la intervenció**. Assumint que allò que diferencia ambdós grups des de l'inici roman constant en el temps, la diferència entre aquestes dues diferències assenyalaria l'impacte del programa.

REQUADRE 18: PROGRAMA CONTRA LA DROGODEPENDÈNCIA – DOBLES DIFERÈNCIES

Imaginem que els gestors del programa s'estan plantejant incrementar les hores de suport psicològic durant l'estada al programa i volen saber, abans d'escalar el programa, si seria beneficiós de cara a reduir la reincidència en el consum de drogues. Com que l'aleatorització de l'increment d'hores entre els pacients de cada centre pot resultar desaconsellable, una possibilitat seria que un grup de centres (potser aquells que apleguen perfils poblacionals més vulnerables) disposessin de la possibilitat d'incrementar les hores i la resta no. Aleshores, la tècnica de les dobles diferències consistiria a: i) calcular la diferència en els outcomes dels centres amb la ampliació horària abans i després de començar el programa; ii) calcular la diferència en els outcomes dels centres sense ampliació també abans i després de la posada en marxa del programa; iii) calcular la diferència entre aquestes dues diferències. El supòsit clau que estem fent és que qualsevol diferència pre-programa entre uns i altres centres, sigui més gran o més petita, es manté constant encara que el programa no hagués existit.

Gràfic 2. Exemple de dobles diferències



Font: Elaboració pròpia.

En el cas que ens ocupa, assumim que els pacients que reben el tractament ampliat en l'any 2015 (línia vermella contínua del gràfic 2), en cas que no l'haguessin rebut, haguessin evolucionat entre 2014 i 2016 d'igual forma com ho fan els no tractats (línia vermella de punts), mantenint-se constant la diferència inicial (2014) entre uns i altres.

El principal punt feble dels models de dobles diferències llavors és la possible existència de factors inobservables rellevants amb relació a l'outcome que variïn al llarg del temps, i que ho puguin fer amb intensitats diferents entre participants i no participants. Si volem que els resultats d'una avaluació d'impacte que utilitzi un disseny de dobles diferències resultin creïbles, caldrà oferir arguments que permetin descartar l'existència d'aquesta limitació (Casado i Todeschini, 2013; Alegre, 2015).

Matching o aparellament

Aquesta tècnica crea un grup de control ex-post que s'assembla al màxim possible al grup de tractament en aquelles característiques o atributs dels participants que són alhora rellevants i observables. L'aplicació d'aquest mètode per avaluar l'impacte d'una política pot aplicar-se en aquells casos en què, amb posterioritat a la intervenció pública, disposem d'un ampli volum d'informació tant d'una mostra d'individus que han estat beneficiaris del programa com d'una mostra de persones no beneficiàries. En concret, per a cadascun dels individus d'ambdós grups, cal tenir informació sobre el valor que pren en cada cas l'outcome d'interès i també sobre tots aquells factors (característiques individuals, familiars, de l'entorn, etc.) que, d'una banda, poden haver determinat el procés pel qual els individus han decidit

participar en el programa i, de l'altra, poden tenir efectes sobre el valor que pren l'outcome en qüestió.

El matching proposa utilitzar tota la informació disponible per construir un grup de comparació entre els individus que no es beneficien del programa. Per fer-ho, el mètode busca, **per a cadascun dels individus que componen la mostra de tractats, una parella o match (d'aquí el nom de la tècnica) que sigui el més semblant possible** en el sentit que acabem de descriure.

La pretensió darrera de la tècnica del matching és obtenir, mitjançant procediments estadístics, allò que obtindríem a partir de l'assignació aleatòria del tractament. Però mentre que l'assignació aleatòria distribueix de forma equitativa les característiques observables i les no observables entre el grup de control i el de tractament, amb el matching només podem garantir que es distribueixin equitativament les característiques observables. En altres paraules, hem d'assumir que no hi ha cap variable rellevant no observable que difereixi sistemàticament entre el grup de tractament i el de comparació i que, per tant, l'outcome del grup de tractament si no hagués participat o s'hagués beneficiat del programa (és a dir, el contrafactual) equival a l'outcome del grup de comparació que, realment, no hi ha participat.

REQUADRE 19: PROGRAMA CONTRA LA DROGODEPENDÈNCIA – MATCHING

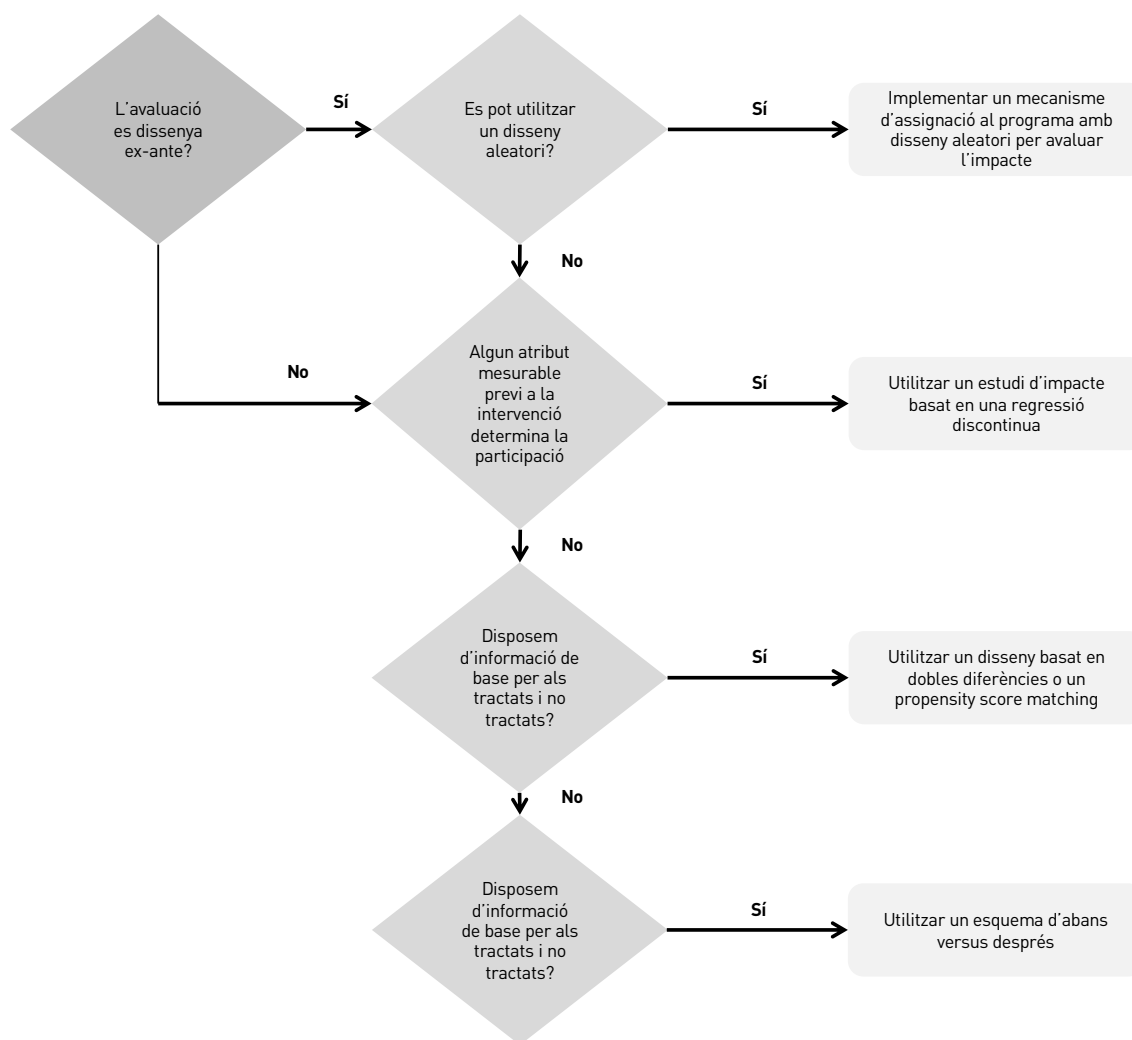
Posem per cas que es posa en marxa una iniciativa destinada a donar un fàrmac nou als homes entre 18 i 35 anys que participen en el programa i no hagin registrat cap incidència en els darrers tres mesos. Tanmateix, la decisió de prescriure o no la nova medicació és facultativa del psiquiatra de referència i, per tant, no totes les persones amb els atributs bàsics passen a ser tractades amb la nova medicació. Aquesta circumstància permetria construir un grup de comparació amb les persones "elegibles" no seleccionades per rebre la medicació, a partir dels atributs formals i incorporant altres característiques observables (estudis, història laboral, procedència, situació familiar, etc.). Imaginem que un dels pacients que rep la nova medicació té 20 anys, no ha registrat cap incidència, ha assolit un nivell educatiu universitari, viu amb els seus pares i mai ha treballat. Llavors es tractarà de buscar una persona d'entre els no participants que tingui exactament aquests atributs i comparar els resultats de l'un i l'altre. Així, la tècnica realitza emparellaments entre participants i no participants que són semblants, i calcula l'impacte com la diferència mitjana dels outcomes de les parelles.

El supòsit bàsic d'aquest mètode és que, en mitjana, un cop s'ha tingut en compte l'efecte de les variables observables, els participants haguessin obtingut el mateix outcome que els no participants si la política no hagués existit. Per tal de minimitzar el biaix de selecció per "inobservables", el matching requereix la utilització d'un conjunt de variables de control el més ampli possible, que tingui en compte totes aquelles variables per a les quals existeix evidència d'interacció amb la participació i amb l'outcome d'interès. Altrament, la credibilitat dels resultats obtinguts quedarà erosionada.

Recapitulant: l'elecció dels mètodes d'avaluació d'impacte

En els punts anteriors hem passat revista als mètodes que més adeqüen i més aplicació han tingut en l'avaluació de programes i intervencions socials. I hem argumentat que l'elecció d'un o altre mètode depèn tant de les preguntes d'impacte que vulguem respondre com del mecanisme d'assignació que governi la participació en el programa. La figura 10 resumeix les alternatives metodològiques ordenades segons la seva robustesa (de més a menys robustes) i d'acord amb les opcions d'assignació pròpies de cada programa.

Figura 10. Alternatives per fer una avaluació d'impacte



Font: Elaboració pròpia.

REQUADRE 20: EL BALANÇ SOCIAL I LES SEVES LIMITACIONS

El Balanç Social és una eina de rendició de comptes i mesura de la responsabilitat social, ambiental i de bon govern per a les organitzacions i empreses de l'Economia Solidària. L'eina consta d'un qüestionari de 60 indicadors que es pot completar en línia, així com de qüestionaris d'avaluació externa a complimentar per persones treballadores, voluntàries i clientes o usuàries. Es tracta d'un exercici que permet comparar una organització i el seu funcionament envers uns estàndards ètics, com la igualtat d'oportunitats, la qualitat democràtica o la sensibilitat vers l'apoderament social del territori. En aquest sentit, la filosofia de l'eina s'assimila als sistemes de gestió i acreditació de la qualitat, que també defineixen unes dimensions bàsiques i instruments per diagnosticar com es troba l'organització respecte uns estàndards de referència. En aquest cas, la filosofia és similar, si bé es treballa amb les dimensions clau i els estàndards propis de l'Economia Social i Solidària.

Tanmateix, si bé l'eina permet mesurar la responsabilitat social de l'organització, és important no confondre aquesta idea amb l'avaluació de l'impacte. En aquest sentit, cal destacar que el Balanç Social no és una forma d'avaluar l'impacte social, ni de l'organització ni del conjunt de la Xarxa d'Economia Solidària, ja que no es recull informació que permeti mesurar canvis socials més enllà de la pròpia organització o xarxa, ni tampoc estimar què hauria passat en la seva absència, és a dir, no podem saber com de diferent seria la realitat sense ella.

<http://ensenyaelcor.org/balanc-social-vaadin/login#!login>

http://mercatsocial.xes.cat/wp-content/uploads/sites/2/2016/04/Manual-us-questionari-BSXES-2015_29042016.pdf

Font: *Elaboració pròpia.*

De forma general, si volem fer avançar l'avaluació d'impacte dins del tercer sector social, caldrà fer passes decidides en tres direccions, ja tractades en capítols anteriors:

- **Una teoria del canvi ben definida.** El sentit, rellevància i utilitat de les preguntes d'impacte només s'entenen amb relació a la teoria del canvi que l'avaluació pretén posar a prova, en particular a la relació esperada entre els outputs del programa (els productes de les activitats) i els outcomes d'interès (a curt, mig i llarg termini). Disposar de bons models lògics i fonaments causals ben establerts en el disseny dels programes és, per tant, condició necessària per donar sentit a tot intent d'avaluar-ne els impactes.
- **Uns indicadors d'outcome ben dissenyats i amb informació disponible.** Díficilment es podran resseguir els impactes de les intervencions si no es té ben especificat sobre quines variables es volen observar o si no existeix informació (o obtenir-la és excessivament costós) sobre els indicadors en qüestió. En aquest punt, torna a fer-se evident la necessitat de comptar amb fonts i sistemes de gestió de la informació adequats a les necessitats d'implementació i avaluació del programa.
- **Una bona estratègia d'identificació del contrafactual** (grups de control o comparació). Ja sigui mitjançant l'establiment de mecanismes d'aleatorització de la participació en situacions d'excés de demanda o identificant mostres de no participants resseguibles en el temps o comparables a les mostres de participants, és de vital importància preveure d'on s'obtindrà el contrafactual requerit per tota avaluació d'impacte i com se n'obtindrà la informació necessària per dur-la a terme.

5.4. RECURSOS

En l'actualitat existeix una multiplicitat de recursos disponibles en forma gratuïta a la web, a partir dels quals aprofundir sobre les característiques dels mètodes descrits en aquest capítol. Fem aquí esment d'aquells recursos que es poden trobar a les plataformes Coursera o EdX, ordenats per nivell de dificultat.

Anàlisi de dades i estadística bàsica

- **Técnicas cuantitativas y cualitativas para la investigación**

<https://www.edx.org/course/tecnicas-cuantitativas-y-cualitativas-upvalenciav-tcci201-x>

- **Analyzing and Visualizing Data with Excel**

<https://courses.edx.org/courses/course-v1:Microsoft+DAT206x+4T2016/info>

- **Statistical Inference**

<https://www.coursera.org/learn/statistical-inference>

- **Master statistics with R**

<https://www.coursera.org/specializations/statistics>

Estadística intermèdia i econometria

- **Critically Analyze Research and Results Using R**

<https://www.coursera.org/specializations/social-science>

- **Regression models**

<https://www.coursera.org/learn/regression-models>

- **Policy analysis using interrupted time series**

<https://www.edx.org/course/policy-analysis-using-interrupted-time-ubcx-itsx-2>

Avaluació amb disseny aleatori

- **Evaluating social programs**

https://courses.edx.org/courses/MITx/JPAL101_1x/3T2014/info

- **Evaluación de Impacto de Programas Sociales**

<https://www.edx.org/course/evaluacion-de-impacto-de-programas-mitx-jpal101spax>

- **Designing and running randomized evaluations**

<https://www.edx.org/course/designing-running-randomized-evaluations-mitx-jpal102x>

5.5. REFERÈNCIES

Alegre, M.A (2015). *Com avaluar l'impacte de les polítiques educatives*. Barcelona: Ivàlua, 2015. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 11)

Angrist, J. Pischke J.S. *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton university press, 2009.

Angrist, J. and Pischke J.S. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton Press. 2016.

Blasco, J. I Casado, D. *Avaluació de l'impacte*. Barcelona: Ivàlua, 2009. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 5)

Casado, D. Todeschini, F.A. *Com avaluar l'impacte de les polítiques actives d'ocupació*. Barcelona: Ivàlua, 2013. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 10)

Duflo, E. Glennerster, R. Kremer, M. *Using Randomization in Development Economics Research: A Toolkit*. A: T. Schultz, Strauss, J. *Handbook of Development Economics*. Vol. 4. Elsevier, 2008

Heynes, L. Service, O. Goldacre, B. Torgerson, D. *Test, Learn and Adapt. Developing Public Policy With Randomised Controlled Trials*. Behavioural Insights Team, Cabinet Office, 2012.

Imbens, G. W. Rubin, D. B.. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

Skovron, C. Titiunik, R. *A practical guide to regression discontinuity designs in political science*. Working Paper. University of Michigan, 2015.

6. AVALUACIÓ ECONÒMICA O D'EFICIÈNCIA

6.1. PUNT DE PARTIDA I MOTIVACIÓ

La gran majoria d'objectius de canvi que es proposa el tercer sector social poden afrontar-se mitjançant diferents estratègies d'intervenció. Per exemple, per tal d'alleugerir la càrrega o estrès que pateixen els cuidadors informals de persones dependents es poden promoure els ajuts a domicili, els centres de dia especialitats, cursos individuals o grupals sobre com proporcionar les cures i/o l'ús de les tecnologies per oferir suport als cuidadors. Quan del que es tracta és d'ajudar a trobar feina a persones en risc d'exclusió, són diverses les opcions que es poden prioritzar, segons es doni més importància a la formació, a la realització de pràctiques laborals o a un acompanyament personalitzat en el procés de recerca de feina. En un sentit o un altre, l'escassetat de recursos acaba obligant a apostar per unes alternatives en detriment d'altres.

L'objectiu de l'avaluació econòmica és contribuir a què en la presa de decisions, on necessàriament han de contribuir d'altres factors, es prengui també en consideració **la relació que les diverses alternatives exhibeixen en termes de resultats i costos; això és, la seva eficiència.**

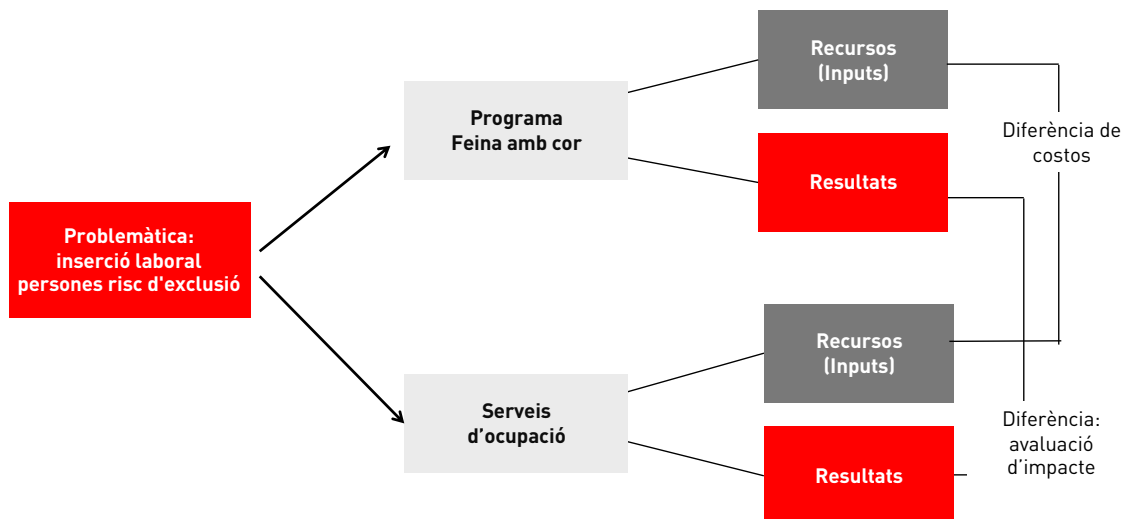
En els darrers anys, la preocupació per l'eficiència de les actuacions i els programes ha calat amb força dins el tercer sector social. Aquesta preocupació ha vingut de la mà de l'aposta per conèixer l'impacte d'aquestes actuacions. En conjunt, hi ha en joc la necessitat de donar arguments (als beneficiaris de les accions, però també a la ciutadania i a possibles finançadors) sobre quines estratègies d'intervenció acaben sent més efectives i compensant més i millor els recursos que implica la seva implementació; uns arguments que interessin tant a les entitats i als seus beneficiaris, com als seus finançadors (reals o potencials) i a la ciutadania en general.

6.2. QUÈ VOL DIR AVALUAR ECONÒMICAMENT UN PROGRAMA

L'avaluació econòmica consisteix bàsicament a **comparar els costos i els resultats que es deriven de diverses alternatives d'actuació, amb el propòsit fonamental d'identificar aquelles intervencions més eficients:** això es, aquelles que maximitzen els resultats assolits a partir d'una quantitat donada de recursos o, alternativament, aquelles que aconsegueixen un determinat nivell de resultats minimitzant els recursos emprats. L'avaluació econòmica descansa, per tant, sobre dos supòsits fonamentals: d'una banda, l'existència d'alternatives per assolir un determinat resultat i, de l'altra, l'heterogeneïtat entre elles en termes tant de la magnitud dels resultats assolits com de la quantia de recursos esmerçats en cada cas. D'una forma molt esquemàtica, la figura 11 il·lustra amb un exemple les principals variables

implicades en una avaluació econòmica. En aquest cas, es compara el que representaria en termes de recursos i resultats la participació en un programa d'inserció laboral per a persones vulnerables (programa Feina amb cor) amb el que suposaria, com a alternativa, la no provisió de cap servei específic per a aquest col·lectiu (ús dels serveis d'ocupació "convencionals").

Figura 11. Esquema d'una avaluació econòmica



Font: Elaboració pròpia.

L'avaluació econòmica passa necessàriament per la comparació de, com a mínim, dues alternatives d'acció. Així, s'avalua l'impacte que té la implementació del programa objecte de l'avaluació respecte d'una situació alternativa, que pot ser la implementació d'un altre programa o bé l'absència de cap programa específic ("status quo"). I el mateix succeeix amb l'avaluació dels recursos emprats per dur a terme el programa: es consideren els recursos addicionals necessaris respecte d'un altre possible programa o respecte del status quo.

Per tant, una anàlisi que solament consideri els recursos necessaris per dur a terme un programa i els resultats obtinguts no és una avaluació econòmica. I tampoc ho és una anàlisi únicament dels costos de dues (o més) alternatives. En altres paraules, una avaluació econòmica és el procés sistemàtic d'identificar, mesurar i valorar els costos i els resultats de dues o més intervencions o programes alternatius, i fer-ne l'anàlisi comparativa posterior.

6.3. PREGUNTES I ELEMENTS CLAU D'UNA AVALUACIÓ ECONÒMICA

Hi ha una sèrie d'elements clau que han de ser considerats en qualsevol avaluació econòmica. Conèixer-los pot ser d'utilitat tant per realitzar anàlisi com per valorar críticament estudis realitzats (Raya i Moreno, 2013):

- **Quines són les alternatives que s'avaluen?** Cal tenir ben clar quina és la situació alternativa (o situacions alternatives) amb què es compara el programa objecte d'estudi, i s'han de comparar tant els costos com els beneficis de les alternatives en qüestió. També cal descriure la població objecte de la intervenció i la durada i intensitat de les diferents alternatives avaluades, fins i tot quan no existeixi un programa específic com a comparador i l'alternativa al programa avaluat sigui el status quo.
- **Quin és l'horitzó temporal de l'anàlisi?** L'horitzó temporal d'una avaluació econòmica és el període de temps durant el qual es mesuraran els beneficis i els costos de les diferents alternatives. Idealment, aquest període hauria d'incloure la totalitat dels impactes rellevants de les intervencions avaluades, tant pel que fa als costos com als beneficis. És cert, però, que en no poques ocasions s'estableixen horitzons amb una durada menor a causa de la impossibilitat de disposar d'informació adequada per estimar aquests impactes a llarg termini.
- **Quina es la perspectiva de la anàlisi?** Un element clau en qualsevol avaluació econòmica és la perspectiva o punt de vista des del qual es consideren i s'avaluen els costos i els resultats o beneficis. La seva elecció s'ha de derivar de l'objectiu de l'avaluació i dels seus destinataris. Qualsevol intervenció afecta els interessos de múltiples agents, i cadascun dels costos i beneficis d'un programa recau sobre un determinat grup de persones. Depenent de quina perspectiva s'adopti s'inclouran uns costos i beneficis o uns altres. **La perspectiva de la societat**, també denominada perspectiva social, es la que considera tots els costos i resultats rellevants independentment dels agents que els suporten o se'n beneficien. Aquesta és la perspectiva més comprensiva i la més adequada en el cas d'avaluacions que tinguin com a finalitat informar la decisió d'assignar recursos públics. És recomanable oferir la informació de forma desagregada perquè el destinatari de l'estudi pugui calcular els costos i beneficis que es derivarien de l'adopció **d'altres perspectives diferents (finançador del programa, individus, administracions públiques etc.)**
- **Com s'ha estimat l'efectivitat del programa?** És necessari explicitar la procedència de les dades sobre l'efectivitat de les actuacions (això és, sobre els seus impactes), si s'han obtingut d'assajos controlats o proves pilot, si procedeixen de la revisió de la literatura, si s'ha recorregut a models teòrics o a exercicis d'altres països i, en aquest cas, quins han estat els supòsits utilitzats.
- **S'han identificat, mesurat i valorat adequadament els costos i resultats rellevants de cada alternativa?** Una vegada determinada la diferència en resultats a través de l'avaluació de l'impacte, el pas següent consisteix a determinar la diferència en l'ús de

recursos i, per tant, en els costos de les alternatives considerades. La determinació del cost dels programes implica valorar monetàriament l'impacte de les diferents alternatives sobre el consum de diferents recursos. Això suposa incloure com a cost tot allò que suposi un sacrifici, és a dir, els recursos que, pel fet d'utilitzar-se en el programa, no es poden dedicar a un ús alternatiu: això és, els **costos d'oportunitat**.

Una vegada identificats els recursos afectats per les intervencions, cal assignar-los un valor monetari. Per fer-ho, és necessari obtenir informació sobre la quantitat consumida de cadascun dels recursos i multiplicar-los pel seu cost unitari. Com s'obté la informació sobre aquest cost unitari depèn del tipus de recurs. En ocasions es pot utilitzar un preu de mercat (p.ex. lloguer d'uns determinats espais o sou d'un treballador social); quan aquest preu no existeix (p.ex. el temps dels usuaris o el dels cuidadors informals), s'ha de recórrer a altres mètodes de valoració.

Car recalcar que només s'ha de considerar l'ús de recursos que sigui diferent entre les dues alternatives avaluades. Per exemple, si dos enfocaments per ajudar a la reinserció laboral comporten el mateix temps d'ús d'un orientador, el cost d'aquests recursos no s'ha d'incloure en l'avaluació econòmica. Un aspecte clau a l'hora de determinar quins costos s'inclouen en el càlcul del cost de les diferents alternatives avaluades és la perspectiva de l'estudi. Com s'ha comentat anteriorment, si s'adopta la perspectiva de la societat s'han d'incloure tots els costos rellevants, independentment de qui els suporta. Però si s'adopta una perspectiva més limitada, com la de les administracions públiques, no caldrà considerar la repercussió del programa sobre l'ús de recursos que afecta altres agents, com per exemple individus i empreses.

- **S'han ajustat els costos i resultats en funció del moment del temps en què es produeixen?** Si els costos i els resultats s'estenen al llarg de successius períodes de temps, s'han de homogeneïtzar tenint en compte la valoració desigual dels beneficis i costos en funció del moment en què es donen. Per dur a terme aquesta homogeneïtzació es recorre al procediment del descompte, que consisteix a expressar tots els valors futurs en el seu valor equivalent en el moment actual. Per fer-ho, s'utilitza el factor de descompte que quantifica la preferència temporal dels individus.
- **S'ha realitzat una anàlisi incremental dels costos i resultats?** L'interès no rau en els costos i beneficis mitjos de cada opció, sinó en els costos i beneficis incrementals – marginals – que suposa un programa respecte de l'utilitzat com a referència.
- **Es presenten els resultats mitjançant un índex de decisió?** Com a norma general, però particularment en el cas que l'anàlisi estigui específicament orientada a informar la presa de decisions, el resultat final s'ha de formalitzar a un índex de decisió. Aquest índex dependrà del mètode d'avaluació econòmica utilitzat. En una anàlisi cost-efectivitat, la mesura que s'utilitza per comparar les diferents alternatives entre si s'anomena ràtio cost-efectivitat incremental (RCEI) i, tal com el seu nom indica, s'obté dividint els costos incrementals actualitzats del programa entre els resultats addicionals obtinguts, mesurat en unitats naturals.

En una **anàlisi cost-benefici** s'utilitzen tres mesures alternatives: la **ràtio benefici/cost**, el valor actual net (**VAN**) i la taxa interna de rendiment (**TIR**). La primera d'aquestes mesures, tal i com el seu nom indica, s'obté dividint els beneficis actualitzats entre els costos actualitzats. Al seu torn, el VAN s'obté restant dels beneficis que es deriven del programa, degudament actualitzats, els costos que ha implicat dur-lo a terme, també actualitzats. Finalment, una mesura alternativa és la TIR, que es defineix com la taxa de descompte que iguala els beneficis i els costos.

- **S'ha incorporat l'efecte de la incertesa?** Les avaluacions econòmiques haurien d'incloure el tractament de la incertesa mitjançant anàlisis de sensibilitat. En l'aplicació pràctica de l'avaluació econòmica esdevé necessari establir hipòtesis sobre alguns dels factors que determinen la valoració que s'acaba fent dels costos, dels resultats i dels beneficis. Això és especialment cert en les avaluacions econòmiques ex-ante, que es formulen sobre intervencions que encara no s'han produït. Cadascun d'aquests supòsits tindrà un efecte sobre el resultat final de l'anàlisi. Per conèixer l'impacte que pot tenir la modificació d'aquests supòsit sobre els resultats de l'avaluació, és necessari dur a terme allò que es coneix com a anàlisi de sensibilitat¹⁰.
- **S'ha explicat la distribució dels costos i beneficis?** La presentació dels resultats en forma de mitjanes amaga la distribució dels guanys i les pèrdues derivades del programa avaluat. L'informe ha d'explicitar, com a mínim, i de forma detallada, qui guanya i qui perd, ja que aquesta informació constitueix un input bàsic per a la presa de decisions.

6.4. MÈTODES PER A L'AVALUACIÓ ECONÒMICA

Existeixen diferents mètodes per realitzar una anàlisi d'avaluació econòmica. Aquests es diferencien bàsicament en un únic aspecte: les unitats utilitzades per a la quantificació dels resultats o impacte de les diferents alternatives incloses en l'avaluació. Tal com s'ha detallat anteriorment, en una avaluació econòmica s'analitzen i es comparen entre si tant els costos (ús de recursos) com els resultats de dos o més programes o polítiques. En **tots els mètodes d'avaluació econòmica existents, la valoració de l'ús de recursos motivat per la implementació de les diferents alternatives avaluades es realitza en unitats monetàries**. En altres paraules, el cost que suposa aquest ús de recursos sempre es quantifica en euros, dòlars, lliures esterlines o qualsevol altra moneda. No obstant això, hi ha cinc maneres diferents de quantificar els resultats o impactes, que es corresponen amb quatre tipus d'avaluacions econòmiques diferents (Sabes-Figuera, 2016):

¹⁰ Per a més informació sobre el sentit i característiques dels càlculs de sensibilitat en el marc de l'avaluació de polítiques i programes, podeu consultar Sabes-Figuera, R. *L'avaluació econòmica en l'àmbit de les polítiques de justícia Una guia introductòria: Guia pràctica 14*. Barcelona: Ivàlua, 2016.

6.4.1 Anàlisi de minimització de costos (AMC)

En aquest tipus d'avaluació econòmica, els resultats o beneficis de les alternatives se suposen essencialment idèntics. Per tant, només es valoren els costos de les alternatives avaluades i la decisió únicament té en compte les diferències en aquests costos. El criteri clau per poder emprar aquesta tècnica és demostrar o justificar que els resultats són idèntics entre les alternatives avaluades, ja que aquesta situació no és molt freqüent. Recordem que una anàlisi que només valori els costos de diferents alternatives no és una avaluació econòmica completa, llevat que es justifiqui la igualtat en resultats.

6.4.2. Anàlisi cost-conseqüències (ACC)

En aquest tipus d'estudis, els costos i les conseqüències de les diferents alternatives es presenten de manera desagregada, utilitzant més d'una mesura de resultat, i es deixa al destinatari de l'anàlisi la valoració de cadascun dels resultats. Per exemple, respecte a un programa de prevenció de consum de drogues en adolescents, una anàlisi cost-conseqüències presentaria el cost d'aquest programa i de l'alternativa (no fer res o els serveis del ajuntament existents, per exemple) juntament amb els resultats de les dues alternatives segons diferents mesures: per exemple, nombre de joves amb problemes d'addiccions, taxes de graduació en ESO, o nombre de joves que cometen algun tipus de delictes. Per tant en una avaluació econòmica que utilitza la tècnica de l'anàlisi cost-conseqüències, llevat que tant els costos com totes les mesures de resultats utilitzades siguin favorables a una alternativa (menors costos i millors resultats), l'usuari de l'estudi ha de ponderar les diferències entre les distintes mesures de resultats i entre els costos per poder prendre una decisió.

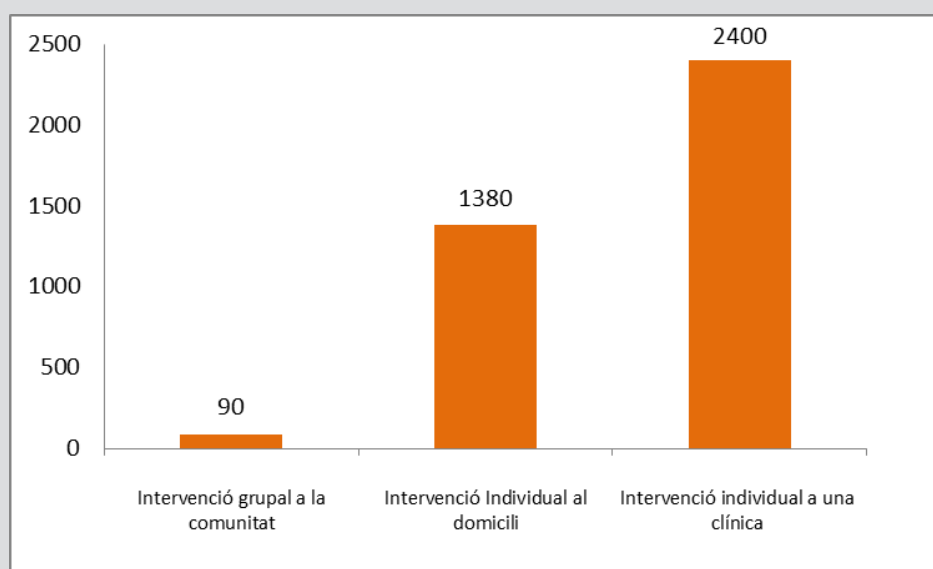
6.4.3. Anàlisi cost-efectivitat (ACE)

En l'enfocament cost-efectivitat, els resultats es mesuren en unitats d'efectivitat no monetàries equivalents per a totes les alternatives. L'objectiu final d'aquest tipus d'avaluació és determinar una **ràtio cost-efectivitat que expressi el cost addicional per unitat de resultat addicional obtingut mitjançant el programa avaluat** (comparat amb l'alternativa). És a dir, les ràtios o proporcions cost-efectivitat s'han de calcular de manera incremental, això és, com el quocient entre la diferència en els costos del programa i la diferència en els resultats.

REQUADRE 21: ANÀLISI COST-EFECTIVITAT DELS PROGRAMES DE "PARENTING"

Per exemple, el gràfic 3 mostra els resultats d'una avaluació econòmica mitjançant el mètode d'ACE de diferents programes de millora de les habilitats parentals ("parenting") que tenen l'objectiu de prevenir problemes de conducta en nens (McCabe, C i altres 2005). En aquesta anàlisi es van comparar cadascuna de les intervencions respecte a l'alternativa del status quo, i els resultats o impacte de les intervencions s'expressen en nombre de nens amb millores de conducta. D'aquesta manera, l'ACE ofereix informació del cost addicional necessari per millorar la conducta d'un nen problemàtic segons els diferents programes o intervencions.

Gràfic 3. Cost addicional (£) per millorar problemes de conducta en infants (de 3 a 8 anys) mitjançant programes de "parenting"



Font: Elaboració pròpia a partir de McCabe et al. (2005)

El gràfic anterior mostra un dels avantatges d'aquest mètode; permet comparar avaluacions econòmiques de programes sempre que els resultats s'expressin en la mateixa unitat, fins i tot si els programes són molt diferents. En aquest cas, la intervenció grupal a la comunitat és el més eficient atès que té la menor ràtio cost-efectivitat dels tres. És a dir, es requereixen menys recursos (addicionals) per obtenir una unitat de resultat (addicional), en aquest cas un nen sense problemes de conducta.

L'ACE presenta algunes limitacions a l'hora de decidir com assignar recursos. En primer lloc, la ratio cost-efectivitat incremental fa difícil pronunciar-se sobre quan "val la pena" aplicar un programa nou. Vèiem en l'exemple anterior com la intervenció grupal a la comunitat resultava l'actuació més eficient de les tres considerades, en el sentit que requereix menys recursos addicionals per nen sense problemes de conducta. No obstant això, l'ACE no ofereix informació sobre si val la pena destinar 90 lliures esterlines per evitar que un nen tingui problemes de comportament.

Una segona limitació d'aquest tipus d'anàlisi la trobem en la reducció dels resultats del programa a una única unitat mesura d'efectivitat. Així, per exemple, un programa de formació ocupacional a presos es pot pensar que pot tenir impacte tant sobre les taxes d'atur com sobre les de reincidència. Tanmateix, una anàlisi cost-efectivitat només podrà utilitzar una de les mesures.

En tercer lloc, la tècnica de l'ACE no permet comparar intervencions que tinguin objectius diferents i que, per tant, mesuri els seus impactes en base a unitats diferents.

Un cas particular de la tècnica d'ACE, l'**anàlisi cost-utilitat**, encara que en alguns manuals es considera una tècnica, es va desenvolupar per superar les dues últimes limitacions esmentades de les ACE, especialment en l'àmbit sanitari. En una anàlisi cost-utilitat, els resultats de les intervencions sanitàries es mesuren utilitzant sempre la mateixa unitat de mesura, que es basa a quantificar la variació en utilitat individual relacionada amb la salut que les intervencions proporcionen. Per tant, es tracta d'una tècnica que podria aplicar-se en l'avaluació econòmica d'una part important dels programes sociosanitaris que actualment desenvolupen les entitats del tercer sector.

6.4.4. Anàlisi cost-benefici (ACB)

En aquest tipus d'avaluacions econòmiques, tant els resultats dels programes com l'ús de recursos que aquests impliquen es valoren en termes monetaris. Així, **els diferents resultats de les intervencions es monetitzen mitjançant l'assignació d'un valor monetari** (en euros, dòlars, etc.) i s'agreguen per obtenir el valor monetari total dels resultats o impacte de les intervencions, que rep el nom de beneficis. Aquesta quantitat, com està denominada en les mateixes unitats de mesura, es compara directament amb la valoració de l'impacte de les intervencions sobre l'ús de recursos (els costos). Aquesta comparació directa de beneficis i costos permet determinar clarament quan una intervenció és eficient (sempre valorant-la respecte d'un altre programa o el status quo) i ho serà quan els beneficis siguin superiors als costos, és a dir, quan el valor net sigui positiu. És important recalcar la necessitat que en les ACB (com en totes les avaluacions econòmiques) s'analitzin com a mínim dues alternatives.

Un dels principals atractius de la tècnica de l'ACB és que permet fer comparacions entre projectes de naturalesa molt diversa, ja que homogeneïtza la unitat de mesura dels resultats en valors monetaris. Així mateix, aquesta homogeneïtzació en unitats monetàries, les mateixes que s'usen per valorar l'ús de recursos que impliquen les diferents alternatives, **facilita la comparació directa entre costos i resultats**, fet que permet obtenir conclusions inequívokes sobre la conveniència d'adoptar una política o posar en marxa un programa nou. Tanmateix, la necessitat d'assignar un valor monetari als diferents resultats és la crítica principal que rep aquesta tècnica, atesa la dificultat d'aquest requisit.

REQUADRE 22: ANÀLISI COST-BENEFICI D'UN PROGRAMA DE REINSERCIÓ LABORAL

Un programa de reinserció laboral per a persones que surten de la presó pot tenir un impacte positiu sobre la seva ocupació, però també sobre les seves taxes de reincidència en el delictes. En una anàlisi cost-benefici d'aquest programa el primer que caldria fer es comptabilitzar l'impacte sobre els seus costos: la diferència entre els recursos necessaris per implementar el nou programa (formadors, pràctiques, acompanyament i seguiment posteriors, etc.) i els recursos que aquests individus haguessin requerit per part dels serveis d'inserció laboral destinats a la població general. Tot seguit, es monetitzarien els resultats del programa, és a dir, s'assignaria un valor monetari tant a l'augment de la inserció laboral com a la reducció de les taxes de reincidència.

Per realitzar la monetització dels resultats es poden considerar els beneficis (o costos evitats) tangibles i intangibles. Per exemple, la reducció de les taxes de reincidència i, per tant, del nombre de delictes, implicarà un cost menor per al sistema judicial i de presons, un cost menor directe per a les víctimes dels delictes i una major productivitat laboral per a les víctimes i per als delinqüents. Així mateix, una reducció dels delictes tindrà també uns resultats intangibles (com ara menys sofriment per part de les víctimes), que s'hauran també d'incorporar, una vegada valorats monetàriament, com a beneficis.

Taula 8. Exemple d'una ACB (dades no reals)

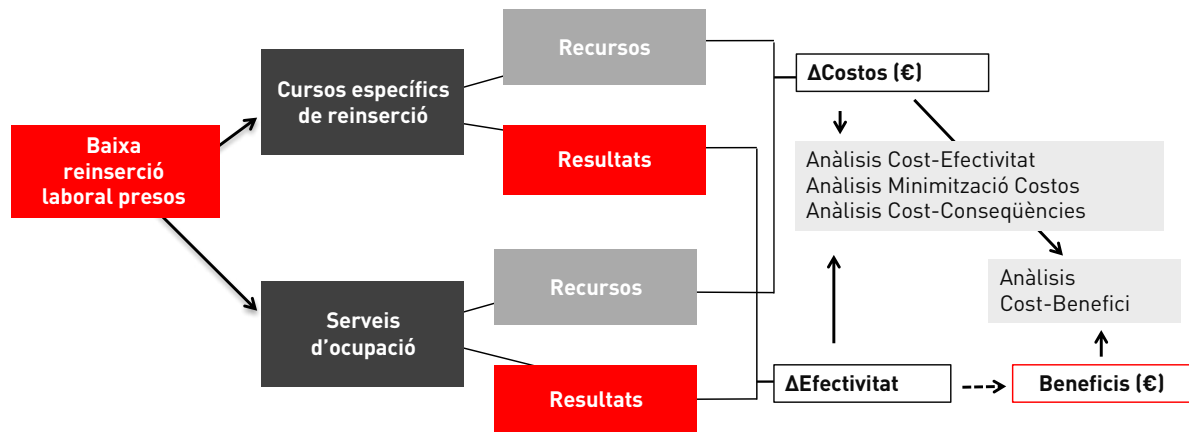
INTERVENCIÓ	COSTOS	BENEFICIS (VALORACIÓ MONETÀRIA DELS RESULTATS)	VALOR NET
Programa de reinserció laboral per a persones que surten de la presó	32.000 €	40.000 €	10.000 €
Serveis d'ocupació	12.000 €	10.000 €	
Diferència	20.000 €	30.000 €	

Font: Elaboració pròpia

6.4.5. Recapitulant: síntesi de tècniques

La figura 12 ofereix una síntesi dels diferents mètodes d'avaluació econòmica exposats amb anterioritat.

Figura 12. Tècniques d'avaluació econòmica: exemple intervencions laborals



Font: Elaboració pròpia

És necessari subratllar que, de vegades, es presenten com a avaluacions econòmiques estudis que en realitat, malgrat tenir un component o enfocament econòmic, no són avaluacions econòmiques. Per exemple, una anàlisi de costos de diferents programes en què no es justifiqui que els resultats dels programes siguin els mateixos, no és una avaluació econòmica. De la mateixa manera tampoc ho és una anàlisi financera que valori els ingressos i costos pressupostaris durant la vida d'un projecte.

REQUADRE 23: EL SROI I LES SEVES LIMITACIONS

La metodologia del SROI (Social Return of Investment) es troba relativament estesa entre les organitzacions del tercer sector, i va sorgir justament amb la intenció de documentar el valor que l'acció del sector aportava a la societat (Cabinet Office, 2009).

El SROI es pot definir com un mètode per mesurar i comunicar el valor creat per les activitats de les organitzacions, tenint en compte els impactes socials, ambientals i econòmics de les citades activitats. De forma molt esquemàtica, un SROI considera la inversió necessària per realitzar un projecte o programa i obté una valoració monetària de tot tipus de resultats i impacte del projecte.

La diferència respecte a una anàlisi cost-benefici és el mètode per determinar quins resultats són rellevants i la magnitud dels mateixos. En el mètode SROI es realitza un procés participatiu amb els actors i agents rellevants amb l'objectiu de determinar quins són els outputs rellevants del programa i als quals, posteriorment, se'ls assignarà un valor monetari. A més, a la pràctica, els mètodes utilitzats habitualment per determinar l'impacte del programa avaluat solen tenir problemes de robustesa, per exemple utilitzant en molts casos valors basats en opinions d'experts. També el fet de donar rellevància a l'opinió dels diferents agents respecte als resultats a considerar i per tant valorar, fa que es produeixen de vegades errors de doble comptabilització. Aquests errors es deuen al fet que un mateix impacte o resultat és considerat, de vegades amb diferents denominacions, per més d'un agent i per tant és valorat monetàriament dues vegades.

Aquest conjunt de limitacions obliga a considerar amb molta cautela els resultats d'anàlisis realitzades mitjançant el SROI, un mètode que no seria pròpiament una avaluació econòmica. I malgrat tot, si una anàlisi SROI d'un programa es basa en mètodes d'estimació d'impacte robusts, evita els problemes de doble comptabilització i emprà mètodes de valoració monetària dels resultats adequats, els seus resultats no haurien de diferir d'una anàlisi cost-benefici del mateix programa.

Font: *Elaboració pròpia*

REQUADRE 24: VALORSOCIAL.COOP I LES SEVES LIMITACIONS

Valorsocial.coop és una eina online de la Federació de Cooperatives de Treball de Catalunya que vol posar en valor l'aportació que les cooperatives de treball realitzen a la societat. L'eina consisteix en un qüestionari, adreçat a les cooperatives, que busca mesurar i valorar les seves activitats i productes així com l'evolució d'alguns outcomes, com la percepció dels usuaris sobre la seva qualitat de vida. Així, l'eina permet fer un seguiment d'aquesta informació a nivell individual, d'una cooperativa, i agregat, pel conjunt del sector.

Cal puntualitzar que aquestes mesures no permeten extreure conclusions de tipus causal, és a dir, per conèixer l'impacte de les cooperatives: l'absència d'un contrafactual ens impedeix saber què hauria passat en la seva absència i comparar-ho amb la realitat. Tanmateix, l'eina suposa un avenç respecte la valoració dels impactes econòmics de l'activitat de les cooperatives, si bé els indicadors suggerits pateixen certes limitacions metodològiques i tenen un enfocament que no abasta totes les dimensions rellevants, per la qual cosa l'eina no permet avaluar l'eficiència com ho faria una avaluació econòmica.

Font: *Elaboració pròpia*

6.5. QUINA ÉS LA UTILITAT DE LES AVALUACIONS ECONÒMIQUES PER A LES ENTITATS DEL TERCER SECTOR SOCIAL?

La utilitat principal de les avaluacions econòmiques de programes és ajudar a la presa de decisions proporcionant informació sobre quin és el millor ús dels recursos disponibles. En altres paraules, davant d'un problema o necessitat ofereixen informació als decisors sobre l'eficiència de les diferents alternatives disponibles o suggerides per afrontar-lo (Pritchard i Svistak, 2014).

A banda d'aquesta utilitat general, l'avaluació econòmica pot aportar altres beneficis a les entitats del tercer sector i al desenvolupament de les seves activitats. Per exemple:

- **Rebre *feedback*.** Tenir informació sobre els resultats i els costos de les activitats de les organitzacions, a través dels resultats de les avaluacions econòmiques, proporciona beneficis per als integrants de les organitzacions així com per a altres agents involucrats en el seu funcionament.
- **Agregar l'impacte dels programes en un únic nombre.** En els mètodes d'avaluació més utilitzats, l'anàlisi cost-efectivitat i l'anàlisi cost-benefici, el conjunt dels efectes del programa avaluat es compila en un únic nombre. Això és útil per a les organitzacions que duen a terme programes amb efectes en diferents dimensions. Per exemple, un programa per prevenir el consum de drogues en adolescents pot tenir efectes en múltiples dimensions: salut, criminalitat, i assoliments acadèmics. Una avaluació econòmica mitjançant un ACB permet agregar aquests impactes en un únic valor.
- **Produir un resultat que és de fàcil comprensió per a un públic ampli.** Aquest és especialment el cas dels resultats de l'ACB, que, com hem vist, es presenten en termes monetaris.
- **Incrementa el coneixement sobre com el programa genera valor.** Dur a terme una avaluació econòmica ajuda a visualitzar i explicitar com les activitats realitzades generen valor i aporten beneficis als diferents agents involucrats i a la societat en el seu conjunt.
- **Ajuda a trobar finançament per a programes i projectes.** Realitzar o encarregar avaluacions econòmiques dels seus programes pot contribuir a incrementar la capacitat de les organitzacions d'aconseguir fons per a les seves activitats.

6.6. RECURSOS

A continuació es llista una relació de recursos disponibles a Internet que poden resultar útils per a l'avaluació econòmica d'intervencions en l'àmbit de les entitats del tercer sector social:

Guies i documents:

- Casado D. Sabes-Figuera R. *L'aplicació de l'avaluació econòmica en l'àmbit dels serveis socials. Una guia introductòria*. Barcelona: Ivàlua, 2016. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 13).
- Casado D. Tarrach A. *L'avaluació econòmica de polítiques educatives. Una guia introductòria*. Barcelona: Ivàlua, 2015. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 12)
- Sabes-Figuera R. *L'avaluació econòmica en l'àmbit de les polítiques de justícia. Una guia introductòria*. Barcelona: Ivàlua, 2016. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 14)
- Sefton T. Byford, S. McDaid, D. Hills, J. Knapp, M. *Making the most of it. Economic evaluation in the social welfare field*. York: Joseph Rowntree Foundation, 2002.
<https://www.jrf.org.uk/report/making-most-it-economic-evaluation-social-welfare-field>
- Svistak, M. Pritchard D. *Economic evaluation: what is it good for? A guide for deciding whether to conduct an economic evaluation*. NPC, UK 2014.
<http://www.thinknpc.org/publications/economic-analysis/>

Repositoris i institucions:

- Center for Benefit-Cost Studies of Education <http://cbcse.org/>
- The Cost-Benefit Knowledge Bank for Criminal Justice <http://cbkb.org/>
- Washington State Institute for Public Policy <http://www.wsipp.wa.gov/BenefitCost>

6.7. REFERÈNCIES

Cabinet Office. *A guide to social return on investment*. London, UK: Cabinet Office, 2009.

McCabe, C. Sutcliffe, P. Kalthenthaler, E. *Parent Training Programs in the Management of Conduct Disorder*. Sheffield: NICE Decision Support Unit, 2005.

Raya, J.M. Moreno, I. *Introducció a l'avaluació econòmica*. Barcelona: Ivàlua, 2013 (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 9).

Sabes-Figuera, R. *L'avaluació econòmica en l'àmbit de les polítiques de justícia Una guia introductòria*. Barcelona: Ivàlua, 2016. (Col·lecció Ivàlua de Guies pràctiques sobre avaluació de polítiques públiques; 14).

Svistak, M. Pritchard D. *Economic evaluation: what is it good for? A guide for deciding whether to conduct an economic evaluation*. UK: NPC, 2014.

7. CONCLUSIÓ

7.1. RECAPITULACIÓ: PREGUNTES, MOMENTS I MÈTODES D'AVALUACIÓ

Ens els capítols anteriors ens hem ocupat de descriure els principals conceptes, procediments i aplicacions relatius als diferents tipus d'avaluació que es poden arribar a plantejar les entitats del tercer sector social. Com dèiem a l'inici d'aquesta guia, l'objectiu d'aquest recorregut no és altre que el de proveir marc conceptual i metodològic comú i bàsic que contribueixi a motivar i orientar la pràctica avaluativa dins el tercer sector.

I partíem també de la idea que avaluar vol dir, principalment, formular-se preguntes sobre la intervenció en qüestió i tractar de respondre-les. En aquesta guia hem parlat d'avaluació de necessitats, del disseny, de la implementació, d'impacte i econòmica. Que, com a entitat social interessada a fer-se preguntes sobre les seves intervencions, ens orientem cap a un tipus d'avaluació o un altre, dependrà, en primer lloc, de **quines preguntes volem respondre**. Ens interessa actualitzar el diagnòstic de necessitats o la magnitud de la problemàtica social que atenem o volem conèixer el nivell de cobertura de la nostra acció? Ens interessa dotar d'evidència el disseny de les activitats, o mesurar l'impacte sobre les persones que hi participen? Ens interessa conèixer els costos de la intervenció o comparar la relació entre costos i resultats de diferents alternatives d'intervenció?

És ben probable que la resposta als interrogants anteriors sigui un genèric "ens interessa tot, volem saber-ho tot". Tanmateix, difícilment ens trobarem en disposició d'intentar saber-ho tot, per motius diversos. Alguns d'ells tenen a veure amb el fet que l'avaluació consumeix recursos, principalment temps i diners; i sabem que els recursos són, sempre, necessàriament limitats.

Més enllà de les preguntes d'interès i dels recursos disponibles, altres motius que ens condicionen a l'hora d'apostar per un tipus d'avaluació o un altre són:

- **El moment del cicle del programa.** En aquest cas distingiríem, bàsicament, entre fase ex-ante del programa (això és, abans de la seva implementació) i fase ex-post (un cop el programa ha començat a implementar-se o ha ja finalitzat). Les preguntes que haurem de prioritzar dependran d'allà on es trobi el programa. Alguns exercicis d'avaluació són de naturalesa principalment ex-ante (avaluació de necessitats i del disseny), mentre que d'altres ho són ex-post (implementació, impacte i avaluació econòmica). Tanmateix, dels capítols anteriors es deriva una consigna general ben clara: com més anticipem la definició d'allò que voldrem saber a les fases inicials del programa, això és, en el moment de dissenyar-lo, més avaluable l'acabarem fent tot al llarg del seu recorregut de desplegament.
- **Les dades i fonts d'informació.** Avaluar implica treballar amb dades, amb informació sobre els participants, sobre les activitats que cursen, sobre el seu context i, si volem

acabar mesurant l'impacte d'un programa, sobre individus no participants comparables als participants. Quina informació disposem necessàriament condicionarà la possibilitat que puguem tenir de respondre unes o altres preguntes, per tant, el tipus d'avaluació que estiguem en disposició d'afrontar. Novament, quan incloem la preocupació per les condicions d'avaluabilitat del programa en el seu moment de disseny, incrementem les probabilitat que tindrem de respondre aquelles preguntes d'avaluació que realment ens interressi afrontar.

- Els **mètodes d'avaluació**. Finalment, cada exercici d'avaluació comporta l'ús d'uns determinats mètodes d'anàlisi. Alguns d'aquest mètodes són de naturalesa qualitativa i impliquen tècniques com les entrevistes i els grups de discussió, d'altres són de tipus quantitatiu i es basen en tècniques estadístiques i econòmiques més o menys sofisticades. En aquesta guia hem fet referència a alguns d'aquests mètodes i tècniques, associant-los als tipus d'avaluació en què acostumen a utilitzar-se. Hem indicat també alguns recursos i bibliografia per a saber més sobre les seves característiques i aplicacions. Tanmateix, és evident que aquesta contribució no resol els dubtes de coneixement i procediment que es poden plantejar a l'hora d'afrontar uns o altres tipus d'anàlisi. L'accés a aquest coneixement o, si més no, a personal expert en aquest camp, és també un factor condicionant del tipus d'avaluació que ens puguem arribar a plantejar. Alhora, ser-ne conscient des del mateix moment en què ens formulem aquelles preguntes que ens interessarà respondre, pot dur a una millor planificació i dimensionament dels recursos que requerirem per a l'avaluació.

A mode de cloenda, la taula 9 sintetitza la manera com es plasma cadascun dels atributs anteriors (preguntes rellevants, moment del cicle del programa, dades requerides i mètodes utilitzats) en l'àmbit de cadascun dels exercicis d'avaluació que hem descrit en els capítols anteriors.

Taula 9. Trets principals dels diferents tipus d'avaluació

TIPUS D'AVALUACIÓ	PREGUNTES QUE VOL RESPONDRE	MOMENT DEL CICLE DEL PROGRAMA	DADES REQUERIDES	MÈTODES DE REFERÈNCIA
Avaluació de necessitats	<ul style="list-style-type: none"> • Quina és la magnitud, la distribució, les causes i les conseqüències de la problemàtica social? • Quins tipus de serveis haurien d'utilitzar les persones que pateixen la problemàtica? • Com haurien de ser aquests serveis i quin n'hauria de ser l'ús? 	<ul style="list-style-type: none"> • Es tracta d'una anàlisi ex-ante, previ al disseny i implementació del programa. D'aquesta manera pot contribuir a definir i orientar els seus continguts. • Pot ser també un exercici continu, en tant que la realitat social, així com els programes que s'hi adrecen, canvien amb el temps. 	<ul style="list-style-type: none"> • Dades sobre la magnitud i distribució de la problemàtica. • Dades demogràfiques i socioeconòmiques de les persones que pateixen la problemàtica. • Dades sobre les característiques i cobertura dels serveis existents. 	<ul style="list-style-type: none"> • Anàlisi de registres administratius. • Disseny i explotació d'enquestes. • Entrevistes i grups de discussió amb persones que pateixen la problemàtica i/o responsables de determinats serveis.
Avaluació del disseny	<ul style="list-style-type: none"> • Quina és la teoria del canvi del programa? Per què s'espera que funcioni correctament i tingui impacte? • El programa reuneix les condicions necessàries d'avaluabilitat? Els seus objectius, indicadors i sistemes d'informació estan definits en aquesta direcció? 	<ul style="list-style-type: none"> • És recomanable que les preguntes sobre el disseny es formulin i es responguin ex-ante, d'aquesta manera augmenten les seves perspectives d'èxit (implementació i impacte) i avaluabilitat. • De forma ex-post, pot induir millores en el disseny de programes ja implementats. 	<ul style="list-style-type: none"> • Ex-ante, es fa referència a les característiques formals dels indicadors operatius i d'impacte del programa (han de ser vàlids, mesurables, sensibles, comprensibles i comunicables, contrastats i rellevants). • Ex-post, caldrà material documental i informants que permetin donar compte del disseny original del programa. 	<ul style="list-style-type: none"> • Revisió d'evidències sobre el funcionament i impacte de programes similars a l'avaluat (aproximació "què funciona"). • Altres mètodes: entrevistes i grups de discussió amb persones implicades en el disseny, estudi documental, observació directa reprocessos i activitats.

Avaluació de la implementació	<ul style="list-style-type: none"> • S'ha disposat dels recursos planificats i aquests han estat suficients? S'han realitzats totes les activitats previstes? • El perfil dels usuaris es correspon amb el de la població diana? Quin ha estat el nivell d'ús i completió del programa per part dels usuaris? 	<ul style="list-style-type: none"> • Anàlisi ex-post, posterior (o en paral·lel) a la implementació, orientada a conèixer com el programa funciona a la pràctica i en quina mesura s'ajusta al previst en el disseny. • Les preguntes d'implementació s'han de formular ex-ante, així com les condicions d'avaluabilitat. 	<ul style="list-style-type: none"> • Característiques individuals i socials dels usuaris. • Dades sobre recursos, activitats i productes resultants, que permetin respondre les preguntes d'avaluació escollides. • Aquestes dades poden provenir de registres administratius, d'informants clau o d'instruments de mesura específicament dissenyats per l'avaluació. 	<ul style="list-style-type: none"> • Revisió i anàlisi documental. • Anàlisi de registres del programa • Entrevistes o grups de discussió amb informants clau. • Disseny i explotació d'enquestes (incloent qüestionaris de satisfacció). • Altres: observació directe, opinió d'experts i estudis de casos.
--------------------------------------	---	---	--	---

<p>Avaluació de l'impacte</p>	<ul style="list-style-type: none"> • Ha aconseguit el programa corregir la problemàtica i en quina mesura? • Quins components de la intervenció són més efectius? • Ha tingut els mateixos efectes sobre els diferents col·lectius atesos? 	<ul style="list-style-type: none"> • Anàlisi ex-post, posterior (o en paral·lel) a la implementació, orientada a conèixer l'efectivitat del programa. • Les preguntes d'impacte s'han de formular ex-ante, així com les condicions d'avaluabilitat. 	<ul style="list-style-type: none"> • Dades relatives al disseny i la implementació del programa. • Dades sobre els outcomes rellevants (grup de participants i grup de comparació). • Dades sociodemogràfiques dels participants i controls. • Dades de context que puguin estar condicionant els outcomes. 	<ul style="list-style-type: none"> • Disseny experimental, amb assignació aleatòria de la participació (mètode més robust). • Mètodes no experimentals: <ul style="list-style-type: none"> - Regressió discontinua - Dobles diferències - Emparellament (matching)
<p>Avaluació econòmica</p>	<ul style="list-style-type: none"> • Quin és l'esforç (mesurat en termes de cost) que requereix una intervenció per assolir un determinat nivell de resultats? • Quina és l'eficiència de les alternatives que es presenten per adreçar una mateixa problemàtica? Això és, quina relació exhibeixen en termes de resultats i costos? 	<ul style="list-style-type: none"> • Anàlisi ex-ante, orientada a la presa de decisions sobre programes en fase de disseny (les treballs amb escenaris plausibles). • Anàlisi ex-post, posterior (o en paral·lel) a la implementació, orientada a valorar l'eficiència de les accions. 	<ul style="list-style-type: none"> • Ex-ante, per a la modelització es requereixen evidències d'altres estudis sobre l'impacte i sobre l'ús de recursos. • Ex-post, es requereixen: <ul style="list-style-type: none"> - Dades d'impacte de la intervenció sobre els resultats rellevants - Dades de l'ús de recursos que les alternatives avaluades impliquen - Dades de costos unitaris, per a la valoració monetària dels recursos 	<ul style="list-style-type: none"> • Els mètodes es diferencien en les unitats utilitzades per a la quantificació dels resultats de les diferents alternatives considerades: <ul style="list-style-type: none"> - Anàlisi de minimització de costos (AMC) - Anàlisi cost-conseqüències (ACC) - Anàlisi cost-efectivitat (ACE) - Anàlisi cost-utilitat (ACU) - Anàlisi cost-benefici (ACB)

Font: Elaboració pròpia

7.2. EL PLANTEJAMENT DE L'AVALUACIÓ: CINC PASSOS BÀSICS

Dediquem aquest apartat final a recapitular sobre els punts clau que sustenten el plantejament inicial de l'avaluació de qualsevol programa social. Exposem aquests punts a mode de passos a seguir en la definició d'aquest plantejament, des de la descripció preliminar de la intervenció i la seva teoria del canvi, fins a la redacció d'un pla inicial d'avaluació.

PAS 1. ANÀLISI PRELIMINAR DE LA INTERVENCIÓ

Descriure la teoria del canvi

Es tractaria de revisar quines respostes caldria donar a interrogants com ara:

- Quins recursos mobilitza o té previst mobilitzar el programa? Són adequats i suficients per dur a terme la intervenció? (ens preguntem pels inputs)
- Quines activitats desenvolupa el programa (o té previst desenvolupar)?
- Quin és o s'espera que sigui l'abast d'aquestes activitats en termes de població atesa, hores de formació, nombre de seguiments, etc.? (ens preguntem pels outputs)
- Quins objectius estratègics persegueix el programa i per què esperem assolir-los mitjançant les activitats i productes del programa? (ens preguntem pels impactes)

Identificar els destinataris principals de l'avaluació

En aquest punt, seria convenient disposar d'una breu descripció de la **població diana teòrica** a la qual s'adreça el programa i, en cas que el programa estigui en marxa, de la **població real**, efectivament atesa.

Identificar la fase en què es troba el programa

Cal clarificar el moment o fase en què es troba el programa, bàsicament:

- Planificació
- Primeres fases d'implementació
- Fase madura
- Finalitzat

Identificar la fase en què es troba el programa

De cara a reforçar la teoria del canvi del programa i incrementar les seves probabilitats d'èxit en la seva implementació i impacte, és convenient seleccionar i revisar estudis o informes d'avaluació de programes similars i/o que adrecin la mateixa problemàtica. Aquesta selecció pot obeir a diferents nivells de rellevància:

- Proximitat a la problemàtica en qüestió
- Proximitat al tipus d'intervenció en qüestió
- Proximitat geogràfica
- Rigorositat metodològica

PAS 2. IDENTIFICACIÓ DEL PROPÒSIT DE L'AVALUACIÓ

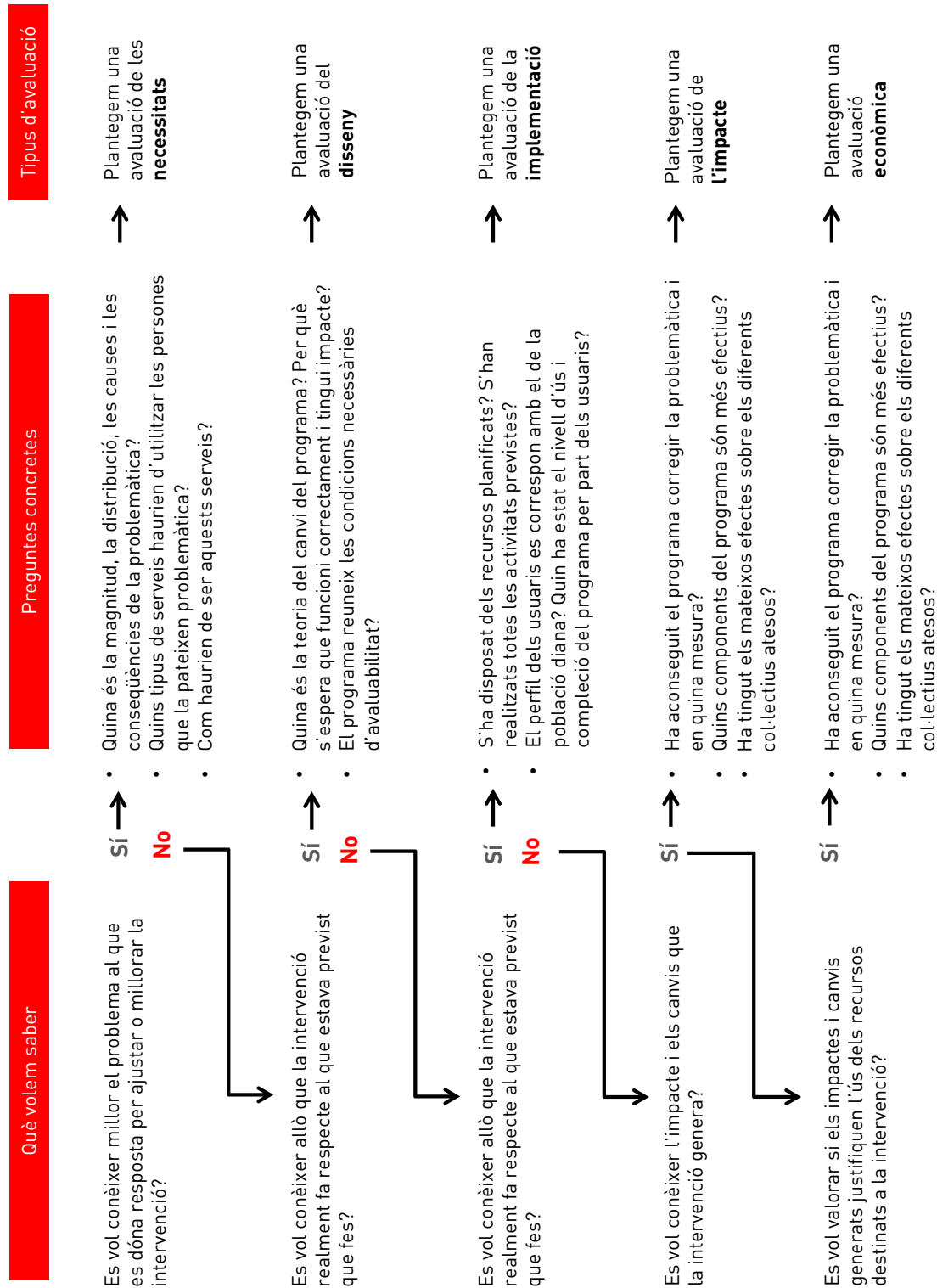
Bàsicament es tractaria de clarificar i prioritzar els objectius reals que es persegueixen encarregant o duent a terme l'avaluació. Aquests objectius poden ser diversos i no excloents els uns dels altres. Tanmateix convindria diferenciar aquells que són prioritaris, d'aquells que no ho són. Algunes d'aquests objectius poden ser:

- Ajuda a la presa de decisions
- Contribució a l'aprenentatge de l'organització
- Ús com a instrument de gestió
- Rendició de comptes a la ciutadania o finançadors
- Coneixement bàsic sobre programes socials
- Acompliment d'un requeriment formal (p.ex. contractació pública)

PAS 3. IDENTIFICACIÓ I PRIORITZACIÓ DE PREGUNTES D'AVALUACIÓ (TIPUS D'AVALUACIÓ)

Es tracta de seleccionar aquelles preguntes sobre el programa (sobre la seva formulació i funcionament) a les quals es provarà de donar resposta. La selecció i priorització d'aquestes preguntes donarà lloc al plantejament d'uns exercicis d'avaluació o uns altres. En tot cas, la identificació de les preguntes d'avaluació haurà de tenir en compte, entre d'altres factors, la fase en què es troba el programa. Una forma d'il·lustrar el procediment que es pot seguir en aquest punt és mitjançant l'arbre de decisions representat en la figura 13.

Figura 13. Identificació de les preguntes d'avaluació, arbre de decisions



PAS 4. IDENTIFICAR ELS RECURSOS DE L'AVALUACIÓ

L'objectiu és disposar d'una relació dels **recursos que són necessaris** per dur a terme l'avaluació prevista, i, al mateix temps, d'aquells que dels **que pot arribar a disposar l'entitat** per dur-la endavant. Parlem de recursos del tipus:

- Finançament aproximat: pressupost de l'avaluació.
- Tipus d'expertesa i coneixement: mètodes d'avaluació, anàlisi estadística, redacció d'informes, etc.
- Suport del personal del programa: dotació pròpia de personal i grau de predisposició inicial d'aquestes persones.
- Informació i accés a dades: llista de bases de dades i altres fonts d'informació existents i rellevants per a l'avaluació, considerant el seu grau d'accessibilitat.
- Temps disponibles per a l'avaluació: període previst per a l'avaluació i temps requerit per les diferents tasques i dedicacions relacionades.

PAS 5. PLA INICIAL DE L'AVALUACIÓ

Aquest representa el darrer pas en el procés de formulació de l'avaluació. Serà el moment de **plasmar en un document de base els elements exposats** amb anterioritat: característiques del programa, propòsit de l'avaluació, preguntes i tipus d'avaluació prioritzada, recursos necessaris i disponibles. Aquest document servirà en tant que **pla inicial de l'avaluació**, i serà utilitzat com a base de cara al disseny d'un projecte específic d'avaluació.

ivàlua ✓

Institut Català d'Avaluació
de Polítiques Públiques

Pau Claris, 108
4t 1a
E-08009 Barcelona

T. +34 93 554 53 00
info@ivalua.cat
www.ivalua.cat

Amb el suport de:



Ajuntament de
Barcelona

Economia
Cooperativa,
Social i Solidària
i Consum



Generalitat de Catalunya
**Departament de Treball,
Afers Socials i Famílies**