

JACINT JORDANA CASAJUANA

**ELS LÍMITS DE LA HIPÒTESI DE**  
**RACIONALITAT COM A FONAMENT PER A**  
**LES TEORIES DE L'ACCIÓ.**

**Una recerca sobre les aplicacions de la teoria de jocs a la interpretació de  
l'acció col·lectiva**

Tesi Doctoral dirigida pel Dr. Fèlix Ovejero Lucas

DEPARTAMENT DE SOCIOLOGIA I METODOLOGIA DE LES CIÈNCIES SOCIALS  
FACULTAT DE CIÈNCIES ECONÒMIQUES I EMPRESARIALS  
UNIVERSITAT DE BARCELONA

Barcelona, 1991

## Índex

<b>Introducció .....</b>	<b>7</b>
--------------------------	----------

## **I. PRIMERA PART: RACIONALITAT I INTERACCIÓ.**

### **Capítol 1: Racionalitat, preferències i utilitat.**

<b>1.1. Racionalitat i model d'home a les ciències socials.....</b>	<b>28</b>
1.1.1. Introducció: una visió general del concepte de racionalitat.....	28
1.1.2. L'individualisme metodològic i les teories basades en l'acció racional.....	34
1.1.3. Racionalitat i interès propi: l'«homo oeconomicus».....	42
1.1.4. Limitacions a la racionalitat en les ciències socials.....	48
<b>1.2. Comportament racional i explicacions basades en preferències.....</b>	<b>56</b>
1.2.1. La lògica del model d'home amb comportament racional.....	56
1.2.2. La consistència en la «gestió» de les preferències.....	64
<b>1.3. Risc i racionalitat: la Hipòtesi de la Utilitat Esperada.....</b>	<b>75</b>
1.3.1. Gènesi i evolució del concepte d'utilitat esperada.....	75
1.3.2. La Hipòtesi de la Utilitat Esperada i el seu vincle amb la noció de probabilitat.....	81
1.3.3. Noves discussions de la Hipòtesi de la Utilitat Esperada: els intents de relaxació axiomàtica.....	91
<b>1.3.A. Annex: La formalització de la Hipòtesi de la Utilitat Esperada.....</b>	<b>102</b>

**Capítol 2: Fonaments de la teoria de jocs: Una «gramàtica» per a les teories de l'acció?**

2.1. Les relacions de la teoria de jocs amb la racionalitat.....	106
2.1.1. Una perspectiva sobre els conceptes de la teoria de jocs.....	106
2.1.2. L'evolució de la teoria de jocs en el context de les ciències socials.....	113
2.2. La vinculació entre acció i teoria de jocs.....	121
2.2.1. Ciència social i teoria de jocs.....	121
2.2.2. Evidència psicològica i modelització interactiva de les orientacions d'interès.....	131
2.2.3. Normes i jocs: variables exògenes o preferències endògenes...	139
2.2.4. Problemes en la modelització dels actors a la teoria de jocs....	148
2.2.5. Racionalitat limitada, aprenentatge i recerca de l'equilibri.....	158
2.3. Teoria de jocs, ciència social i acció col.lectiva.....	169

**II. SEGONA PART: ACCIÓ COL.LECTIVA I RACIONALITAT**

**Capítol 3: L'acció col.lectiva interpretada des del paradigma de l'elecció racional.**

3.1. L'anàlisi de l'acció col.lectiva.....	176
3.1.1. Definició de l'acció col.lectiva.....	176
3.1.2. La gènesi del problema de l'acció col.lectiva.....	178

3.2. El model neoclàssic d'acció col.lectiva i les seves limitacions.....	184
3.2.1. Racionalitat econòmica i empelts institucionalistes en la teoria d'Olson sobre l'acció col.lectiva.....	184
3.2.2. Cruïlles teòriques i ambigüitats en l'obra d'Olson.....	197
3.3. Confrontacions per una teoria general de l'acció col.lectiva (1965-1990).....	201
3.3.1. L'aparició de diferents perspectives analítics.....	201
3.3.2. Acció col.lectiva i mobilització de recursos.....	206
3.4. Reconstruccions parcials i crítiques al model teòric d'Olson.....	209
3.4.1. Individu i bé públic.....	209
3.4.2. Interacció, grandària i les temptacions del 'free-rider'.....	214
3.4.3. Bé públic, interacció i càlcul individual.....	221
3.5. Les intervencions de segon ordre en l'acció col.lectiva.....	226
3.5.1. Sancions com a garantia.....	227
3.5.2. Incentius com a foment.....	228
3.5.3. Líders i empresaris com a impulsors de les intervencions de segon ordre.....	231
3.6. Una discussió com a exemple: classe social i lògiques de l'acció col.lectiva.....	235
3.6.1. La tesi dels privilegis estructurals en l'acció col.lectiva.....	235
3.6.2. Les interpretacions sobre l'existència de dues lògiques en l'acció col.lectiva.....	239

<b>Capítol 4: La teoria de jocs com a nucli explicatiu de l'acció col.lectiva.</b>	
4.1. Cap a una caracterització teòrica de l'acció col.lectiva.....	246
4.2. Qüestionant l'aplicació de la teoria de jocs a l'anàlisi de l'acció col.lectiva.....	251
4.2.1. Interdependència i acció col.lectiva.....	251
4.2.2. La teoria de jocs no és la teoria de l'acció col.lectiva.....	255
4.3. Béns públics i lògica interactiva: estratègies de modelització.....	264
4.3.1. No-unicitat en la interpretació de l'acció col.lectiva.....	264
4.3.2. L'aplicació del cas particular d'eleccions binàries entre dos jugadors.....	268
4.3.3. Els models convencionals d'eleccions binàries amb 'N' jugadors.....	275
4.3.4. Aproximacions parcials al cas general de 'N' estratègies i 'N' jugadors.....	289
4.4. Problemes en la dissolució dels supòsits de racionalitat i interès propi a la teoria de l'acció col.lectiva.....	294
4.4.1. La variabilitat de les solucions als problemes de l'acció col.lectiva.....	294
4.4.2. Orientacions d'interès i provisió interactiva de béns públics....	299
4.4.3. Mecanismes de decisió i elecció d'estratègies.....	304
<b>Capítol 5: Conclusions.....</b>	<b>316</b>
<b>Referències bibliogràfiques.....</b>	<b>340</b>

## INTRODUCCIÓ

La noció de racionalitat ha centrat les discussions dels científics socials sobre la definició d'una hipòtesi de comportament aplicable al conjunt dels éssers humans, amb la qual poder realitzar proposicions sobre la societat i els seus processos. Tot i els grans esforços realitzats per refinar conceptualment aquesta hipòtesi -a través de les perspectives teòriques més rigoroses d'aquest segle-, les observacions empíriques han generat greus dubtes sobre l'efectivitat de mantenir-la com a principi explicatiu. Les freqüents constatacions de fallides en el comportament racional semblen suggerir la necessitat d'abandonar la pretensió d'explicar els fenòmens socials emprant el supòsit que els individus són racionals.

Malgrat la consideració anterior, destaca la importància que ha adquirit la hipòtesi de racionalitat en els enfocaments teòrics que elaboren explicacions intencionals. Això és degut a un motiu molt concret: no és possible fer proposicions de caire general sobre accions intencionals a partir de supòsits sobre comportament irracional. La hipòtesi de racionalitat ofereix, en canvi, la possibilitat de suposar la seva pròpia autonomia ontològica. Aquesta hipòtesi dona per assentat que els éssers humans disposen d'una estructura mental racional de forma innata, o pel fet de conviure en societat i fer ús del llenguatge. Cal afegir que també facilita l'adopció de la hipòtesi de racionalitat la coincidència entre els procediments d'anàlisi -basats en propietats lògiques comunament acceptades, és a dir, racionals- i el comportament dels subjectes a analitzar, als quals tanmateix se'ls pressuposa la racionalitat.

Dins un dels grans paradigmes sobre el fonaments de la ciència econòmica, conegut com la teoria econòmica neoclàssica, s'ha desenvolupat una refinada formulació de la hipòtesi de racionalitat per tal d'ésser aplicada a molts tipus de condicions d'elecció en què es trobi un individu, limitada a articular medis per a uns fins prèviament establerts. Aquesta ha estat precisament la formulació que s'ha estès amb més vigor -des de l'àmbit de la teoria econòmica on va néixer- a diferents camps de les ciències socials, destacant especialment la ciència política. La presentació axiomàtica d'aquest paradigma estableix generalment a partir de la

hipòtesi de racionalitat proposicions d'abast prescriptiu o normatiu que alhora volen tenir també capacitat descriptiva. Els resultats no estan exempts de problemes. Introdueixen un cert confusionisme sobre la forma d'explicar plausiblement fenòmens socials, polítics o econòmics, degut a que no totes les proposicions sobre el que ha d'ésser un comportament racional en determinades situacions, han de ser també correctes per explicar els comportaments observats, per molt escaients que siguin aquestes proposicions.

Són estèrils els propòsits d'aquest paradigma, com argumenten molts crítics que justament tan sols comparteixen aquesta opinió? L'envergadura del propòsit que aquest paradigma desenvolupa invita a considerar que, almenys, no es poden formular respostes simples i contundents. De fet, aquesta formulació de racionalitat no és acceptada de forma totalment unívoca. Ni els orígens contemporanis ni les discussions més recents sobre la hipòtesi de la racionalitat a l'economia tenen res a veure amb la simplicitat. Tot i que la simplicitat és un dels objectius més cercats però, degut als propòsits que es desitgen, existeix una considerable polèmica sobre procediments, pressupòsits i plausibilitat.

La tasca que s'emprèn a les pàgines que segueixen neix d'aquests dilemes, plantejats a la ciència econòmica, per tal d'abordar una preocupació sobre els efectes de la seva transposició a altres àmbits de les ciències socials. L'objectiu que es proposa és examinar els intents de construir formulacions teòriques en camps externs al que va donar origen al propi paradigma, dedicant una especial atenció a la ciència política -l'àmbit en que s'han desenvolupat més aplicacions-. La finalitat d'aquest examen és aclarir, més enllà de discussions internes sobre la definició de la racionalitat, els límits a la seva utilització com a hipòtesi de comportament. Però la implementació de limitacions pot fer aparèixer molts tipus de problemes: començant pel manteniment de la capacitat analítica de la hipòtesi, també presenten problemes els criteris de connexió amb altres supòsits de comportament, així com l'establiment dels límits en que la hipòtesi de racionalitat pot predominar.



Per afrontar aquest conjunt de qüestions, donada l'amplitud d'aspectes en què es troben presents, així com l'enorme camp de propostes teòriques existent, l'itinerari que es segueix a les pàgines següents té com a punt essencial l'anàlisi de les elaboracions teòriques existents a l'entorn de la noció d'acció col·lectiva, escollida com un àmbit de discussió que serveix per exemplificar alguns dels problemes generals del paradigma econòmic. Connectant amb un dels problemes centrals de la ciència política, analitzar els problemes de l'acció col·lectiva serveix d'espai de discussió sobre les dificultats que genera l'adopció de la hipòtesi de racionalitat per interpretar fenòmens que no han ocupat tradicionalment a l'anàlisi econòmica. Escollint un terreny aliè a l'economia, es pretén introduir elements de reflexió sobre l'aplicació de la hipòtesi de racionalitat a algunes ciències socials on havien dominat altres perspectives. Des d'una visió interna de les especificitats de la ciència política s'argumenta que, per incorporar elements explicatius d'aquest paradigma, és necessari caracteritzar de forma distinta la rellevància de la hipòtesi de racionalitat.

L'aparició d'interpretacions dels fenòmens polítics basades en l'anàlisi econòmica es va produir de forma decidida a partir de finals dels anys cinquantes, tot i que amb anterioritat ja s'havien produït algunes aportacions. Els autors més destacats d'aquest corrent (A. Downs, J. Buchanan, M. Olson, entre d'altres), amb plantejaments diversos, tenien com a punt de partida comú en les seves obres la hipòtesi de racionalitat en les actuacions dels individus, proposant aplicacions del model d'«homo economicus» en molts tipus de fenòmens no essencialment econòmics. Aquest model d'home afegeix un altre supòsit important a la hipòtesi del comportament racional: que les accions dels individus estan guiades per la recerca del seu propi interès. Des de la lògica de les organitzacions fins a l'estructura de les democràcies, passant per l'anàlisi de les constitucions, són qüestions força diverses les que han estat abordades des d'aquesta òptica, tant a la seva vessant prescriptiva, com a la descriptiva o la comparativa. Entre aquestes, s'ha desenvolupat una interpretació sobre els problemes de l'acció col·lectiva dels individus, que amb uns

principis molt simples, integra explicativament una àmplia diversitat de fenòmens. Aquest enfocament, amb algunes varietats segons les orientacions dels diversos autors que l'han desenvolupat, es conegut, en la seva aplicació al conjunt de les Ciències Socials, amb el nom d'«elecció racional».

Degut a què les noves aportacions a la ciència política presenten una forta dependència d'un cos teòric ja establert, els objectius esmentats no es poden assolir qüestionant directament els instruments de l'anàlisi econòmica. Per aquest motiu l'estructura d'aquest treball presenta dues parts clarament definides. Una primera part, titulada 'Racionalitat i interacció', es dedica a analitzar els problemes generals de l'enfocament, i posa en evidència alguns dels seus punts més discutits. Al primer capítol, dedicat als problemes de la definició de la hipòtesi de racionalitat, s'introdueixen polèmiques - presents i passades - sobre l'elecció en situacions de risc; al segon capítol, dedicat a la teoria de jocs, es destaquen els problemes de diverses modalitats d'integrar l'anàlisi de les eleccions en aquesta teoria formal. La segona part de la tesi, titulada 'Acció col·lectiva i racionalitat', es centra plenament en un terreny tractat per politòlegs i altres científics socials que s'han aproximat a un dels grans temes de la ciència política: l'anàlisi dels problemes de l'acció col·lectiva. Aquesta part també es divideix en dos capítols. El primer analitza com el problema de l'acció col·lectiva ha estat interpretat des del paradigma de l'«elecció racional». El següent capítol formula una interpretació general dels problemes de l'acció col·lectiva en termes de teoria de jocs, on la hipòtesi de racionalitat manté un important paper, però perd la seva hegemonia. Amb aquesta formulació s'argumenta que, en sortir dels seus terrenys tradicionals, l'instrumental de l'anàlisi econòmica no ha de perdre el seu rigor en el tractament de les qüestions substantives, però que el seu entramat teòric s'ha de transformar per integrar-se en el plantejament propi dels problemes de cada camp específic de les ciències socials. Unes conclusions finals serveixen per nous suggeriments, tot reflexionant sobre les respostes parcials que s'han anat discutint al llarg del treball.

Un objectiu complementari de la tesi és suggerir les bases per a una transformació de la interpretació que l'«elecció racional» realitza sobre els problemes de l'acció col·lectiva. Aquest propòsit es desenvolupa paral·lelament a la discussió sobre la hipòtesi de racionalitat a la segona part del present treball. L'objectiu implica revisar la hipòtesi d'aparició de la cooperació entre els individus únicament com a resultat dels interessos per al seu propi benestar, hipòtesi que entén l'actitud de participació en un procés de cooperació com a resultat d'un càlcul que maximitza la relació entre el cost i el benefici personal en la decisió d'efectuar una acció. Davant aquesta concepció, s'argumenta que els interessos individuals poden ésser definits en formes més diverses, no sempre encaminats a maximitzar la utilitat personal, i que els processos de decisió que vinculen preferències amb accions no han de ser sempre guiats necessàriament per la hipòtesi de racionalitat. Però el propòsit no és tant arribar a formular una nova interpretació teòrica, sinó discutir els problemes metodològics presents en l'elaboració de mecanismes explicatius de menor generalitat però de major eficàcia per a l'anàlisi de configuracions socials d'acció col·lectiva força diverses.

La interpretació convencional que proposa l'enfocament de l'«elecció racional» sobre els problemes de l'acció col·lectiva presenta greus anomalies, repetidament detectades per estudis de caràcter empíric. Per resoldre aquestes anomalies, els teòrics de l'«elecció racional» que treballen en l'àmbit de l'acció col·lectiva han seguit dues grans línies d'investigació. La primera, refinant els supòsits sobre la capacitat racional dels individus per comprendre les situacions en què intervenen; la segona, integrant en l'explicació basada en la racionalitat consideracions de caire institucionalista, com poden ésser l'existència de restriccions, sancions, o incentius selectius per a la realització d'una acció col·lectiva.

Davant aquestes dues línies d'investigació, en aquest treball es valoren els efectes de variacions en els supòsits bàsics del comportament individual i dels objectius de l'acció col·lectiva sobre la capacitat explicativa de proposicions

teòriques ja establertes. L'estratègia esmentada examina com és possible integrar i ampliar un conjunt de propostes teòriques sobre racionalitat i acció col·lectiva desvinculades entre sí, que comparteixen la perspectiva de l'individualisme metodològic. És proposen en darrer terme algunes limitacions a aquesta heurística (almenys, pel que fa a l'estat actual de la ciència social), però això no impedeix que la perspectiva de l'acció individual racional no es pugui considerar un camp fèrtil per explicar diversos aspectes dels processos d'acció col·lectiva. Es suggereix que si es diferencia entre mecanismes de decisió, tant procediments de decisió racional com formes no racionals, pot disposar-se d'un major ventall de possibilitats per establir formulacions amb capacitat generalitzadora sobre les formes d'emergència i estabilitat de l'acció col·lectiva.

En aquest sentit, es tracta de proposar un canvi de perspectiva, per establir fins a quin punt és possible explicar amb combinacions de diversos supòsits de comportament la producció de fenòmens d'acció col·lectiva sense haver de recórrer a la introducció d'altres factors explicatius, tals com filtres que precisin les condicions d'acció dels actors.

\*\*\*\*

Després de presentar el paper de la hipòtesi de racionalitat en els models econòmics aplicats a la ciència política, i proposar una interpretació de l'acció col·lectiva aplicant la teoria de jocs, s'introdueix de forma més detallada l'estructura que pren l'argumentació a través dels diversos capítols, indicant els lligams i fils conductors més destacats. En el primer capítol, es dedica una atenció especial als intents de formalitzar la noció de racionalitat individual, sobretot en el seu cas més complex, la teoria de la decisió en les situacions de risc i, al mateix temps, s'aïllen i s'examinen els components bàsics del model d'home que acompanyen la hipòtesi de racionalitat. Oferint una interpretació de l'evolució de les relacions entre els conceptes d'utilitat i de racionalitat a partir de les seves incorporacions als models

formals de la microeconomia, es defensa que la definició d'una teoria formal de la decisió ha tingut constantment la necessitat de treballar amb una mesura de satisfacció -o quelcom similar- per establir escales d'actituds, tot i que contínuament s'ha intentat trencar aquesta vinculació, per evitar una dependència del substrat psicològic que implicava. En canvi, aquesta dependència no comprenia necessàriament una definició concreta de la racionalitat, ni el supòsit d'interès propi, que també s'hi va incorporar. Per entendre el rerafons d'aquesta formulació, el capítol s'introdueix amb una breu referència a la concepció epistèmica de la racionalitat, per seguir posteriorment amb una discussió sobre les relacions entre els elements claus de la noció de racionalitat i els altres components del model d'home conegut com «homo economicus». Aquest model té els seus orígens en una visió de l'individu i la societat sorgida del pensament social dels segles divuit i dinou, en especial de l'utilitarisme, on autors com J. Bentham i J. Mill oferien una complexa visió global de la societat, que posteriorment seria reduïda per l'anàlisi econòmica a un conjunt d'axiomes fonamentals.

A partir dels anys cinquantes d'aquest segle es va produir un refinament dels fonaments axiomàtics de la racionalitat, alhora que es produïa una expansió del conjunt dels pressupòsits formulats per la teoria econòmica cap a altres ciències socials. Així, mentre que tant aquests fonaments com la formulació concreta dels seus models més usuals rebien crítiques dins la pròpia teoria econòmica -crítiques que actualment es continuen produint-, la discussió sobre la racionalitat a la ciència política tendia a ésser més una crítica interparadigmàtica. Aquest supòsit era adoptat sense recollir les formulacions i alternatives que des de l'interior del propi paradigma econòmic s'havien proposat: tant modificacions a la hipòtesi de racionalitat per donar resposta a l'observació d'imperficcions, com interpretacions substancialment distintes de la racionalitat instrumental, entre les que destaca -tot i els seus problemes conceptuals- la noció de racionalitat limitada proposada per Herbert Simon.

Les discussions sobre la utilitat esperada mostren els problemes plantejats en relacionar probabilitat i racionalitat. Aquesta teoria de la utilitat és la peça clau mitjançant la qual s'ha anat bastint una part important de la teoria econòmica formalitzada axiomàticament des dels anys cinquantes, en el moment en què es va considerar necessari incorporar les situacions de risc en l'anàlisi de les accions dels individus. El motiu per realitzar aquesta incorporació era l'escàs realisme del supòsit d'informació completa i, per tant, la necessitat d'incorporar un enfocament en termes probabilístics als models d'acció racional. A partir de les discussions sobre aquesta hipòtesi, i de les fortes crítiques sobre la plausibilitat dels seus axiomes, se n'ha derivat en els darrers anys una àmplia polèmica sobre la aplicació de la teoria de la decisió racional a la teoria econòmica, apareixent moltes propostes i refinaments alternatius.

Abans de discutir l'ús de la utilitat esperada, en un apartat anterior es presenta la noció de preferència en teoria econòmica, exposant els seus límits i la possible utilització d'aquesta noció fora de l'entramat teòric de l'economia. També s'analitza la possibilitat d'establir interpretacions més complexes dels processos de formació de preferències, en relació amb els contextos i els efectes de les decisions que s'han d'afrontar (per exemple, els treballs d'A. Sen o H. Margolis). A l'apartat següent, en analitzar els problemes de l'aplicació de les probabilitats per afrontar les decisions en situacions de risc, es destaquen dos moments claus: primer, les discussions dels anys cinquantes, sobre el principi d'independència i el tipus de probabilitat a considerar (objectiva o subjectiva); i segon, les noves polèmiques dels darrers anys vuitantes, quan s'ha reexaminat profundament la solidesa dels diversos axiomes i s'han posat en dubte molts dels seus postulats -especialment l'axioma d'independència-, tant formalment com des de les crítiques de psicòlegs especialitzats en processos de decisió. Aquesta línia de discussió interessa per dos motius bàsics: primer, per a il·lustrar quines són les dificultats bàsiques que tenen els economistes i els teòrics de la decisió per a construir una modelització més complexa de la racionalitat. Segon, per poder mostrar posteriorment com aquestes dificultats

es transmeten a àmbits teòrics externs a l'economia, en els que no es plantegen habitualment els problemes de fonamentació, encara que aquests hi són de forma subjacent al traslladar-s'hi directament l'estructura interpretativa i les preguntes bàsiques. Això produeix sovint que des de fora del propi paradigma es critiquin els seus supòsits. La crítica però, resta en un enfrontament de principis, ja que no es disposen de punts de referència adequats. El repte és, doncs, com transformar internament l'enfocament per tal que sigui possible afrontar problemes importants d'àmbits externs a l'economia, aprofitant la capacitat del seu instrumental analític.

Les dificultats metodològiques i teòriques plantejades al primer capítol són molt àmplies i, a més, força interdependents. Naturalment, és impossible de realitzar una discussió crítica de tots ells. No és aquesta la intenció del treball. En el primer capítol es plantegen aquests problemes, tot destacant les seves connexions més importants, per seleccionar i accentuar les qüestions que es consideren centrals per l'argumentació de la tesi. Aquesta pretensió té l'objectiu de mostrar que tant la integració del concepte de racionalitat en un model d'home, com la seva construcció de forma prou general per poder bastir teories sobre el comportament agregat a partir d'unes poques hipòtesis bàsiques sobre el seu comportament i les seves característiques no són tasques simples. Al contrari, aquestes tasques són sotmeses a una gran tensió, amb propostes i contrapropostes, amb el conflicte constant entre la precisió de complexitat i la necessitat de simplicitat.

Cal remarcar que el conflicte afecta directament, més enllà de les variacions del model d'home que es presentin, a totes les teories que depenen solament d'uns pocs principis sobre el comportament de l'individu. El problema central és que les transformacions teòriques que impliquen una major flexibilitat dels supòsits bàsics de l'explicació de l'acció col·lectiva, no són suficients per a solucionar algunes de les dificultats interpretatives que mostren els fenòmens d'acció col·lectiva. Aquestes variacions permeten una aproximació més realista a diversos processos de mobilització i cooperació, però aleshores cal renunciar a la generalització anterior

sobre la capacitat d'aplicació dels models resultants. Malgrat tot, aquesta generalització tampoc convenia degut a la seva manca de plausibilitat en la major part dels àmbits en què s'aplicava; i quelcom similar succeiria en el cas d'adoptar com a fixes altres principis de comportament.

El segon capítol discuteix una aplicació formal dels principis sobre racionalitat presentats al primer capítol. La teoria de jocs o interacció estratègica, formalitza els processos d'interacció mútua entre els individus i en proposa una explicació sobre els resultats de la interacció. Generalment es tracta d'una anàlisi de les situacions en què es produeix la interacció entre individus racionals; en altres termes, quan el model d'home exposat en el primer capítol interactua. Des de la perspectiva adoptada, la teoria de jocs constitueix una de les bases instrumentals més importants que es proposen per desenvolupar teories específiques, amb capacitat predictiva mitjançant la definició d'equilibris estables, en diversos camps de les ciències socials. Per a la segona part de la tesi interessa especialment la teoria de jocs com un instrument matemàtic que modela interaccions entre individus, ja que es discuteix si és possible presentar de forma més bàsica els resultats de la lògica de l'acció col·lectiva a través d'aquesta teoria instrumental. La vinculació entre la teoria de jocs i l'anàlisi de l'acció col·lectiva s'ha desenvolupat en la tradició de l'anàlisi econòmica, amb posterioritat a les aportacions inicials de M. Olson, com una derivació de la teoria dels béns públics.

La teoria de jocs es relaciona amb la teoria de la racionalitat mitjançant la seva pròpia fonamentació, la Hipòtesi de la Utilitat Esperada, ja que la seva primera versió fou desenvolupada per J. von Neumann i O. Morgenstern com el mecanisme explicatiu de les eleccions en situacions de desconeixement de les intencions del contrari típiques de la teoria de jocs (en un procés d'interacció simultani, un jugador no sap l'estratègia que elegeix el contrari). Així, hom pot entendre també la teoria de la decisió racional com una simplificació de la teoria de jocs, considerant que es tracta d'un joc contra la natura, malgrat trobar-se força més desenvolupada la



primera que la segona. D'altra banda, la relació és aparent perquè degut a la seva autonomia, la teoria de la decisió racional, amb la noció d' 'utilitat' que té associada, pot ser considerada com un cas particular de la teoria de jocs, en ser possible que les solucions als jocs -l'equilibri d'una interacció- també siguin efectives sense cap consideració de comportament racional. Aquesta gran capacitat de generalització, que es discuteix amb especial atenció als casos que combinen regles i racionalitat, troba utilitat a l'anàlisi dels fenòmens d'acció col·lectiva. Gràcies a la teoria de jocs, és possible introduir i analitzar de forma comparable les diferents variacions de la hipòtesi de racionalitat prèviament discutides, així com els efectes per a la seva capacitat explicativa d'aquests fenòmens.

Per la relació amb els problemes de l'acció col·lectiva, en el segon capítol es tracten bàsicament els jocs no-cooperatius de suma diferent de zero, quan la cooperació pot augmentar el producte conjunt. A partir de la vinculació d'aquests jocs amb les discussions sobre la teoria de la decisió racional i les varietats que aquesta presenta, es realitza una aproximació a algunes de les possibilitats actuals de formalització desenvolupades pels especialistes en teoria de jocs. En el context de la teoria de jocs, les modelitzacions de diversos autors s'entenen com exemples de distintes perspectives de construir teories sobre la realitat social, que van des de reduir algunes de les propietats del model formal de decisió racional, que comporten la definició de nous conceptes d'equilibri, fins a incorporar condicionants derivats de la introducció de supòsits addicionals als models d'home que es formulen. Racionalitat, utilitat, interès propi, maximització, són conceptes bàsics per a la construcció posterior de teories, que, com la de l'acció col·lectiva, es proposen per l'explicació en condicions molt distintes de fenòmens socials. Però el que es planteja és si la rellevància i el paper de cada un d'aquest conceptes s'han de mantenir en formes idèntiques que a la teoria econòmica, o bé si s'han de transformar en apropar-se a altres terrenys de les ciències socials.

La segona part de la tesi desenvolupa el supòsit que els fenòmens d'acció

col.lectiva es poden formalitzar en termes de teoria de jocs, i que aquesta formalització permet explicar a partir d'accions individuals força resultats -a vegades contraintuïtius- dels processos d'acció col.lectiva. Per aconseguir aquesta capacitat d'aplicació general de la teoria de jocs, hom ha de limitar, a través de diversos mecanismes, la utilització de la hipòtesi de racionalitat. A la vegada, però, s'argumenta que la teoria de jocs, aïllada, no té suficient capacitat per establir una teoria general de l'acció col.lectiva suficientment convincent, i que les línies d'investigació per connectar la teoria de jocs amb proposicions teòriques més substantives sobre l'acció col.lectiva, encara presenten greus dificultats de desenvolupament -com es veurà al segon capítol-. Bàsicament es discuteixen les interpretacions sobre l'acció col.lectiva descentralitzada i l'aparició del fenomen de la cooperació; sols de forma puntual s'analitzen algunes especificitats de l'acció col.lectiva que té el seu origen dins organitzacions formals ja establertes. Un cop detallades les característiques i principals components d'aquestes interpretacions, s'analitza l'estructura de les propostes teòriques utilitzant la crítica realitzada en la primera part de la tesi als seus fonaments, posant en evidència les dificultats explicatives existents mitjançant algunes discussions sobre la interpretació de processos específics d'acció col.lectiva.

En el tercer capítol es presenta la interpretació de l'acció col.lectiva que desenvolupa l'«elecció racional». La revelació d'una contradicció entre interessos individuals i interessos col.lectius, formulada per M. Olson durant els anys seixantes, serveix com a punt de partida per a mostrar les limitacions de l'acció col.lectiva a partir de propòsits voluntaristes. El treball destaca que l'obra d'Olson ha aconseguit una gran influència a la ciència política degut a què proposa una síntesi entre elements de caire institucionalista -propers a postulats ja existents en aquesta ciència- i l'adopció de la hipòtesi de racionalitat com a forma de comportament individual invariable. Amb aquest argument, Olson intenta formular una teoria general sobre les associacions d'interessos, i desenvolupa uns principis bàsics sobre el funcionament de l'acció col.lectiva. Malgrat tot, aquest autor deixa

molts més problemes oberts que problemes resolts, ja que el seu intent de síntesi és excessivament simplificador, sense connectar realment els elements dels dos grans paradigmes que utilitza, sinó tan sols agrupant-los per explicar amb un el que no aconsegueix explicar amb l'altre, aplicant el seu conjunt al tema concret de les associacions d'interessos.

Amb posterioritat aquest enfocament sobre els problemes de l'acció col·lectiva s'ha desenvolupat força, oscil·lant entre les perspectives basades gairebé exclusivament en la hipòtesi de racionalitat i les perspectives que donaven prioritat explicativa als elements institucionals, tals com els incentius selectius. Hi ha hagut també diversos intents integradors, però amb pocs resultats nous. En aquest treball es defensa que la teoria del béns públics, ja utilitzada per Olson, pot ser desenvolupada força més àmpliament per tal d'incloure en l'anàlisi de l'acció col·lectiva fenòmens i problemes que la formulació d'Olson no podia recollir. Es pot destacar que a la seva obra, Olson considera que el problema de l'acció col·lectiva es degut a una categoria específica de béns públics, que presenten unes característiques concretes. No obstant això, aportacions posteriors han mostrat que béns públics amb propietats distintes també comporten altres problemes específics d'acció col·lectiva; amb la qual cosa es pot oferir una caracterització més completa dels processos d'acció col·lectiva, que trenca amb la hipòtesi reduccionista d'Olson sobre la incompatibilitat general entre racionalitat col·lectiva i racionalitat individual.

Malgrat aquesta ampliació de la capacitat de l'enfocament de l'«elecció racional» per afrontar els problemes de l'acció col·lectiva, resta sense resoldre el problema bàsic de com integrar els elements explicatius d'origen institucionalista -o intervencions de segon ordre- a la lògica de l'acció dictada per la racionalitat, o inclús per les intencions en general, si s'abandona la hipòtesi de racionalitat. Les discussions al voltant d'aquest problema han estat força àmplies i en part es reflecteixen en diversos apartats d'aquest capítol. L'assumpte és d'una importància

crucial perquè reflecteix plenament el problema existent en molts altres àmbits temàtics, en relació a l'expansió del paradigma econòmic al conjunt de les ciències socials i la possible transformació del paper de la hipòtesi de racionalitat en aquest procés d'expansió.

L'estratègia d'investigació que es segueix a la tesi per aprofundir en aquesta pregunta es desenvolupa al capítol quart seguint un itinerari d'anada i tornada en relació a la primera part de la tesi. El que es proposa és integrar conceptualment el conjunt dels problemes de l'acció col·lectiva mitjançant la teoria de jocs, tant a través d'una interpretació de la teoria dels béns públics en termes de teoria de jocs com integrant en aquesta una teoria de la decisió més àmplia, en la qual la racionalitat apareix com un cas particular. Es tracta de vincular distintes formes de modelització simplificada que permet la teoria de jocs amb les diverses modalitats en què poden presentar-se els fenòmens d'acció col·lectiva, per cobrir tot el ventall d'accions possibles. Malgrat aquestes transformacions, s'argumenta que la teoria de jocs no és una teoria general de l'acció col·lectiva, ni específicament un fonament d'aquesta. En tot cas, el que es pot fer és construir interpretacions de l'acció col·lectiva a partir d'alguns elements que la teoria de jocs pot integrar. La discussió efectuada al segon capítol sobre les possibilitats de construir teoria social a partir de la teoria de jocs, ja permet entendre perquè la teoria de jocs resta limitada quan es volen establir anàlisis de tots els elements importants que intervenen en un procés d'acció col·lectiva. El més remarcable és que pot incorporar un marc conceptual clar i amb capacitat unificadora, on la noció d'equilibri hi juga un paper fonamental.

Discutint sobre les limitacions de la teoria de jocs, s'obren a la vegada perspectives per a una major aplicació d'aquesta teoria formal. En poder ésser la teoria de jocs dissociada de la hipòtesi de racionalitat, i al mateix temps poder fragmentar el procés de decisió en diverses formes de comportament -incloent les no racionals, com el seguiment de convencions o normes-, els models construïts a partir de la teoria de jocs poden servir per analitzar alguns elements comuns a tot

tipus de problemes d'acció col·lectiva. Encara que els resultats que ofereixen aquestes formalitzacions, basats en l'establiment d'equilibris en la interacció, siguin d'aplicació menys general que quan es manté una sola hipòtesi de comportament -i apareixen menys solucions ben determinades-, faciliten l'adopció d'un enfocament unificat per analitzar els processos d'acció col·lectiva. Cal afegir que aquests resultats poden mostrar quin és realment l'estat de les connexions entre les formulacions teòriques no dependents de la hipòtesi de la racionalitat o d'altres comportaments individuals i les derivades de les teories de l'acció que segueixen els eixos del paradigma analitzat.

\*\*\*\*

Finalment, per concloure la introducció, s'ofereixen algunes consideracions sobre el procediment utilitzat en l'elaboració d'aquesta tesi. Molts dels conceptes utilitzats tenen un ampli ventall de significats, relacionats amb els diversos corrents i paradigmes existents a les ciències socials. A les pàgines que segueixen no es pretén realitzar un seguiment interparadigmàtic del significat de cada concepte, analitzant i matisant la interpretació que realitza cada corrent. Existeixen altres concepcions de la racionalitat, però per conveniència, el que s'ha anomenat racionalitat instrumental és la noció que s'analitza. Al llarg de la tesi hom es refereix a aquesta interpretació concreta, que té una gran riquesa de varietats internes, les quals presenten un gran desenvolupament formal gràcies a la seva continuada discussió a la teoria econòmica. A més, és necessari afegir que molts supòsits de conducta (interès propi, altruisme o altres) són també entesos dins el marc conceptual creat per l'enfocament de l'«elecció racional». Aquests supòsits, sense ésser indissolubles de la teoria de la racionalitat, podrien també assolir significats diferents des d'altres perspectives teòriques.

La discussió sobre les teories analitzades es presenta amb un tractament informal. Això és degut a què en fer ús en l'exposició d'una quantitat força gran de

models i propostes en gran part molt recents -sense cap pretensió exhaustiva-, el seu tractament conceptual permet discutir les seves possibles vinculacions, assolint un enfocament global que situa els problemes més importants i els seus fonaments. Procedint d'aquesta manera, s'intenten oferir interpretacions força precises, relacionant conceptes, i analitzant propostes i elaboracions teòriques. En aquest sentit, també s'incorporen en la discussió sobre els problemes teòrics actuals, elements esclaridors sobre la gènesi històrica d'aquests problemes. Tot i que aquest procediment pot augmentar la tensió interpretativa en el tractament d'alguns temes, la manca de context històric dificultaria enormement l'anàlisi d'alguns problemes conceptuals.

Hi ha altres paradigmes en la teoria econòmica, però si no es menciona explícitament, al llarg del text, hom es refereix a economia o teoria econòmica en el sentit del paradigma conegut actualment com 'economia neoclàssica'. Cada paradigma genera una sèrie de preguntes i dificultats bàsiques que molt possiblement no tenen sentit des d'altres enfocaments teòrics, ja que són punts de desenvolupament i controvèrsia sobre la capacitat explicativa de la pròpia perspectiva en relació amb l'evidència empírica -modelada ja per les preguntes específiques del propi enfocament-. En aquest sentit, no és l'objectiu del treball que es presenta realitzar una comparació entre paradigmes: s'analitzen les dificultats existents quan l'expansió del paradigma es confronta amb nous problemes substantius, i es planteja l'ús d'algunes de les formulacions bàsiques de l'enfocament en aquests nous problemes.

Una dedicació tan gran a la fonamentació teòrica de la racionalitat, la utilitat i la interacció -més de la meitat del text- col·loca aquest treball a mig camí entre la metodologia i la teoria social. Això es veu motivat, apart de per l'estratègia de recerca ja exposada, per la concepció prèvia de què sense qüestionar els fonaments i les assumpcions d'una teoria, difícilment es poden entendre les limitacions teòriques internes d'una interpretació i, per tant, investigar sobre les seves

possibilitats de desenvolupament, en relació a les seves necessitats d'aplicació a fenòmens socials i polítics. Cal afegir que tot i que en aquest paradigma els problemes s'acostumen a discutir en termes teòrics molt abstractes, per tenir en la hipòtesi de la racionalitat el seu nucli essencial i en postular una presentació axiològica-deductiva, la consistència interna de les seves formulacions és més metodològica que teòrica. Aquest propòsit no implica que una altra via d'aproximació, menys atenta a la discussió metodològica, no pugui intentar amb èxit una transformació de la teoria de l'acció col·lectiva (o qualsevol altre proposta teòrica), però el que no podrà rellevar amb claredat seran els lligams interns d'aquest paradigma.

\*\*\*\*

Durant la realització d'aquest treball he contret nombroses deutes. Molts amics i amigues, mestres i col·legues, han seguit amb atenció els meus dubtes i vacil·lacions, aportant els seus ànims, consells i suggeriments, sempre d'un ajut inestimable quan hom es proposa una tasca sobre la que no té clar el seu resultat final. A tots ells vull expressar el meu sincer agraïment, així com manifestar que tots els possibles errors del treball són sols de la meua responsabilitat. He d'afegir també el suport rebut de la meua companya Dunia, dels meus pares i dels meus germans. Els seus estímuls han estat sempre insubstituïbles per avançar en totes les fases del desenvolupament de la tesi.

Aquest treball no hagués estat possible sense la inestimable ajuda i comprensió de Rosa Virós, qui en tot moment va seguir la seva evolució, oferint-me l'estímul i la confiança necessaris. A ella, i als companys del Departament de Ciència Política i Dret Constitucional de la Universitat de Barcelona, dec també el suport necessari per a poder elaborar aquesta tesi alleugerit de moltes tasques acadèmiques. També agraeixo als nous companys de la Universitat Pompeu Fabra l'interès mostrat per la finalització de la tesi durant els darrers mesos, així com els interessants comentaris

realitzats per Ferran Requejo i Ricard Zapata sobre diverses versions d'aquest treball.

Les diverses estades realitzades a la Universitat de Bielefeld, facilitades per un ajut de la Fundació Jaume Bofill, han contribuït sens dubte a millorar els possibles mèrits d'aquest treball. La immensa biblioteca especialitzada en ciències socials d'aquesta universitat ha permès la descoberta i el seguiment de línies d'investigació que han acabat formant una part important de l'argumentació que es desenvolupa. No obstant això, l'entorn humà de la Universitat de Bielefeld ha estat també decisiu en l'evolució que ha seguit aquest treball. Apart de les apassionades discussions amb Joachim Milles sobre teoria política i acció col·lectiva, he d'agrair també els comentaris i suggeriments rebuts per C. Offe, F. Scharpf, V. Wellhöner, H. Wiesenenthal i A. Windhoff-Hèritier.

Finalment, vull destacar l'atenció i la dedicació que Félix Ovejero, director de la tesi, ha esmerçat en el seguiment d'aquest treball, tot i la distància física que durant algun temps ens ha separat. Els seus consells i reflexions al llarg de les diverses fases que ha travessat aquest treball han constituït un punt de referència inestimable per seguir avançant.



PRIMERA PART:

**RACIONALITAT I INTERACCIÓ**

**CAPÍTOL PRIMER:**

**Racionalitat, preferències i utilitat**

## 1.1. Racionalitat i model d'home a les ciències socials.

### 1.1.1. Introducció: una visió general del concepte de racionalitat.

El terme 'racionalitat' és una noció que té una àmplia gamma de significats, que van des de les concepcions formals de consistència en les creences i les accions fins a les conceptualitzacions de caire més tradicional sobre la capacitat de judici autònom, la lliure determinació dels fins i la importància dels processos de deliberació en la presa de decisions col·lectives. A les ciències socials es freqüent la utilització d'aquesta noció, com a heurística per a analitzar les accions humanes<sup>1</sup>. Entre les diverses conceptualitzacions que s'han produït sobre la racionalitat, aquest treball es concentra en l'enfocament anomenat «elecció racional». Aquest enfocament, força formalitzat i dependent -per tant- d'uns axiomes, es pot considerar actualment com una derivació d'un dels paradigmes més importants que empenen el concepte de racionalitat a les ciències socials<sup>2</sup>.

---

<sup>1</sup> Una comparació entre diverses concepcions de la racionalitat en l'àmbit de la política, en relació a la presa de decisions col·lectives, es pot trobar a l'article de J. Elster "The Market and the Forum" (a J. Elster i A. Hylland, (Eds.), Foundations of Social Choice Theory, Cambridge University Press, Cambridge, 1986). Vegeu també l'article de J. Johnson, "Habermas on Strategic and Communicative Action", Political Theory, vol. 19, 1991. Com un exemple del paper clau dels processos de deliberació en la formació de decisions col·lectives, entès com una possible alternativa a les explicacions basades en teories de la racionalitat formal, vegeu B. Manin, "On Legitimacy and Political Deliberation", Political Theory, vol. 15, 1987. En aquest sentit, una mostra del debat existent a la ciència política és el treball de C. Offe i U.K. Preuss, "Democratic Institutions and Moral Resources", Arbeitspapier 5/90, Zentrum für Sozialpolitik, Bremen, 1990.

<sup>2</sup> Les aproximacions generals a aquest paradigma d'H. Wiesensthal ("Rational Choice, Ein Überblick über Grundlinien, Theorienfelder und neuere Themenadquisition eines sozialwissenschaftlichen Paradigmas", Zeitschrift für Soziologie, vol. 16, 1987) i J. Elster ("Introduction", a J. Elster, (Ed.) Rational Choice, Oxford University Press, Oxford, 1986) insisteixen que en els darrers decennis s'ha estès al conjunt de les ciències socials, trencant amb les divisions del treball existents tradicionalment.

Seguint la proposta de J. Elster<sup>3</sup>, es poden dividir les concepcions de la racionalitat en dos grans grups: les que formulen teories estrictes de la racionalitat, i les que estableixen teories àmplies sobre la racionalitat. A les teories estrictes no es considera la natura de les creences i dels desitjos, sinó tan sols la consistència lògica entre aquests, així com la consistència de creences i desitjos amb les accions que se'n deriven. Per les teories àmplies de la racionalitat, sense desestimar la consistència que implica la teoria estricta, el que es considera fonamental és que la formació de creences i desitjos sigui també racional. Per tant, cal que els processos d'elecció racional siguin autònoms en la seva formulació de l'aparició de distorsions contextuais o emocionals, que potser podrien també portar als resultats desitjats però no d'una forma completament intencional, i que es derivin de valoracions i judicis que tenen en compte tota l'evidència i el coneixement disponible. Les teories estrictes de la racionalitat sovint es limiten a afirmar l'existència d'un conjunt de connexions entre objectius (desitjos), creences, i medis emprats, per tal de caracteritzar (i explicar) una acció com racional. Es considera que una acció (racional) és aquella que satisfà de la millor manera possible el conjunt de desitjos d'un agent. Les relacions causals o lògiques que es suposen implícites en una decisió racional són que l'acció és deguda als desitjos i a les creences, i que aquestes darreres són el fruit de l'evidència existent<sup>4</sup>.

Aquestes teories també discuteixen específicament sobre objectius, però quan

---

<sup>3</sup> J. Elster, Sour Grapes: studies in the Subversion of Rationality, Cambridge University Press, Cambridge, 1983, pàgs. 1-10 (trad. esp. Uvas Amargas, Península, Barcelona, 1988). Elster pren al seu torn aquesta distinció de l'obra de J. Rawls, A Theory of Justice (Oxford University Press, Oxford, 1972), desenvolupant-la àmpliament.

<sup>4</sup> D. Davidson, Essays on Actions and Events, Clarendon Press, Oxford, 1980; i J. Elster, "The Nature and Scope of Rational-Choice Explanation", a E. LePore i B.P. McLaughlin (Eds.), Actions and Events, Basil Blackwell, 1985. No obstant, cal destacar que la vinculació entre creences i accions com una relació causal nomològica és força discutida entre els filòsofs de l'acció (per exemple, A. Rosenberg, "Obstacles to the Nomological Connection of Reasons and Actions", Philosophy of Social Sciences, vol. 10, 1980). Per una discussió sobre les diferències entre la teoria de l'acció i la teoria de la racionalitat, vegeu M. Patzak, "Rationalmodell und analytische Handlungstheorie", a M. Holler, (Ed.), Homo Oeconomicus III, Verlag N. Leudemann, Munich, 1984. Aquesta qüestió torna a sortir amb posterioritat, en proposar-se una distinció analítica entre la formació de preferències i la realització d'una decisió (vid. apts. 1.2.2. i 4.4.2.).

apareix la qüestió dels objectius finals, mostren dificultats per a analitzar-los només en termes de la noció de racionalitat que defensen, ja que no pot ser exclòs un element d'arbitrarietat en el procés de decisió que s'analitza<sup>5</sup>. Hom pot presentar en termes força definitius aquesta qüestió: *És racional estar viu?, Per què?.* Es podrien fer un munt de racionalitzacions per sustentar aquesta afirmació, però realment, no hi ha cap possibilitat de proveir una resposta substancial: la pregunta no té sentit. No cal insistir en què aquesta ha estat una tensió continuada a la reflexió filosòfica de caire ètic i polític.

En un sentit general, és possible suggerir que totes les teories de la racionalitat estricta, encara que no només aquestes, són bàsicament teories sobre mitjans: quines «regles» cal postular per articular els mitjans, si existeixen prèviament els objectius i les creences. Però aquestes teories tenen sovint associades assumpcions, hipòtesis, o visions ideals -explícites o implícites- sobre la formació de les creences i dels objectius. Hom hi pot aplicar un pressupòsit sobre el raonament per a explicar com es processa la informació, però hi ha la necessitat d'alguna capacitat substantiva per discriminar l'evidència, estigui aquesta basada en les necessitats vitals, en l'elaboració d'ideals o inclús en un sentit lúdic, no materialista, de la vida<sup>6</sup>.

D'una forma extremadament simplificada, es pot apuntar que al llarg del

---

<sup>5</sup> H. Simon (a *Reason in Human Affairs*, Stanford University Press, Stanford, 1983) destaca la qüestió com el problema del pecat original: "*Reasoning processes take symbolic inputs and deliver symbolic outputs. The initial inputs are axioms, themselves not derived by logical but simply induced from empirical observations, or even more simply posited (...) Axioms and inference rules together constitute the fulcrum on which the lever of reasoning rests; but the particular structure of that fulcrum cannot be justified by the methods of reasoning. For an attempt at such a justification would involve us in an infinite regress of logic, each as arbitrary in its foundations as the preceding one*" (pàgs. 5-6).

<sup>6</sup> E. Gellner introdueix la idea del capital cognitiu fix i el capital cognitiu variable, destacant que en l'àmbit del darrer és possible aplicar l'eficàcia instrumental, però que existeix un nombre -creixent- de decisions úniques, fonamentals, no tractables essencialment amb racionalitat instrumental (per exemple, el casament, la mort,..) i que l'intent de «racionalitzar» aquestes decisions implica tractar el capital fix com a variable, amb la qual cosa els càlculs prenen un aire d'irrealitat ("L'animal qui évite les gaffes, ou un faiseaux d'hypothèses", a P. Birnbaum i J. Leca (Eds.), *Sur l'Individualisme*, FNSP, Paris, 1986).

segle vint s'han desenvolupat dues grans interpretacions de la racionalitat en sentit estricta que separen la racionalitat de les accions un cop els fins s'han determinat: (i) la noció de racionalitat plantejada per algunes corrents politològiques i sociològiques, que entenen que racionalitat és escollir els millors mitjans per una finalitat que està sotmesa a un procés de deliberació. En general, aquestes interpretacions intenten apropar-se als processos de definició d'objectius, sense abandonar l'atenció primordial sobre l'acció instrumental, en discutir els problemes de connexió entre els dos plans de decisió<sup>7</sup>. (ii) La noció de racionalitat desenvolupada pels economistes, que generalment considera els fins com a donats, i es proposa analitzar com s'articulen els mitjans. En aquesta noció es poden distingir dos aspectes primordials. Per un costat, la descripció del comportament racional com a teoria de la decisió, d'altre costat, la utilització de la racionalitat com a hipòtesi per a establir prediccions sobre el comportament observat<sup>8</sup>.

L'enfocament de l'«elecció racional» aplicat a la ciència política s'ha desenvolupat especialment a partir de la noció de racionalitat proposada pels economistes, seguint en les seves aproximacions l'evolució de la teoria econòmica.

---

<sup>7</sup> Entre altres autors, aquesta formulació va ser desenvolupada per Max Weber. Aquest, apart de la seva visió de la racionalitat econòmica com a un tipus ideal (*Zweckrationalität*, per exemple, a "Die Grenznutzlehre und das 'psychologische Grundgesetz'", a *Gesammelte Aufsätze zur Wissenschaftslehre*, J.C.B. Mohr, Tubinga, 1968<sup>3</sup>, pàgs. 396 i ss.), destaca la necessitat de suposar l'existència d'una 'llibertat de voluntat' del individu per adoptar la interpretació racional dels seus actes (Cf. *Gesammelte...op.cit.*, 1968<sup>3</sup>, pàgs. 132 i ss.), ja que si la seva voluntat es afectada per factors emocionals -per exemple-, pot distorsionar l'aplicació del model. D'altra banda, Weber suposa que la llibertat de definir els objectius a partir dels valors (*Wertrationalität*) ha de ser gestionada per la capacitat de la raó humana (*Vernunft*) d'articular les normes d'acció, els conflictes i les contradiccions que són presents a la societat -per exemple, els processos de racionalització tècnica- (*Gesammelte...op.cit.*, 1968<sup>3</sup>, pàgs. 500 i ss.). Sobre la noció de racionalitat de Weber vegeu, entre d'altres, els treballs de G. Hartfiel (*Wirtschaftliche und Soziale Rationalität*, Ferdinand Enke, Stuttgart, 1968, pàgs. 181 i ss.) i R. Brubaker (*The Limits of Rationality. Essay on the Social and Moral Thought of Max Weber*, Allen and Unwin, Londres, 1984, pàgs. 49 i ss.). Caldria afegir que la idea de racionalitat limitada proposada per H. Simon s'aproxima a aquesta perspectiva, ja que intenta apropar la noció de racionalitat a les limitades capacitats cognitives dels éssers humans, introduint supòsits conductuals (per exemple, "From Substantive to Procedural Rationality", a J. Latsis, (Ed.), *Method and Appraisal in Economics*, Cambridge University Press, Cambridge, 1976).

<sup>8</sup> Per a una atapeïda síntesi dels problemes d'aquesta interpretació, alguns dels quals són analitzats extensament al llarg de la tesi, vegeu l'article d'A. Sen, "Rational Behavior" (a *The New Palgrave. A Dictionary of Economics*, vol. 4, 1987).

En aquest sentit, al llarg de tot el treball hom es concentra en examinar diversos problemes d'aquesta perspectiva, realitzant només algunes petites incursions en la perspectiva de la racionalitat «sociològica». Deixant a banda possibles polèmiques, el motiu és que a la sociologia i a la ciència política no s'ha produït un procés de construcció teòrica basat en una concepció pròpia de la racionalitat comparable a l'existent a l'anàlisi econòmica. La teoria de la racionalitat en sentit estricte desenvolupada pels economistes s'ha produït de forma refinada mitjançant un instrumental matemàtic, que permet elaborar models que confronten els desitjos individuals amb les diverses variables presents en el procés de decisió, i assignar un resultat com la millor solució. El punt de partida és una ordenació de preferències prèviament definida per l'agent que realitza la decisió, per valorar els possibles resultats finals del procés de decisió. És possible considerar aquesta teoria com un dels intents més importants per excloure del corpus de proposicions teòriques sobre les decisions racionals, tot tipus d'assumpcions substantives sobre objectius, creences i regles. Mantenint les mínimes assumpcions de valor possibles, a partir d'una hipòtesi del comportament individual, es desenvolupen potents teoremes sobre els processos econòmics d'una societat.

Tal com hom pot esperar, els fonaments heurístics de les funcions matemàtiques formulades per aquest enfocament estan basats en conceptes lògics desenvolupats des dels grecs com una modalitat coherent de «bon pensament», que al mateix temps s'apliquen -primer implícitament i posteriorment formalitzats- com a principis de comportament que serveixen per establir els axiomes de la teoria estricta de la racionalitat que es proposa. Aquesta concepció ha estat defensada, entre altres justificacions, suposant que la gent és efectiva en la persecució dels seus objectius i que els axiomes que serveixen de base són intuïtius<sup>9</sup>.

---

<sup>9</sup> Cf. G. Quartone i A. Tversky, "Contrasting Rational and Psychological Analyses of Political Choice", *American Political Science Review*, vol. 82, 1988. Possiblement serveix per il·lustrar aquesta situació una analogia amb les discussions sobre la geometria euclidiana durant el segle XIX. Després de considerar durant uns dos mil anys que els axiomes proposats per Euclides eren els únics possibles per fonamentar un sistema geomètric, degut a que eren els únics que semblaven lògicament sensats, diversos autors varen proposar altres sistemes, igualment consistents, que no tenien aplicacions concretes quan es van formular,

*A la teoria econòmica, la gènesi i popularització de la noció de racionalitat com a principi de comportament es va produir en relació amb la difusió d'un model de l'home optimitzador i egoista. La difusió d'aquest conjunt conceptual es pot vincular amb el rerafons de l'expansió triomfadora del capitalisme com a forma d'organització econòmica, i amb l'aparició de noves institucions polítiques de caire liberal, que pressuposen una separació conceptual entre el públic i el privat<sup>10</sup>. En aquest sentit, cal pensar que aquest concepte es troba estretament vinculat -més enllà de la presumpció d'egoisme estricte- amb l'aparició en el pensament polític de finals del segle XVIII d'aquest nou model d'home individualista. Perquè es suposa, ja des del principi, que l'acció racional -sigui com sigui definida aquesta- comporta també un concepte de mesura per a cada individu, aïlladament, de la seva satisfacció personal i intransferible. Com a principi heurístic, no s'inclouen -no són pensables des de l'enfocament- conceptes que impliquin reconèixer la compatibilitat al mateix temps entre l'interès individual i l'interès dels altres<sup>11</sup>.*

---

però si que ho van ser per la física del segle XX. (vid. M. Kline, Matemáticas. La pérdida de la certidumbre, Siglo XXI, Madrid, 1985, pàgs. 91 i ss.).

<sup>10</sup> A. Domènech discuteix aquesta qüestió: el sorgiment, amb el capitalisme, d'un nou model ideal d'home en la filosofia política, que no té manera d'integrar el públic i el privat d'una forma no antagònica. Aquesta dificultat s'explica com a fruit de la pèrdua -o impossibilitat de recuperació- d'un model ideal d'home clàssic que no generava l'antinòmia esmentada, sinó al contrari, aconseguia dissoldre la pròpia tensió d'una forma especial - a l'anomenada 'tangente àtica', que derivava la defensa de la «cosa pública» de les pròpies activitats privades (Cf. A. Domènech, De la Ética a la Política, Crítica, Barcelona, 1989). En un sentit anàleg, S.K. Kolm apunta les claus del model ideal d'home contemporani contraposant-lo a la concepció budista del món (Cf. S.K. Kolm, "Le bouddhisme et les hommes économiques", Bulletin du MAUSS, vol. 6, 1986). Vegeu també A. Hirschman, Interés privado y acción pública, FCE, Mèxic, 1982.

<sup>11</sup> Aquesta situació comporta una paradoxa important, encara que allunyada de l'objecte de la tesi: Si el nostre instrumental analític està basat en aquests principis esmentats, es podria argumentar que la nostra aproximació a altres principis de comportament substancialment diferents és problemàtica, per ésser relativament inassolible, ja que cal considerar que les noves referències teòriques sempre són transformacions amb origen en el mateix punt de partida, i potser són, per tant, poc exactes, parcials o llunyanes per establir reconstruccions efectives d'altres concepcions de la relació entre la cosa privada i la pública (vid. A. Domènech, De la Ética a la...op. cit., 1989, pàgs. 77 i ss.; i també A. Hirschman, Interés privado y...op.cit., 1982, pàgs. 72 i ss.).



### 1.1.2. L'individualisme metodològic i les teories basades en l'acció racional.

L'individualisme metodològic té com a punt de partida el supòsit que tot fenomen social pot ésser explicat per les accions dels individus que en formen part. En altres termes, que l'explicació dels fenòmens agregats s'ha de proposar a partir de propietats predicables dels individus. El tipus d'explicació que pressuposa l'individualisme metodològic és la intencional, i té com a punt de partida per explicar la societat les intencions dels individus que la componen<sup>12</sup>. Des d'altres perspectives es defensa també la irreductibilitat d'entitats supraindividuals per explicar els fenòmens socials, argumentant que poden tenir igual capacitat teòrica, derivada de l'existència de propietats emergents de la interacció social. Un punt extrem serien les concepcions que defensen l'holisme radical. Aquestes no tracten els individus ni com éssers privats de voluntat, ja que els actors individuals no tenen sentit pel seu enfoc, en suposar que tots els fenòmens socials ja estan determinats per les estructures i els sistemes socials.

Com a principi, l'individualisme metodològic és una heurística que no presenta cap afirmació sobre les intencions i les motivacions dels individus, ni sobre l'existència d'un comportament que segueixi les pautes de la racionalitat, ni sobre la recerca o no de l'interès propi<sup>13</sup>. No obstant això, es pot afirmar que les explicacions individualistes passen sempre per la construcció de models d'home. Aquests models es constitueixen a partir de la definició de diverses característiques

---

<sup>12</sup> Una anàlisi dels tipus d'explicacions presents a les ciències socials, així com de l'adopció del individualisme metodològic es pot trobar a la primera part del llibre de J. Elster, Explaining Technical Change: A Case Study in the Philosophy of Science, Cambridge University Press, Cambridge, 1983 (trad. cast., Gedisa, 1990). Degut a que en aquest treball el principi de l'individualisme metodològic es tan sols un punt de referència per a altres qüestions, només es presenta la seva definició convencional, sense entrar en les discussions contemporànies o en les seves referències clàssiques, que són força polèmiques.

<sup>13</sup> La teoria econòmica és la branca de les ciències socials que ha adoptat més decididament aquest principi. En aquest sentit, com assenyala B. Barry, "there is nothing in the rot assumptions of the economic approach that forces it toward predetermined conclusions about politics and society" (cf. "Methodology versus Ideology. The 'Economic Approach Revisited'", a E. Ostrom, (ed.), Strategies of Political Inquiry, Sage, Beverly Hills, 1982). Vegeu també A. Hadari, "What are Preference Explanations? The Interpretative Core of Economic Modeling", Social Science Quarterly, vol. 68, 1987.

o propietats dels individus. Això es degut a que les modelitzacions del comportament humà han de ser suficientment simples com per poder ésser aplicades en explicar fenòmens socials, polítics o econòmics<sup>14</sup>, però no són necessàries unes premisses determinades sobre interessos o formes de decisió en els models d'home.

Deixant pel següent apartat (1.1.3.) el problema de representar els interessos i desitjos -que es defineixen a través de les preferències-, en aquest apartat és discuteix si es necessari suposar l'existència d'una forma de decisió hegemònica en els models d'home. Específicament, *la qüestió és si cal considerar la racionalitat com el «mecanisme» de decisió exclusiu per escollir entre les preferències. Un segon problema que es presenta en aquest apartat és si les explicacions intencionals poden ser suficients per abordar tots els fenòmens socials, o be si la seva aplicació presenta algunes limitacions.* Els dos problemes constitueixen una referència constant al llarg de la tesi, en confrontar models basats en la racionalitat econòmica amb les necessitats teòriques de la ciència política. És precisament per aportar elements que contribueixin a aclarir-los, que es proposa un determinat itinerari conceptual pel desenvolupament del treball: abans de tractar els problemes de la construcció teòrica a la ciència política amb explicacions intencionals -a través de la lògica de l'acció col·lectiva-, s'analitzen distintes formes de predir el comportament observat mitjançant supòsits de racionalitat individual.

Existeix una important tensió per definir quina ha d'ésser la relació entre una teoria estricta de la racionalitat i l'explicació intencional. Sovint s'argumenta que predicar la racionalitat en el comportament dels individus és difícil d'evadir, perquè hom necessita alguna hipòtesi per analitzar l'acció, i la més generalitzable és alguna

---

<sup>14</sup> R. Boudon, "Individualisme et Holisme dans les Sciences Sociales", a P. Birnbaum i J. Leca (eds.), Sur l'Individualisme, op. cit., 1986. Per contra, es pot mostrar que les teories que no adopten l'individualisme metodològic, generalment les definicions de models d'home són inoperants (Cf. D. Wrong, "The Oversocialized Conception of Man in Modern Sociology", American Sociological Review, vol. 26, 1961; P. Weise, "Homo Oeconomicus und Homo Sociologicus. Die Schreckenmänner der Sozialwissenschaften", Zeitschrift für Soziologie, vol. 18, 1989).

formulació que tingui com a punt de partida la racionalitat<sup>15</sup>. Dins la noció de racionalitat econòmica, hom pot considerar dues formes d'aproximació distintes per la provisió de supòsits de racionalitat aplicables de forma generalitzada. A partir d'unes ordinacions de preferències, es defensa per un costat l'aplicació d'una teoria de la decisió racional que connecti lògicament les preferències amb l'elecció d'una determinada acció; i d'altre costat, es defensa la necessitat de limitar-se a aplicar la racionalitat com a hipòtesi per a predir el comportament observat, considerant l'existència de relacions causals entre preferències i accions<sup>16</sup>. Més endavant s'analitzen els problemes concrets d'aquestes aproximacions, que pressuposen -amb distints graus de plausibilitat- que la racionalitat és aplicable com a premissa bàsica de les explicacions intencionals, i es discuteix també la possibilitat de desenvolupar altres aproximacions que connectin preferències i decisions.

Hi ha una llarga i important polèmica epistemològica a les ciències socials entorn de la necessitat o no necessitat de disposar de pressupòsits plausibles en formular teories sobre conjunts de fenòmens socials. Com s'ha apuntat, l'explicació intencional fa servir el principi de racionalitat per les seves construccions teòriques, però el problema és que amb aquest principi no queda clar l'estatus metodològic de les formulacions que es realitzen. Si allò que hom intenta elaborar són prediccions basades en regularitats de comportament (o lleis), s'ha de defensar que les premisses d'aquestes prediccions han de correspondre amb un grau important de versemblança

---

<sup>15</sup> En aquest sentit defensa la seva aplicació K. Popper, separant la utilització de la racionalitat com a «principi» de la noció de racionalitat com actitud personal, que pressuposaria considerar que els homes adopten sempre un comportament racional: *"Il constitue en réalité un principe minimum (...), qui anime presque tous nos modèles situationnels explicatifs et que, bien que nous sachions qu'il n'est pas vrai, nous considérons avec quelque raison comme une bonne approximation. L'adoption de ce principe réduit considérablement le caractère arbitraire de nos modèles; un caractère arbitraire qui deviendrait un véritable esprit de caprice si nous tentions de construire les modèles en nous passant de Lui"* (Cf. "La rationalité et le statut du principe de rationalité", a E.M. Claassen (ed.), Les Fondements philosophiques des systèmes économiques, Payot, Paris, 1967). Per a una aproximació a aquesta perspectiva, vegeu els articles de S. Giner, "Intenciones humanas y estructuras sociales: Aproximación crítica a la lógica situacional" (Cuadernos Económicos del ICE, vol. 3-4, 1977) i d'A. Gómez Rodríguez, "Acción individual, lógica social" (Zona Abierta, vol. 48-49, 1988).

<sup>16</sup> Vid. A. Hadari, "What are Preference Explanations?...", art. cit., 1987.

a les observacions que es realitzin per la seva comprovació:

*"If by an assumption of a theory we understand one of the theory's fundamental statements (...), a theory with an unrealistic assumption (...) is patently unsatisfactory; for such a theory entails consequences that are incompatible with observed facts, so that on pain of rejecting elementary logical canons the theory must also be rejected"*<sup>17</sup>

Així, degut als problemes empírics de la hipòtesi de racionalitat, alguns autors que volen continuar defensant un mètode de construcció teòrica, mantenent l'argument de què, tot i que les premisses no siguin realistes, els models sí permeten fer prediccions interessants, i justifiquen, per tant, la hipòtesi utilitzada<sup>18</sup>. Però aquest procediment aconduïxen fàcilment a formular models irracionals del comportament racional, cosa que per la seva manca de coherència, cal descartar-ne l'adopció generalitzada. Des d'un altre punt de vista s'assenyala que la manca de plausibilitat del principi de racionalitat pot ser «excusada» defensant que es tracta de la millor aproximació possible a un principi d'aplicació general: L'argument seria adoptar la racionalitat per considerar que és el tipus de comportament més estès<sup>19</sup>. Però la qüestió que sorgeix en aquesta situació és que caldria entendre les

---

<sup>17</sup> E. Nagel, "Assumptions in Economic Theory", American Economic Review. Papers and Proceedings, vol. 53, 1963, pàgs. 214-215.

<sup>18</sup> L'argument clàssic d'aquesta posició és el de M. Friedman, "*Cuanto más significativa sea una teoría, serán menos realistas los supuestos (...) para ser importante, por lo tanto, una hipótesis deberá ser descriptivamente falsa en sus supuestos. [...] La posibilidad de que una teoría sea 'suficientemente' realista solo puede explorarse observando si genera pronósticos lo bastante buenos para el propósito de que se trate.*" (Cf. "La metodología de la economía positiva", a F. Hahn i M. Hollis (eds.), Filosofía y teoría económica, FCE, Mèxic, 1986, pàgs. 58 i 73 -ed. original: M. Friedman, Essays in Positive Economics, Chicago University Press, Chicago, 1953). D'altre costat, com destaca R. Ackermann "*From Friedman's perspective, theories of economic choice involving maximizing utility or maximizing profit are good scientific theories because, under a wide range of circumstances, firms and individuals seem to act as though they were maximizing expected utility or expected profit and were in full control of the relevant information affecting choice*" (cf. "Methodology and Economics", a The Philosophical Forum, vol. 14, 1983, pàg. 390). Vegeu també T. Moe, "On the Scientific Status of Rational Models", American Journal of Political Science, vol. 23, 1979.

<sup>19</sup> Vegeu, entre d'altres, K. Popper, "La rationalité et...", art. cit., 1967.

elaboracions teòriques resultants com a construccions lògiques, més properes a la noció de «tipus ideal» que a la de llei científica. Si s'accepta la consideració, no es podria atorgar excessiva capacitat de predicció a aquestes elaboracions, degut a la impossibilitat de calibrar (amb el propi instrumental teòric) la distància existent amb el conjunt de fenòmens observats empíricament<sup>20</sup>.

Aquestes dificultats són presents als models de la teoria econòmica que utilitzen la hipòtesi de racionalitat, i apareixen igualment a la ciència política que adopta l'enfocament de l'«elecció racional». És per aquesta raó que són discutides en els propers capítols. En els termes plantejats, la discussió entre la utilització de la racionalitat com a «tipus ideal» o com a hipòtesi de comportament apareix com un problema irresoluble, entorn del qual hom tan sols pot escollir en un sentit o altre. Això no obstant, el propòsit d'analitzar els límits de la racionalitat té una estreta relació amb aquest problema. Un cop analitzades al llarg d'aquest capítol les possibilitats de la racionalitat com a principi plausible de comportament, en els següents es valoren les dificultats que apareixerien si s'establissin aproximacions a la ciència política basades en l'individualisme metodològic, que junt amb la racionalitat, adoptessin altres formes d'explicació intencional. Així, *un nou problema que sorgeix és si -almenys com a principi- és necessari integrar un enfocament causal -no derivat directament de les accions- en els models explicatius, que permeti articular la racionalitat amb altres hipòtesis de l'explicació intencional.*

Per discutir el problema de si l'individualisme metodològic pot configurar un

---

<sup>20</sup> Malgrat això, tal com ho defensava M. Weber, aquesta aproximació en termes de «tipus ideal» sí té sentit per a calibrar i analitzar fenòmens concrets, en una determinada constel·lació històrica: *"Für die Forschung will der idealtypische Begriff das Zurechnungsurteil schulen: er ist keine 'Hypothese', aber er will der Hypothesenbildung die Richtung weisen. (...) In seiner begrifflichen Reinheit ist dieses Gedankenbild nirgends in der Wirklichkeit empirisch vorfindbar, es ist eine Utopie, und für die historische Arbeit erwächst die Aufgabe, in jedem einzelnen Falle festzustellen, wie nahe oder wie fern die Wirklichkeit jenem Idealbilde steht,..."* (cf. *Gesammelte...op.cit.*, 1968<sup>3</sup>, pàgs. 190-191). En aquest context també es pot entendre l'aplicació d'una teoria de la racionalitat com un instrument per a realitzar propostes normatives, que permeti disposar d'un marc de referència en l'explicació de les accions humanes (així pren també un altre sentit la discussió per una teoria refinada de la decisió racional, una de les dues formes en que es presenta la racionalitat econòmica).

enfocament unitari d'interpretar la societat -independentment dels principis que s'adoptin-, es pren la recent polèmica sobre el «marxisme analític» com a exemple, i s'il·lustren les contradiccions a les quals està sotmès l'individualisme metodològic quan la seva capacitat analítica s'aplica a explicar processos socials. La tensió que s'apunta seguidament és present al llarg de la tesi, amb l'objectiu d'aportar major claredat sobre algunes de les dificultats existents per desenvolupar aquest enfocament. La disputa sobre marxisme i individualisme metodològic té com a punt de partida la crítica als tradicionals enfocaments marxistes de caire estructural i la reivindicació de la necessitat de formular-los des d'una altra perspectiva<sup>21</sup>. Una línia d'argumentació dels anomenats marxistes analítics, que des dels primers anys vuitantes iniciaren aquesta revisió del marxisme, era la defensa de l'individualisme metodològic associada a la utilització de l'instrumental de teoria de la decisió racional i de la teoria de jocs per explicar institucions i fenòmens socials, afegint en tot cas a l'explicació filtres estructurals amb la finalitat de reduir el marc d'accions possibles pels individus<sup>22</sup>.

La crítica des de la tradició marxista convencional contra aquesta revisió del marxisme es centrava en que l'anàlisi de les estructures productives i socials era més essencial que l'anàlisi de les accions a la major part de les teories d'origen marxista, argumentant que:

*"The role of choice-theoretic methods in Marxist theory is thus necessary limited. The reason is simple -the Marxist problematique revolves around more general structural questions that precede choice-theoretic analyses because these presuppose the structures and relations underlying the manifesto forms of historical social formations and the behavior of*

---

<sup>21</sup> Entre d'altres, vegeu de J. Elster, "Marxism, Functionalism and Game Theory: The Case for Methodological Individualism", *Theory and Society*, vol. 11, 1982; i *Making Sense of Marx*, Cambridge University Press, Cambridge, 1985; també J. Roemer, (Ed.), *Analytical Marxism*, Cambridge University Press, Cambridge, 1985.

<sup>22</sup> J. Elster, "Marxism, Functionalism...", art. cit., 1982.

*individuals within them*"<sup>23</sup>.

Des d'aquestes posicions, no es caracteritzava la teoria de l'«elecció racional» com quelcom completament incompatible amb l'anàlisi marxista, sinó que es destacava que el seu àmbit d'aplicació dins aquesta tradició teòrica era força reduït, indicant que els processos de formació i canvi de preferències quedaven fora de l'esquema formal d'aquesta teoria. S'argumentava que aquesta teoria servia com a instrument per a descriure de forma precisa diverses situacions socials, però no per a explicar com es produïen els possibles canvis estructurals d'aquestes situacions: l'explicació sobre la formació i canvi de preferències havia de derivar-se de referències substantives, com les que ocupaven a les interpretacions teòriques de la tradició marxista. Per a realitzar aquesta integració, es postulava un argument de mútua dependència, que no aconseguia resoldre la qüestió, degut a no aportar cap element per sortir de la tautologia que amenaça tota relació circular:

*"individual and social predicates, like individuals and society themselves, are 'internally related', each provides the conditions of the other's existence"*<sup>24</sup>.

Cal notar que aquesta crítica associa directament individualisme metodològic a racionalitat<sup>25</sup>, quan la major part dels marxistes analítics tan sols defensen la preeminència explicativa de la racionalitat. Per Meiksins, entre d'altres crítics, aquest reduccionisme li és útil per defensar l'existència de categories socials enteses com a entitats supraindividuals, com la utilització del concepte de classe per a

---

<sup>23</sup> J. Weldes, "Marxism and Methodological Individualism", *Theory and Society*, vol. 18, 1989, pàg. 375.

<sup>24</sup> J. Weldes, "Marxism and...", art. cit., 1989, pàg. 362.

<sup>25</sup> Per exemple, en la seva crítica E. Meiksins, destaca que "*The differentia specifica of Rational Choice Marxism is a theory of exploration and class in which various modes of exploitation and classes are logically generated according to the principles of game theoretic analysis*" ("Rational Choice Marxism: Is the Game Worth the Candle?", a *New Left Review*, vol. 177, 1989, pàg. 44).

analitzar el canvi social<sup>26</sup>. Apart de l'objecció esmentada, cal destacar també que aquests raonaments porten a desestimar la rellevància d'una teoria de l'acció, degut a què no consideren la importància explicativa de les accions individuals, i aconduïxen freqüentment a la defensa de formulacions teòriques que acaben establint una sobredeterminació del comportament individual. Per altra banda, hom pot recollir de la crítica ressenyada que acceptar un enfocament centrat solament en la racionalitat presenta el perill de no proveir suficients explicacions, i deixar massa situacions indeterminades o en mans de les probabilitats, per a l'explicació dels fenòmens socials. Aquest perill no impedeix intentar una defensa de l'explicació intencional, entesa com un enfocament que es pot distanciar de la seva vinculació tradicional amb la racionalitat com a fonament únic d'anàlisi<sup>27</sup>, sense rebutjar la seva preeminència explicativa. Finalment, resta l'interrogant -present als propers capítols- sobre si és necessari explicar des d'aquesta perspectiva els processos de formació de preferències mitjançant la complementarietat d'altres aproximacions, tals com la teoria de les normes socials o la teoria de les ideologies<sup>28</sup>, o bé es poden explicar les preferències a través de processos d'interacció social que generen propietats emergents.

---

<sup>26</sup> Vid. la secció 3.6., sobre les formes d'interpretar els processos de canvi social -a través de l'acció col·lectiva- en una anàlisi des de la perspectiva del conflicte de classes.

<sup>27</sup> Així, és clarificador un dels darrers textos de J. Elster referint-se a la relació entre normes socials i individualisme metodològic: "*A norm, in this perspective, is the propensity to feel shame and to anticipate sanctions by others at the thought of behaving in a certain, forbidden way (...) this conception of a network of shared beliefs and common emotional reactions does not commit us to thinking of norms as supra-individual entities that somehow exists independently of their supports*" (cf. J. Elster, *The Cement of Society*, Cambridge University Press, Cambridge, 1989, pàgs. 105-6). Per altre costat, A. Przeworski arriba a una conclusió semblant, en acceptar l'individualisme metodològic "*if we admit that results of individual actions are often indirect, unanticipated and sometimes unwanted by anyone*" ("Marxism and Rational Choice", a *Politics and Society*, vol. 14, 1985, pàg. 400), però destacant que l'aparell formal existent -bàsicament la teoria de jocs- és molt insuficient pel propòsit d'explicar com les accions dels individus en unes condicions donades produeixen noves condicions.

<sup>28</sup> Per exemple, en la línia dels treballs de J. Elster ("Social Norms and Economic Theory", *Journal of Economic Perspectives*, vol. 4, 1990 o *Sour Grapes...op.cit.*, 1983) i R. Boudon (*La lógica de lo social*, Rialp, Madrid, 1982 o *L'Ideologie*, Fayard, Paris, 1986). Acceptar la necessitat d'aquesta complementarietat implicaria reconèixer -amb totes les matisacions corresponents- les dificultats de mantenir en totes les circumstàncies l'exclusivitat d'una explicació intencional basada en els individus, sense renunciar -com a 'desiderata' explicativa- a aquest objectiu.



### 1.1.3. Racionalitat i interès propi: l'«homo oeconomicus».

El model d'home que tradicionalment ha servit de «portador» a la noció de racionalitat econòmica, es conegut amb el nom d'«homo oeconomicus». És precisament el desenvolupament conceptual i l'aplicació continuada d'aquest model el què ha constituït l'impuls bàsic pel refinament de la noció de racionalitat a l'economia -amb totes les seves contradiccions i varietats-. L'aspecte més remarcable de l'«homo oeconomicus» es la incorporació de la presumpció d'egoisme, com un supòsit estable sobre el comportament dels individus. Abans d'entrar a discutir les seves característiques, cal insistir que *com a model d'home, l'«homo oeconomicus» es només una configuració concreta, entre altres hipòteticament possibles d'elaborar amb els mateixos principis de la racionalitat econòmica*<sup>29</sup>. En aquest sentit, A. Sen -formulant una crítica a aquesta configuració- separa clarament la recerca de l'interès propi del comportament consistent, advertint que aquest darrer supòsit

*"Need not invariably involve the pursuit of one's self interest, since other types of goals and objectives may also induce similar consistency"*<sup>30</sup>

Com destaca B. Barry<sup>31</sup>, els supòsits que van configurar una primera versió completa de l'«homo oeconomicus» entre els utilitaristes, foren establerts per J. Bentham en els següents punts: *i) els homes actuen racionalment per aconseguir els seus fins, ii) el poder, l'estatus social i els bens materials són els objectes del desig dels homes en la major part de les societats, iii) les sancions externes per limitar les accions possibles per aconseguir els fins esmentats són més importants que les coercions internes*. Com a continuador i divulgador de J. Bentham, J. Mill va

---

<sup>29</sup> Aquest principis són analitzats detalladament a les properes seccions (1.2., 1.3.).

<sup>30</sup> A. Sen, "Economic Methodology: Heterogeneity and Relevance", Social Research, vol. 56, 1989 (pág. 319).

<sup>31</sup> B. Barry, Los sociólogos, los economistas y la democracia, Amorrortu, Buenos Aires, 1974, pàgs. 17-20.

desenvolupar una teoria política basada en aquest enfocament<sup>32</sup>, que no va tenir continuïtat a la ciència política fins pràcticament cent-cinquanta anys després, mentre que el corrent principal dels economistes va anar desenvolupant i refinant des de mitjans del segle XIX l'enfocament de Bentham. El model de l'«homo oeconomicus» constituït per aquests principis heurístics sobre el comportament de l'individu, va anar incorporant -posteriorment- més precisió sobre les característiques del supòsit de racionalitat. Així, la concepció d'un home bàsicament egoista, el qual per damunt de tot cerca de forma racional el propi interès -bàsicament de natura econòmica- en les seves accions<sup>33</sup>, es va convertir en el model d'home estàndard d'una bona part de la teoria econòmica. Aquest model d'home, es completa freqüentment amb dos supòsits fonamentals sobre el comportament individual: (i) *Que els individus calculen normalment de forma racional quina és la millor forma d'assolir uns objectius propis*, i (ii) *que l'aplicació de la racionalitat per part dels individus requereix un procés d'optimització*. D'aquests dos supòsits, se'n dedueix com a corol·lari que el comportament racional de les persones pren la forma d'optimització de l'interès propi. La teoria econòmica interessada en analitzar les conseqüències d'aquest tipus de conducta en els escenaris econòmics, ha pres com a part de l'explicació que el comportament de les persones s'ajusta als supòsits esmentats. Per a tal fi, aquesta línia teòrica s'ha centrat en cercar arguments consistents que connectin tot tipus de comportament observat amb el principi de la recerca racional de l'interès propi<sup>34</sup>. Però com apunta R. Hardin, els resultats més destacats d'aquesta perspectiva es

---

<sup>32</sup> A Essay on Government (1828), reeditat a J. Lively i J. Rees (Eds.), Utilitarian Logic and Politics, Clarendon Press, Oxford, 1978.

<sup>33</sup> Aquesta concepció és generada per la filosofia política del segle XVIII, i queda fixada amb l'obra A. Smith, que la vincula als principis de la bona economia (cf. A. Hirschmann, Las pasiones y los intereses, FCE, Mèxic, 1978, i F. Ovejero, Una teoria de la historia, Ms., Barcelona, 1988). És interessant destacar que d'aquesta perspectiva neix la idea que la recerca del propi interès és un comportament racional, exclouent altres recerques com irracionals. Enfront d'aquesta idea, en aquestes pàgines es defensa que racionalitat i interès són dues coses analíticament distintes.

<sup>34</sup> Així, la teoria econòmica sofreix la tendència a convertir-se en un «coneixement barroc», que allunyant-se de la validació empírica, desenvolupa un gran aparell formal, el qual esdevé una nova font de problemes (de plantejament i resolució), independitzant-se de la voluntat aclaridora amb què havia estat desenvolupat l'aparell conceptual que va donar origen a la teoria (cf. J. Elster i A. Hylland, "Introduction" a J. Elster i A. Hylland, (Eds.) Foundations of Social Choice..., op. cit., 1986).

refereixen sempre finalment als fenòmens de caràcter macroeconòmic:

*"It is an irony (...) that the notion of economic rationality applies conceptually to the individual, to homo oeconomicus, but that economic explanations based on that notion do not work well for individuals or ever for small numbers of actors (...) is largely a chimera"*<sup>35</sup>.

Una de les grans línies clàssiques de defensa de l'aplicació analítica del principi d'interès propi ha estat la reivindicació dels efectes de la mà invisible: que l'egoisme -l'interès individual- porta al benestar social (A. Smith), mentre que l'altruisme -l'interès social- porta a un malestar global (E. Burke)<sup>36</sup>. D'altra banda, alguns autors fan una defensa d'aquest pressupòsit a través d'un argument de tipus evolutiu, assenyalant que l'egoisme és la millor estratègia de supervivència individual, i afegixen que alhora és aquesta l'estratègia que produeix un millor resultat pel conjunt de la societat (i esmenten que això ho reflecteix precisament els efectes de la mà invisible)<sup>37</sup>. Malgrat aquest arguments, no es pot demostrar que

---

<sup>35</sup> R. Hardin, "Difficulties in the Notion of Economic Rationality", Social Science Information, vol. 23, 1984, pàg. 454. Precisament per aquest motiu, s'han presentat recentment anàlisis que desestimen la necessitat de fonamentar en els supòsits de la racionalitat econòmica individual proposicions teòriques de caràcter macroeconòmic (cf. G. Akerlof i J. Yellen, "Rational Models of Irrational Behavior", American Economic Review, vol. 77, 1987; i D. Kahneman, "Experimental Economics: A Psychological Perspective", R. Tietz, W. Alberts i R. Selten (Eds.), Bounded Rational Behavior in Experimental Games and Markets, Springer Verlag, Berlin, 1988). La línia que s'investiga primordialment a la tesi, però, és la que pretén fonamentar en la racionalitat la construcció teòrica, encara que diverses reflexions sobre aquesta qüestió tornen a aparèixer en els darrers capítols en referència als problemes de la ciència política.

<sup>36</sup> Hirschmann realitza una interessant crítica d'aquest argument en destacar que els *incentius perversos* són efectes econòmics de segon ordre, ja que es generen normalment com casos limitats, derivats de fenòmens molt més amplis. Llavors és de sentit comú que, en ser els beneficis de la *mà invisible* efectes de segon ordre, tenen generalment menys incidència que els perjudicis pels efectes intencionals directes derivats de la recerca del propi interès, que són de primer ordre (cf. A. Hirschmann, "Doscientos años de retórica reaccionaria. El caso del efecto perverso", El Trimestre Económico, vol. 56, 1989)

<sup>37</sup> Vid. H. Kliemt, "Der Homo Oeconomicus zwischen Ethik und Biologie -Aspekte seiner Geschichte und Wirkungsgeschichte", a M.J. Holler (Ed.), Homo Oeconomicus I, Verlag N. Leudemann, Munich, 1983. Hi ha altres línies de defensa (vid. S. Hargreaves Heap, Rationality in Economics, Basil Blackwell, Oxford, 1989; també F. Machlup, "L'Homo Oeconomicus et ses collègues", a E.M. Claassen (Ed.), Les Fondements Philosophiques des Systèmes Economiques, Payot, Paris, 1967), i a molts economistes és una qüestió que no els preocupa en absolut. Vid. en aquest sentit la discussió sobre la noció d'autointerès que presenta D. Parfit, Reasons and Persons, Clarendon Press, Oxford, 1984 (cap. 6).

un comportament egoista continuat porti sempre, en entorns canviants, als millors resultats per l'individu que el practica. Per exemple, en augmentar el nombre d'egoistes en un grup, en determinades circumstàncies poden reduir-se els resultats globals per aquests grup. D'aquesta manera, allò que es pot destacar és que tant les millors estratègies de supervivència individual, com els millors resultats globals són dependents de la configuració concreta de l'entorn en cada període concret, essent un element més la distribució d'actituds dels participants (suposar la consistència dels comportaments en diferents períodes de temps, només es pot entendre com l'adopció d'un supòsit simplificador)<sup>38</sup>.

Tot i aquestes dificultats, des d'alguns sectors de la teoria econòmica es manté la recerca de l'interès propi com a objectiu paradigmàtic de la racionalitat. El supòsit de l'existència d'una motivació única es manté perquè es considera que serveix per a predicar efectes econòmics col·lectius amb facilitat. El que és força significatiu d'aquest argument, és que en relació a l'«homo oeconomicus» es torna a reproduir -com a continuació- la polèmica esmentada en l'apartat anterior sobre la plausibilitat de la hipòtesi de racionalitat, mantenint en essència la mateixa tensió. Seguint aquesta tradició, l'enfocament de l'«elecció racional» -aplicat a àmbits externs de l'economia- atorga sovint un status de privilegi al supòsit d'un comportament egoista, en defecte o absència d'altres comportaments<sup>39</sup>. Però la qüestió no ha estat lliure de problemes<sup>40</sup>, perquè en aquests àmbits es considera generalment

---

<sup>38</sup> Un dels camps en què s'ha desenvolupat més aquesta disputa teòrica és a la microeconomia de l'empresa, com ho il·lustra la discussió sobre aplicacions concretes que presenta A. Jacquemin, La nueva organización industrial, Vicens Vives, Barcelona, 1989, cap. 2 (ed. or. 1987).

<sup>39</sup> Un criteri adoptat és que si serveix per explicar l'economia, també pot servir per explicar la política, ja que es tracta dels mateixos individus. Una versió més refinada defensa el pressupòsit que cal donar privilegi a l'interès propi (amb l'egoisme com a cas extrem) davant l'interès pels altres (amb l'altruisme com a cas extrem), considerant que és més fonamental, ja que no seria pensable un món compost per individus que només tinguessin interès pels altres, sense cap individu amb interès per «rebre». Ara bé, suposar que tothom actua per interès propi, com fa la teoria econòmica -supòsit motivat per la simplificació- és un cas extrem que pot ser contradictori amb moltes situacions reals. J. Elster, Nuts and Bolts for Social Sciences, Cambridge University Press, Cambridge, 1989, pàgs. 53-54.

<sup>40</sup> Vid. R. Zintl, "Ökonomisches Rationalitätskonzept und normorientiertes Verhalten", a Jahrbuch für Neue Politische Ökonomie, vol. 5, 1986.

que les institucions col·lectives afecten el comportament de la gent, i que el caràcter de la gent afecta el funcionament de les institucions, la qual cosa dificulta l'aplicació de l'enfocament econòmic a la interpretació dels fenòmens polítics, ja que als àmbits no econòmics en els quals s'aplica aquest enfocament, es segueix presentant l'explicació dels fenòmens només en la direcció de com els canvis de contextos afecten a un comportament individual determinat per la hipòtesi de racionalitat<sup>41</sup>.

Per introduir més detalladament la caracterització del comportament egoista, es segueixen els tres components bàsics que A. Sen distingeix<sup>42</sup>, associant a cada component un ideal clàssic de comportament: (1) El benestar individual, afegint que aquest ha de ser «autocentrat», o sigui, que no produeixi cap «externalitat» per altres individus, com per exemple ho podria provocar la simpatia, (2) els objectius individuals, que han d'estar enfocats a maximitzar el benestar propi, sense tenir en compte el benestar dels altres, i (3) les eleccions individuals, que s'han de centrar en assolir els objectius personals, sense considerar els objectius dels altres. A. Sen afegeix que les variacions d'aquest model d'home desenvolupades al llarg de la dècada dels vuitantes, poden prendre formes diferents segons el component violat (o la combinació de components). En aquest sentit, no s'ha de pensar la crítica del model de l'«homo oeconomicus» com un conflicte que oposa interès propi a alguna forma d'altruisme general. Al contrari, s'ha de considerar la possibilitat d'un ampli ventall de possibles actituds, que suggereixen dependències i determinacions que cal anar aclarint, cas per cas, o bé a través d'elaboracions teòriques específiques, com podrien ser, per exemple, la definició de mecanismes d'identificació de preferències en processos d'acció col·lectiva. En aquest sentit, s'ha defensat que la racionalitat econòmica no ha d'implicar un determinat model d'home, si es vol utilitzar com a «microfonament» d'aplicació general, i que la constel·lació concreta que representa

---

<sup>41</sup> B. Barry, "Methodology...", art. cit., 1982. El tema es discuteix en detall en la segona part de la tesi, quan es pren com àmbit d'anàlisi d'aquests problemes la teoria de l'acció col·lectiva.

<sup>42</sup> A. Sen, "Economic...", art. cit., 1989, pàgs. 323-4.

l'«homo oeconomicus» només és aplicable a camps força concrets de les activitats humanes<sup>43</sup>. D'altra banda, la freqüent aplicació de l'«homo oeconomicus» a la política hauria d'implicar que les crítiques i discussions existents a la teoria econòmica es traspassessin igualment a l'àmbit de la teoria política que segueix l'enfocament esmentat; però és força infreqüent aquesta actitud entre els politòlegs. D'aquesta forma, apareixeria un camp obert per a modelitzar les motivacions dels individus en la interpretació dels fenòmens polítics.

Malgrat tot el que s'ha esmenat, com es veura més endavant, els problemes més importants del paradigma econòmic aplicat a la política no es deriven tant de les orientacions de l'interès individual, sinó que són degut a altres components d'aquest paradigma explicatiu, destacant dos molt especialment. *El primer problema és l'adaptació de la hipòtesi de racionalitat per a ser utilitzada amb la finalitat d'explicar diversos tipus d'accions intencionals i els seus efectes. Cal augmentar la seva plausibilitat com a teoria descriptiva, tot i separant la seva formulació dels objectius normatius i prescriptius*<sup>44</sup>. *El segon problema és la qüestió de la contextualització. Deixant a banda la hipòtesis de comportament que es mantingui, és necessari formular una explicació de com limiten i interactuen els contextos i les institucions sobre les possibilitats d'acció dels individus? alhora, és possible explicar com la interacció entre individus porta a modificar i alterar aquests contextos?* Aquests problemes generen una important discussió dins l'enfocament de l'«elecció racional», que constitueixen una de les tensions centrals desenvolupades al llarg de tot el treball. La resta del capítol, dedicada a discutir els elements que componen la hipòtesi de racionalitat, ja incorpora diverses consideracions referents a la tensió metodològica que implica la primera pregunta.

---

<sup>43</sup> R. Zinlt, "Der Homo Oeconomicus: Ausnahmeerscheinung in jeder Situation oder Jedermann in Ausnahmesituationen?", *Analyse und Kritik*, vol. 11, 1989.

<sup>44</sup> Un exemple recent d'aquestes crítiques es pot trobar a D. Bell, H. Raiffa i A. Tversky, "Descriptive, Normative and Prescriptive Interactions in Decision-Making", D. Bell, H. Raiffa i A. Tversky, (Eds.), *Decision-Making: Descriptive, Normative and Prescriptive Interactions*, Cambridge University Press, Cambridge, 1988.

Com s'ha esmentat, amb els principis de la racionalitat econòmica es poden configurar formes distintes de models d'home, en combinar diferents orientacions d'interès (preferències) amb els supòsits que configuren un procés de decisió racional. Cal afegir, a més, que aquesta tasca pot servir tant per l'objectiu de proveir una teoria normativa de la decisió racional, com per establir principis que serveixin per a predir el comportament observat, tot i que no són objectius idèntics i, per tant, cal analitzar què és el que comparteixen.

#### 1.1.4. Limitacions a la racionalitat en les ciències socials.

Trencat l'acord tradicional implícit entre economistes i sociòlegs -els primers es dedicaven a l'estudi de la racionalitat i els segons a l'estudi de la no racionalitat-, que autors com W. Pareto varen teoritzar els primers anys d'aquest segle, s'ha produït en els darrers anys una pressió sobre els defensors del paradigma econòmic per sortir de la còmoda posició en què es trobaven. Davant aquesta pressió, els economistes han començat a utilitzar, com destaca S. Lindenberg,

*"psychological propositions and direct measurement of utility rather than unrealistic models and start explaining preferences rather than taking them as given"*<sup>45</sup>.

En termes generals, seguint diverses tradicions teòriques, s'han pres diverses vies: incloure les «anomalies» detectades refinant les argumentacions a partir dels supòsits bàsics del seu model d'home i del concepte de racionalitat econòmica, perfeccionar l'experimentació i l'anàlisi de la conducta econòmica, discutir els

---

<sup>45</sup> S. Lindenberg, "Rational Choice and Sociological Theory: New Pressures on Economics as a Social Science", a Journal of Institutional and Theoretical Economics, vol. 141, 1985, pàg. 244.

problemes de la teoria de la racionalitat, etc.<sup>46</sup>. De forma inversa a la teoria econòmica, hom pot detectar una situació similar a l'esmentada en força enfocaments sociològics i politològics. Aquests intenten incorporar dins els seus plantejaments alguna perspectiva d'autonomia individual, per evitar el «globalisme» determinista i funcionalista. També hi ha alguna corrent sociològica o politològica que ha adoptat l'enfocament de l'«elecció racional» amb plantejaments metodològics similars als dels economistes<sup>47</sup>. Per definir millor els contextos en què actua el model d'home, introdueixen suggeriments sociològics sobre regles i estructures socials, intentant explicar les «anomalies» que observen en les seves prediccions. Tot i que la sociologia tradicional ha defensat la preeminència de la noció de norma per a explicar el comportament, no implica cap contradicció que des de la perspectiva de l'«elecció racional» s'intenti incorporar aquesta noció de forma complementària. Com destaca en aquest sentit Vanberg,

*"The crucial issue in this context is, of course, the definition and theoretical interpretation of the notion of rule -or norm- guided behavior, in particular the possibility of consistently incorporating such a notion into a rational choice framework"*<sup>48</sup>.

---

<sup>46</sup> Per exemple, com a panoràmica d'alguns d'aquests corrents, vegeu P.E. Earl, "Introduction" a Behavioral Economics I, Aldershot, Elgar, 1988. Criticant els treballs de O. Williamson (per exemple, Markets and Hierarchies. Analysis and Antitrust Implications, Free Press, Nova York, 1975), Earl destaca que qualsevol apropiació amb bases neoclàssiques és deformatora del sentit del seu enfocament; però potser caldria matisar que són precisament els intents de connexió una de les vies més interessants per desenvolupar el coneixement teòric a l'economia. En tot cas, el moviment teòric és molt ampli i segueix també altres direccions (una panoràmica de força d'aquestes direccions es pot trobar a R. Marimon i X. Calsamiglia (Eds.), Invitació a la teoria econòmica, Ariel, Barcelona, 1991).

<sup>47</sup> Com a visió general d'aquesta perspectiva, vegeu K.-D. Opp, "Sociology and Economic Man", Journal of Institutional and Theoretical Economics, vol. 141, 1985; i T. Voss, "Rationale Akteure in der soziologischen Analyse", a M.J. Holler (Ed.), Homo Oeconomicus I, op. cit. 1983. Per a un desenvolupament del concepte d'«anomalia» com a categoria general que complementa l'enfocament de l'«elecció racional», en la línia d'explicar l'emergència i persistència de les institucions com respostes a les fallides de la racionalitat individual -anomalies-, vegeu B. Frey i R. Eichenberger, "Anomalies and Institutions", Journal of Institutional and Theoretical Economics, vol. 145, 1989.

<sup>48</sup> V. Vanberg, "Rules and Choice in Economics and Sociology", Jahrbuch für Neue Politische Ökonomie, vol. 7, 1988, pàg. 147.



Donades aquestes tendències contraposades, però fins a cert punt convergents, apareix fàcilment la pregunta sobre les possibilitats de realitzar síntesis i aproximacions integradores. Els enfocaments que trenquen amb els postulats de la racionalitat econòmica, tot i mantenir-se dins la perspectiva de la racionalitat estricta, es podrien entendre com un intent de cercar posicions intermèdies entre els plantejaments d'origen sociològic i els d'origen econòmic. Per aquest motiu, abans d'entrar detalladament en els propers apartats d'aquest capítol a l'anàlisi de la noció de racionalitat dominant a la teoria econòmica, es vol presentar i discutir un dels intents més importants d'establir una formulació alternativa dins el conjunt que s'ha anomenat 'teories estrictes' de la racionalitat. Amb força afinitats a la perspectiva de la racionalitat sociològica (vid. apt. 1.1.1.), la proposta de H. Simon sobre la racionalitat limitada ("Bounded Rationality") neix i es desenvolupa al camp de l'economia, però considera amb especial atenció la influència de la definició d'objectius dins el procés de decisió que vincula instrumentalment els medis amb els fins.

Aquesta proposta, àmpliament difosa i discutida des de la seva formulació durant els anys cinquantes<sup>49</sup>, té un argument centrat en la idea de què el comportament humà no busca maximitzar objectius, sinó que adopta el primer curs d'acció factible que satisfà uns requeriments mínims. Simon distingeix entre racionalitat processal i racionalitat substantiva. La primera seria la teoria de la racionalitat clàssica de l'economia, mentre que la segona implicaria considerar l'existència d'un procés deliberatiu sobre les condicions de l'elecció, tant referent als fins o objectius (derivat de la incertesa computacional interna sobre quines són les accions possibles), com als mitjans i restriccions de l'elecció (o la incertesa

---

<sup>49</sup> Per exemple, del propi H. Simon, entre la seva extensa producció, hom pot destacar Reason in Human...op.cit., 1983; o Models of Man, Wiley, Nova York, 1957. Una visió acurada de la teoria de H. Simon es pot trobar a Ph. Mongin, "Simon, Stigler et les théories de la rationalité limitée", Social Science Information, vol. 25, 1986.

informativa externa sobre les conseqüències de cada acció)<sup>50</sup>. En aquesta perspectiva, la discussió sobre els valors és interna a la pròpia teoria, com el propi Simon destaca:

*"When conflicts in values exist, as they almost always do, it is not even clear how 'optimal' is to be defined. But all is not lost. Reconciling alternative points of view and different weighting of values becomes somewhat easier if we adopt a satisficing point of view: if we look for good enough solutions rather than insisting that only the best solutions will do"*<sup>51</sup>.

Des de la perspectiva de la racionalitat econòmica s'ha intentat recuperar la noció de racionalitat limitada, mantenint el concepte d'optimització<sup>52</sup>. Des d'aquest enfocament, la recerca dels requeriments mínims s'ha intentat entendre en general com una forma de maximització en condicions de restriccions externes a l'acció racional. Es tracta de realitzar el mínim esforç per assolir els màxims resultats, amb uns objectius i unes condicions informatives donades. Alguns autors, com J. Elster, destaquen que el problema de la incorporació de la teoria de Simon és que la racionalitat limitada no pot incloure una argumentació convincent que sigui traduïble en termes de la teoria de la racionalitat econòmica, explicant com l'individu aconsegueix un punt d'equilibri òptim en la seva recerca d'informació per al comportament «satisfaent»:

*"The attempt to re-define satisficing as a form of maximizing leads to an infinite regress, because the choice of a profit maximizing information*

---

<sup>50</sup> H. Simon, "From Substantive to Procedural Rationality", art. cit., 1976; i H. Simon, "Rationality as Process and as Product of Thought", American Economic Review, vol. 68, 1978. Vid. també Ph. Mongin, "Simon, Stigler et...", art. cit., 1986.

<sup>51</sup> H. Simon, Reason in Human..., op. cit., 1983, pàg. 84.

<sup>52</sup> El primer treball en aquest sentit és de J. Stigler, "The Economics of Information", Journal of Political Economy, vol. 69, 1961. Stigler introdueix el concepte de costos de recerca d'informació, com una nova restricció per una decisió racional en termes d'optimalitat.

*structure itself requires information, and it is not apparent how the aspiring profit maximizer acquires this information, or what guarantee that he does not pay an excessive price for it*<sup>53</sup>.

Però aquesta crítica només té sentit si es vol integrar la teoria de la racionalitat limitada com a cas especial de la teoria de la racionalitat econòmica. Així, l'única possibilitat per solucionar el problema seria adoptar el supòsit que la meta-decisió no té cap cost per l'agent<sup>54</sup>. Sense la pretensió d'integrar la racionalitat limitada en les premisses de la racionalitat econòmica, hom pot entendre els nivells d'aspiració -que determinen què és suficient- com quelcom que no ha de ser tractat necessàriament de forma teòrica, sinó que pot deixar-se per definir a les aplicacions que es desenvolupin<sup>55</sup>; i també a l'inrevés: considerar que l'actitud optimitzadora és un cas especial de la teoria de la racionalitat limitada -quan es comptabilitza el cost de la recerca d'informació sobre uns objectius definits prèviament<sup>56</sup>. Un altre intent d'integració ha estat mitjançant un model que incorpora el comportament governat per regles quan el nivell d'incertesa

---

<sup>53</sup> J. Elster, "Anomalies of Rationality: Some Unresolved Problems in the Theory of Rational Behavior", a L. Lévi-Garboua (ed.), Sociological Economics, Sage, Londres, 1979, pàg. 74.

<sup>54</sup> El tema és discutit per H. Gottinger ("Computational Cost and Bounded Rationality", a W. Stegmüller, W. Balzer i W. Spohn (Eds.), Philosophy of Economics, Studies in Contemporary Economics, vol. 2, Springer Verlag, Berlin-Heidelberg, 1982), demostrant en forma de teorema d'impossibilitat aquesta única alternativa. Una proposta alternativa per mantenir el supòsit d'optimització és realitzat per G. Wiston, però a costa de definir diferents nivells de racionalitat (a "Imperfectly Rational Choice", Journal of Economic Behavior and Organization, vol. 12, 1989). El problema de la proposta de Wiston és que, apart de definir limitacions temporals per aplicar el càlcul racional, no defineix explícitament com operen els diferents graus de la racionalitat.

<sup>55</sup> Per exemple, aquest és l'argument de R. Tietz per millorar una teoria descriptiva de la negociació que abandoni el paradigma de l'optimització (R. Tietz, "Experimental Economics: Ways to Model Bounded Rational Bargaining Behavior", a R. Tietz, W. Albers i R. Selten (Eds.), Bounded Rational Behavior in Experimental Games and Markets, Springer Verlag, Berlin-Heidelberg, 1988. D'altra banda, R. Radner proposa un model de satisfacció en el qual el comportament és guiat per una regla que distribueix l'aplicació de l'esforç entre diverses activitats en funció de l'experiència anterior ("Satisficing", Journal of Mathematical Economics, vol. 2, 1975).

<sup>56</sup> H. Simon, "Bounded Rationality", The New Palgrave. A Dictionary of Economics, Vol. 1, 1987.

informativa és força elevat<sup>57</sup>. Tot i el seu interès, aquest i altres intents d'incorporar regles de comportament al model de l'«homo oeconomicus» tenen el problema de que no expliquen els processos d'aprenentatge per a la definició d'objectius, i resolen el problema modelant unes eleccions dominades per regles<sup>58</sup>.

L'enfocament de la racionalitat limitada entra dins l'espectre de teories que defensen el tipus d'explicació intencional, i sembla tenir més plausibilitat intuïtiva que altres enfocaments basats en la hipòtesi de comportament racional. Hi força suggeriments pel seu desenvolupament, que encara presenten greus problemes per la seva construcció teòrica. La més destacada es deguda a les possibilitats de flexibilitzar la teoria de la racionalitat limitada amb motivacions i actituds variades, com les que proposen els psicòlegs socials que estudien els processos de decisió<sup>59</sup>, gràcies a la seva major capacitat per adoptar-les que la teoria normativa de la decisió<sup>60</sup>. Tot i que s'ha defensat que aquesta teoria té grans possibilitats d'expansió

---

<sup>57</sup> La proposta ha estat elaborada per R. Heiner, "The Origin of Predictable Behavior", American Economic Review, vol. 73, 1983 i "Imperfect Choice and Self-Stabilizing Rules", Economics and Philosophy, vol. 5, 1989. Heiner destaca que l'optimització apareix quan la incertesa s'aproxima a zero, però com que normalment la gent no pot abastar tota la informació disponible, apareix la incertesa. Llavors, es pot entendre la incertesa com la font indirecta que permet predir un comportament guiat per normes.

<sup>58</sup> Per a una crítica a aquests intents d'incorporar regles, vid. M. Bianchi, "The Unsatisfactoriness of Satisficing: from Bounded Rationality to Innovative Rationality", Review of Political Economy, vol. 2, 1990; i J. Martin, G. Kleindorfer i W. Brashers, "The Theory of Bounded Rationality and the Problem of Legitimation", Journal for the Theory of Social Behavior, vol. 17, 1987.

<sup>59</sup> Per exemple H. Einhorn i R. Hogarth, "Behavioral Decision Theory: Processes of Judgement and Choice", a D. Bell i altres, Decision-Making...op.cit., 1988; o R. Thaler, "The Psychology of Choice and the Assumptions of Economics", a A. Roth (Ed.), Laboratory Experimentation in Economics, Cambridge University Press, Cambridge, 1987.

<sup>60</sup> Així ho apunten en aquesta línia A. Tversky i D. Kahneman, a "Rational Choice and the Framing of Decisions", a R. Hogarth i M. Reder, (Eds.), Rational Choice, University of Chicago Press, Chicago, 1987 i J. March, "Bounded Rationality, Ambiguity and the Engineering of Choice", Bell Journal of Economics, vol. 9, 1978.

a les ciències socials<sup>61</sup>, presenta dos importants problemes per a la seva aplicació al desenvolupament de propostes teòriques. El primer és com tractar els nivells d'aspiració perquè no siguin determinats de forma exògena en cada decisió; i el segon, com definir i explicar el procés de generació d'alternatives que els individus realitzen quan efectuen una decisió.

Degut als problemes esmentats, *es pot concloure que la teoria de la racionalitat limitada difícilment es pot convertir en una hipòtesi sobre el comportament d'aplicació general -perquè depèn sempre de la informació empírica- que pugui substituir les propostes tradicionals de la racionalitat econòmica.* D'altra banda, la seva capacitat per interpretar els mecanismes de decisió tampoc és prou satisfactòria, ja que difícilment es pot vincular la racionalitat limitada amb propostes teòriques sobre els nivells de satisfacció, si es vol deixar suficient espai per a l'anàlisi dels efectes agregats de l'acció individual. Per aquests motius, encara que la proposta de Simon té un interès conceptual remarcable, *cal cercar altres vies que permetin alterar el predomini de la noció de racionalitat econòmica com a hipòtesi d'aplicació general, o que -alternativament- aconseguixin relacionar elements d'aquesta noció amb altres components teòrics, que mantinguin l'us de l'explicació intencional.* La via desenvolupada a la tesi investiga bàsicament aquesta darrera aspiració. L'objectiu es integrar possibles variacions de la hipòtesi de racionalitat com a part d'un conjunt de mecanismes teòrics de decisió que depenguin d'entorns analítics més amplis. Aquesta estratègia d'investigació segueix la *discussió sobre una teoria de la racionalitat, ja construïda i apta per a propostes normatives, per analitzar si és possible establir altres aproximacions al comportament racional més adaptades a necessitats descriptives i predictives, però no dependents d'aplicacions*

---

<sup>61</sup> Una interpretació en aquest sentit és defensada per R. Boudon: *"The notion of subjective rationality sketches a paradigm in the frame of which all kinds of behavior, including behavior oriented against the interest of the subject, as well as all kinds of attitudes and beliefs, including magical beliefs and more generally beliefs in false ideas, can be more convincingly explained than with other paradigms"* (a "Subjective Rationality and the Explanation of Social Behavior", Discussion Paper, 89/6, Max-Plank-Institut für Gesellschaftsforschung, Colonia, 1989, pàg. 41).

empíriques concretes.

## 1.2. Comportament racional i explicacions basades en preferències.

### 1.2.1. La lògica del model d'home amb comportament racional.

Recuperant la discussió prèvia sobre els models d'home, cal destacar la crítica clau de B. Barry a l'elaboració de teories socials basades en supòsits de comportament individual:

*"¿En qué punto el precio que se paga en falta de realismo, con el objeto de ganar simplicidad y poder deductivo, se torna demasiado alto?"<sup>62</sup>.*

B. Barry argumenta que, degut a la gran complexitat de les situacions polítiques, les aplicacions d'aquesta perspectiva són força limitades per deduir-ne solucions ben definides. La utilitat, limitada o no, dels enfocaments que construeixen models d'home per explicar les interaccions socials i polítiques, resta, per ara, en un interrogant<sup>63</sup>. Abans de discutir aquesta utilitat, cal explorar quina es la flexibilitat que pot incorporar l'enfocament de la racionalitat econòmica a la formulació d'un model d'home. *Com més flexibilitat s'incorpori, major capacitat d'adaptació i més marge de maniobra tindran els models d'home, tot i que dificultaran la tasca de construcció teòrica.* Precisament per aquest motiu, l'anàlisi s'ha de centrar en investigar en quins punts s'equilibren millor les dues tendències contraposades,

---

<sup>62</sup> B. Barry, *Los sociólogos...*, op.cit., 1974, pág. 198.

<sup>63</sup> Una perspectiva a destacar, i que es considera més àmpliament a la segona part de la tesi, és la que defensa B. Hindess, en senyalar que la utilitat d'aquest enfocament depèn de "*how limited those contexts are. Rational Choice theory requires a number of specific postulates which may well have heuristic value in certain cases*" (Cf. "Rational Choice Theory and the Analysis of Political Action", *Economics and Society*, vol. 13, 1984, pág. 258). El tractament de la indeterminació per J. Elster destaca els casos en què la teoria de l'acció racional porta a prediccions no unívocues (existència de molts òptims) o quan no existeix cap elecció racional possible, conclouent que "*The theory is more powerful when applied to medium-sized problems like the purchase of a car or a house, but even here the question of optimal search is largely indeterminate*" ("When Rationality Fails", *Solomonic Judgements*, Cambridge University Press, Cambridge, 1989, pág. 27).

segons els procediments que s'adoptin.

En opinió d'alguns autors, una de les línies de defensa de l'enfocament de l'«elecció racional» en la seva expansió a les ciències socials, és la producció de resultats explicativament poderosos a partir d'uns principis força simples. Aquesta pretensió inclou el supòsit de què les interaccions d'una pluralitat d'actors puguin ésser inferides a través de models basats en principis axiomàtics. En aquest sentit, una part de les discussions d'aquest enfocament es centren en proposar i criticar les característiques concretes que el model d'home ha de tenir, a més de la capacitat d'optimitzar i de decidir racionalment. En aquest apartat -i en el següent- s'analitza la utilització de la racionalitat com a hipòtesi per a predir el comportament observat. Per a tal fi, s'han de realitzar algunes incursions en la teoria normativa de la decisió racional. En aquesta anàlisi, també es qüestionen les formes de connexió entre preferències i decisions racionals, destacant que tant la connexió lògica com la causal presenten dificultats importants per a ésser adoptades en la interpretació del comportament observat.

Seguint la proposta de M. Hollis<sup>64</sup>, es poden destacar com a més importants tres principis bàsics que inspiren el model d'home que es comporta racionalment, sense incloure el principi de comportament egoista: i) Es suposa que els actors tenen una certa predisposició per l'auto-interès; ii) Els actors són àtoms socials (podent ésser escollits aleatòriament dels seus grups), sense considerar estructures socials, posicions, rols, etc.; iii) Els actors són racionals i la seva racionalitat s'entén rigorosament en termes utilitaristes.

Les diferències amb l'«homo oeconomicus» són importants en cada un dels tres principis. En relació a la proposta de Hollis, els principis bàsics que caldria afegir-hi són: en i), el pressupòsit d'un egoisme continuat per aconseguir l'interès

---

<sup>64</sup> M. Hollis, "Rational Man and Social Science", a R. Harrison (Ed.), Rational Action. Studies in Philosophy and Social Science, Cambridge University Press, Cambridge, 1979.



propi; en ii), es redueixen els grups d'actors possibles a un sol tipus d'àtom bàsic; i en iii), el principi és el supòsit de maximització, i l'assoliment d'òptims en la consecució dels fins, donats els recursos disponibles (vid. apt. 1.1.3.). Com a exemple, es pot destacar que una de les versions més ortodoxes de l'«homo oeconomicus» pressuposa l'estabilitat de les preferències de l'individu en tots els seus àmbits vitals. El defensor més destacat d'aquest corrent és Gary Becker, qui argumenta que:

*"The heart of my argument is that human behavior is not compartmentalized, sometimes based on maximizing, sometimes not, sometimes motivated by stable preferences, sometimes by volatile ones, sometimes resulting in an optimal accumulation of information, sometimes not. Rather, all human behavior can be viewed as involving participants who maximize their utility from a stable set of preferences and accumulate an optimal amount of information and other inputs in a variety of Markets"<sup>65</sup>.*

Aquesta perspectiva deriva cap a una comprensió del món que es podria caracteritzar pel principi que 'tot el que s'esdevé és explicable com el resultat d'una acció racional'. Per tant, es cerquen explicacions que parteixen de suposar la racionalitat en tot comportament, per a tal de justificar la maximització continuada d'unes preferències invariables, afegint-hi les premisses que siguin necessàries. Aquestes premisses són bàsicament els canvis dels contextos i les alteracions en la disponibilitat de recursos, que provoquen canvis de comportament, per aconseguir que l'individu maximitzi sempre els seus objectius.

Tant els resultats experimentals com l'experiència quotidiana mostren que amb certa freqüència la gent no actua seguint el seu propi interès -aplicant el principi de l'egoisme-, i menys de forma absoluta, maximitzant sempre unes preferències

---

<sup>65</sup> G. Becker, The Economic Approach to Human Behavior, Chicago University Press, Chicago, 1976, pàg. 14.

invariables<sup>66</sup>. No obstant això, una línia per a defensar el supòsit d'interès propi podria ésser destacar que la seva absència és bàsicament deguda a un comportament irracional dels agents: falla l'anàlisi racional dels subjectes en reconèixer quins són els seus interessos. Però aquesta interpretació contradiu la pretensió del propi model de suposar sempre l'existència d'un comportament racional, i implicaria considerar com inexplicable un fenomen, que potser es podria explicar mantenint el supòsit de racionalitat i una altre orientació d'interès. Un altre argument per a defensar el supòsit d'interès propi és destacar que actituds aparentment altruistes poden encobrir en el fons un càlcul per afavorir al propi interès<sup>67</sup>; però el problema d'aquests raonaments és que fàcilment poden caure en explicacions de caire conspiratiu. Davant la defensa de la invariabilitat de l'interès propi com a supòsit de comportament, altres autors recullen els suggeriments dels resultats empírics. B. Hindess argumenta sobre aquesta qüestió que la

*"Rational choice theory does not deny the existence of other motivations. The claim is rather that the presumption of rational self-interested action provides the starting-point for analysis, which may well be forced to recognize the role of other motivations"*<sup>68</sup>.

No obstant això, la tesi d'Hindess és força discutible, perquè depèn d'una interpretació contemporània sobre el comportament humà. De forma més refinada assenyala Elster que l'interès propi té una certa prioritat metodològica sobre l'altruisme, degut a que l'objectiu de l'altruista és proveir plaers propis als altres,

---

<sup>66</sup> Vegeu per exemple l'experiment realitzat per N. Frohlich i J. Oppenheimer, "Beyond Economic Man. Altruism, Egalitarianism and Difference Maximizing", Journal of Conflict Resolution, vol. 28, 1984.

<sup>67</sup> Vegeu per exemple, A.A. Alchian i W.R. Allen, "The Pure Economics of Giving", a Diversos autors, The Economics of Charity, The Institute of Economic Affairs, Londres, 1973.

<sup>68</sup> B. Hindess, Choice, Rationality and Social Theory, Uwin Hyman, Londres, 1988, pàg. 35.

i per tant, sempre hi ha d'haver un egoista que s'aprofiti dels serveis dels altruistes<sup>69</sup>. Però d'aquesta prioritat metodològica, afegeix Elster, no se'n pot inferir que l'interès propi és la motivació més estesa. Com a cas extrem, podria haver-hi solament un egoista, el qual rebés totes les atencions de la resta de membres del grup o societat altruistes. *La dicotomia entre egoisme i altruisme es una excessiva simplificació de les possibles actituds que pot adoptar la modelització d'un «home racional»*. Els estats possibles no són altruisme i egoisme, sinó que aquests són només dos possibles extrems. No té excessiva dificultat postular l'existència de moltes orientacions d'interès entre els dos caràcters ideals, fent intervindrà més d'una dimensió<sup>70</sup>. Hom pot trobar altres orientacions, com per exemple les oportunistes (quan l'egoisme no està restringit per normes de comportament), o les destructores (per odi o enveja). Les formes de mesurar i analitzar aquestes diverses orientacions, que en gran part prenen sentit com actituds envers els altres, tenen per aquest motiu més facilitat de desenvolupament a partir d'una teoria de la interacció (vid. apt. 2.2.2.).

Un plantejament diferent sobre les motivacions del comportament sorgeix de la distinció clàssica weberiana entre l'ètica de la convicció i l'ètica de les conseqüències, enteses com dos principis de comportament que fonamenten meta-motivacions, a partir de les quals es poden ordenar les preferències<sup>71</sup>. Un desenvolupament d'aquesta perspectiva és elaborat per J. Elster<sup>72</sup>, destacant l'existència de diversos «tipus ideals», o actituds humanes bàsiques, entre els que

---

<sup>69</sup> Com diu Elster, "*we cannot coherently imagine a world in which everyone had exclusively altruistic motivations*" (a *Nuts and Bolts...*, op.cit., 1989, pàg. 53). Afegeix també que "*if some are to be altruistic, others must be selfish, at least some of the time, but everybody could be selfish all the time*" (pàg. 54).

<sup>70</sup> Com ho fan els psicòlegs socials que utilitzen mètodes experimentals per a analitzar diferents actituds en un problema de decisió. Per exemple U. Schulz i T. May, "The Recording of Social Orientations with Ranking and Pair Comparison Procedures", *European Journal of Social Psychology*, vol. 19, 1989.

<sup>71</sup> Una proposta general en aquest sentit és la realitzada per D. George, "Meta-Preferences: Reconsidering Contemporary Notions of Free-Choice", *International Journal of Social Economics*, vol. 11, 1984.

<sup>72</sup> J. Elster, *Nuts and Bolts...* op. cit., 1989.

destaca els individus «kantians» (seguidors de l'ètica de la convicció), els «utilitaristes» i els «circumstancialistes» (ambdós seguidors de l'ètica de les conseqüències). Els primers són els que adopten actituds de principis en les seves decisions, els segons són els que calculen sempre en funció del seu interès propi la seva decisió, i els darrers estableixen primordialment les seves decisions en funció de les decisions dels altres.

El funcionament de l'aparell motivacional també ha provocat propostes més complexes, que no es limiten a un supòsit unívoc sobre l'orientació dels interessos de l'individu, sigui el de l'altruisme o de l'egoisme, o a l'aplicació d'un principi de comportament, sigui aquest el circumstancialisme o l'utilitarisme. En formes diferents, aquestes propostes assenyalen que és necessari afegir més teoria a la definició del model d'«home racional». Les formes possibles de desenvolupar aquest propòsit es refereixen a la qüestió dels ordres de preferències (establint relacions entre més d'un ordre) de l'individu i, en alguns casos, a la introducció d'elements descriptius de caire psicològic. Entre les diverses propostes existents, cal destacar la d'A. Sen, qui insisteix en les seves crítiques contra el supòsit d'egoisme, introduint com alternativa els elements de simpatia i compromís (donant a l'actor més d'un ordre de preferències) per tal d'alterar en determinades circumstàncies la recerca de l'interès propi, que reconeix com a central en el comportament de l'individu<sup>73</sup>. Posteriorment, generalitzant la seva proposta, A. Sen proposa l'anàlisi d'una escala de preferències per a cada actor, en funció dels diversos camps d'acció d'aquests, afectats, en cada cas, per intensitats diferents de factors emocionals<sup>74</sup>.

---

<sup>73</sup> A. Sen, "Rational Fools. A Critique of the Behavioural Foundations of Economic Theory", Philosophy and Public Affairs, vol. 6, 1977. Sobre les incorporacions de característiques psicològiques, veure J. Payne, "Contingent Decision Behavior", Psychological Bulletin, vol. 92, 1982; i J. Morgan, "Multiple Motives, Group Decisions, Uncertainty, Ignorance, and Confusion: A Realistic Economics of the Consumer Requires Some Psychology", American Economic Review. Papers and Proceedings, vol. 68, 1978.

<sup>74</sup> A. Sen, Choice, Welfare and Measurement, Basil Blackwell, Oxford, 1982, i A. Sen, Commodities and Capabilities, North-Holland, Amsterdam, 1985.

Una altra proposta interessant és la d'H. Margolis<sup>75</sup>, qui considera que els seus actors tenen alhora un ordre de preferències altruistes i egoistes, i un mecanisme de distribució que determina quina proporció dels recursos de l'actor es dedicarà en cada cas a finalitats altruístiques o d'interès propi. La proposta d'H. Margolis és també una forma d'ampliar l'estructura teòrica del model, però en aquest cas, sense sortir dels principis explicatius de l'elecció racional en la definició de les preferències.

Més enllà dels problemes d'algunes de les propostes existents sobre incorporació de elements teòrics més complexos en els models d'actor, cal considerar que la utilitat d'aquestes propostes està molt relacionada amb la capacitat que tinguin per servir com a fonaments explicatius de teories substantives sobre processos socials, polítics o econòmics. Per a la major part és aplicable l'anàlisi de les relacions entre preferències i eleccions, utilitzant alguns supòsits de la teoria de la decisió racional, si es defineixen prèviament les preferències resultants<sup>76</sup>.

Per concloure l'apartat, es pot destacar que a l'anàlisi de l'«elecció racional» el més destacat és la preeminència de les propietats de racionalitat, i no els supòsits sobre l'orientació del comportament. Addicionalment, també es pot apuntar que el supòsit d'optimització i la recerca de punts d'equilibri tenen una posició més central que les consideracions sobre l'orientació del comportament en aquest enfocament.

---

<sup>75</sup> H. Margolis, "A New Model of Rational Choice", *Ethics*, vol. 91, 1981, i H. Margolis, *Selfishness, Altruism and Rationality. A Theory of Social Choice*, Cambridge University Press, Cambridge, 1982.

<sup>76</sup> D'altra banda, la proposta de F. Schick presenta una perspectiva distinta, en destacar que la gent pot actuar per raons que no siguin racionals, com per exemple els motius socials (compromís amb un valor ètic). Aquest autor accepta la possibilitat de què la gent no sempre persegueixi els seus interessos, argumentant que "*I have agreed that every action has a reason behind it and also that a rational person pursues his interests in a certain way. This leaves the link between reasons and rationality, with which the rationalist checks his case. My thesis has been that this link is open. Not every reason is rational, and there may sometimes be social reasons where there are no rational ones. Reasons go beyond rationality, and so does the range of action too*" (Cf. F. Schick, *Having Reasons: An Essay on Rationality and Sociality*, Princeton University Press, Princeton, 1984, pàg. 148). Aquesta posició suggereix la necessitat de discutir el paper dels mecanismes de decisió no racional dins les explicacions intencionals (vid. caps. 2 i 4)

Descartant les versions més ortodoxes de l'enfocament de l'«elecció racional», cal assenyalar que la defensa de l'atomisme social està vinculada amb la intenció d'adoptar l'explicació intencional per analitzar els fenòmens socials (vid. apt. 1.1.2.). En aquests models d'«home racional», s'entén que els individus només difereixen entre ells pels interessos que es deriven de les situacions en què es troben, així com per la forma com persegueixen aquests interessos<sup>77</sup>. Però més enllà d'aquestes possibilitats de variació, es manté l'interrogant -ja apuntat- sobre si la vida social és reductible a l'anàlisi de les accions dels individus, o bé sí aquestes accions estan afectades per condicions i determinants socials, moltes de les quals no són possibles de derivar analíticament de la interacció individual. En aquesta disputa, alguns autors matisen l'adopció de l'atomisme social, tot defensant la perspectiva d'analitzar les relacions socials com

*"Structures within which actors, individual and collective, deliberate upon goals, perceive and evaluate alternatives, and select courses of action"*<sup>78</sup>

El problema que apareix en defensar exclusivament aquesta perspectiva de «filtres» explicatius és que s'assumeix parcialment un excessiu determinisme social, difícil d'integrar amb els principis del model d'actor racional<sup>79</sup>: s'infereix que les condicions socials poden sobredeterminar interessos i formes de consecució d'aquests per part dels actors. En aquesta situació es pot suggerir una hipòtesi de caire metodològic: *quan més filtres i condicionants socials s'aconsegueixin evitar per la via de l'anàlisi de les interaccions individuals, menys deterministes seran els*

---

<sup>77</sup> R. Hindess, *Choice, Rationality...*, op. cit., 1988, pàg. 35.

<sup>78</sup> A. Przeworski, *Capitalism and Social Democracy*, Cambridge University Press, Cambridge, 1985, pàg. 96.

<sup>79</sup> La qüestió queda oberta aquí, ja que al darrer capítol s'intenten analitzar les possibilitats de vinculació, més enllà de la teoria dels dos cercles que proposa J. Elster ("Anomalies of rationality.." *art. cit.*, 1979), per acceptar un cert grau substancial -però no determinista- de «conformació» del comportament i les institucions per les relacions socials, en aquest sentit, vid. M. Granovetter, "Economic Action and Social Structure: The Problem of Embeddness", *American Sociological Review*, vol. 91, 1985.

*condicionaments socials que cal imposar sobre els models d'actors individuals* (i no cal utilitzar un model d'home idèntic en cada perspectiva). De la hipòtesi se'n pot deduir que, encara que la perspectiva d'emprar models d'home implica un atomisme social, no cal que aquest sigui radical, sinó que ha d'estar obert a l'anàlisi d'efectes no intencionals produïts a partir d'interdependències entre individus, així com a la introducció -potser temporal- d'altres elements teòrics que redueixin les indeterminacions<sup>80</sup>. No sembla necessari concloure aquest apartat amb una crítica antropològica als postulats de l'«home oeconomicus», perquè ja ha estat força desenvolupada en diversos llocs<sup>81</sup> i, en tot cas, aquesta perspectiva reflectiria una altra forma d'aproximació al problema d'explicar el comportament humà. Els elements que defineixen la racionalitat formal són tractats a l'apartat següent, en el qual s'especifiquen les relacions que es pressuposen als ordres de preferències, el punt clau que permet distingir entre racionalitat i interès propi<sup>82</sup>.

### 1.2.2. La consistència en la «gestió» de les preferències.

Seguint la caracterització proposada per M. Hollis, el tercer principi bàsic del model d'home racional -i més important- és precisament la realització d'eleccions racionals. El principi de comportament racional, sense cap orientació d'interès

---

<sup>80</sup> Això implicaria acceptar un reduccionisme moderat, en el sentit que proposa M. Bunge defensant un materialisme emergent, de base individualista, però capaç d'integrar un entorn no explicable necessàriament a través de la perspectiva individualista, "Si se ignora el ambiente o entorno de una cosa no se la podrá explicar adecuadamente, especialmente si la cosa es un ser humano. Semejante descuido del ambiente es uno de los motivos del fracaso de la economía neoclásica (...) En general, no puede darse cuenta de los actos de una cosa si no se tiene en cuenta el estado en que se encuentra el sistema (o los sistemas) de que forma parte la cosa en cuestión" (a Mente y Sociedad, Alianza, Madrid, 1989, pàg. 94).

<sup>81</sup> Entre molts altres, es poden esmentar les crítiques als seus principis de M. Hollis i E. Nell (Rational Economic Man, Cambridge University Press, Cambridge, 1975) o A. Rosenberg ("If Economics isn't a Science, What is It?", The Philosophical Forum, vol. 14, 1983), fins a la discussió de les seves tautologies (per exemple, K. McCornick, "An Essay on the Assumption of Rational Behavior in Economics", Review of Social Economy, vol. 47, 1989).

<sup>82</sup> Com diu B. Hindess, "What matters for the rational choice model is that all preferences are comparable, not that they are narrowly self-interested" a Choice, Rationality..., op. cit., 1988, pàg. 34.

prèviament definida, es refereix essencialment a una forma concreta de realitzar eleccions. En general, es pot considerar que la realització d'eleccions té dos «moments» claus, sovint interconnectats. Suposant una situació que un individu disposa d'un conjunt d'elecció format per diversos elements (o alternatives), entre els que se n'ha d'escollir un, es considera necessari que en un primer «moment» l'individu estableixi algun tipus de relació entre aquests elements. Seguidament, en un segon «moment», també es considera necessari que l'individu segueixi algun principi per efectuar la seva decisió. Freqüentment, aquest principi serveix per relacionar de forma coherent les relacions prèviament definides sobre els elements del conjunt d'elecció amb l'elecció que efectua.

A partir d'aquest marc conceptual, es presenten els requeriments mínims que una hipòtesi de comportament racional pressuposa sobre la forma de realitzar eleccions (no és una opció única, hi ha autors que consideren diferents requeriments). S'introdueix la noció d'utilitat per facilitar l'exposició de les relacions que són considerades bàsiques a la teoria de la racionalitat, denotant alhora l'estret vincle existent a l'economia entre utilitat i preferència. Considerant només la perspectiva de les preferències ordinals<sup>83</sup>, per el primer «moment d'una elecció racional amb informació completa, i donat un conjunt amb dos elements, A i B, es poden definir les següents relacions (p: prefereix; i: indiferent) per ordenar aquests elements:

A p B : relació de preferència

B p A : relació de preferència

---

<sup>83</sup> La utilitat ordinal és una escala de mesura. La relació entre elements (o corbes d'indiferència) permet proposar relacions d'ordre entre aquests, possibilitant una ordenació dels elements (o combinacions de béns) dels més preferits als menys preferits -amb més o menys «utilitat»- mesurada ordinalment. És una qüestió de facilitat en l'operació, sense tenir en compte les diferències quantitatives. Aquesta caracterització segueix bàsicament la formalització de G. Owen (a *Game Theory*, Academic Press, Nova York, 1982<sup>2</sup>, pàgs. 115-116). Només es presenta la utilitat ordinal per reduir al mínim el nombre de supòsits implícits que hom hauria d'assumir. En el cas de tractar la utilitat cardinal, caldria incorporar el supòsit de la continuïtat (sobre aquesta qüestió veure l'apartat 1.3.2.).



$A \sim B$ , quan ( $A \succsim B$ ,  $B \succsim A$ ): relació d'indiferència

En aquestes relacions, s'entén que si  $A \succ B$ ,  $A$  té més utilitat que  $B$  (per a un individu donat), encara que no és possible saber la diferència d'utilitat entre  $A$  i  $B$ . També es defineixen sovint les següents relacions, de forma complementària, per donar més precisió a la formulació de preferències ordenades racionalment. Donat un conjunt amb tres elements, així com les relacions d'ordre prèviament establertes, tenim

-una relació d'equivalència, si

- i).  $A \sim A$  per a tot  $A$
- ii). si  $A \sim B$ , llavors  $B \sim A$
- iii). si  $A \sim B$ ,  $B \sim C$ , llavors  $A \sim C$

-una relació de transitivitat, si

- i). si  $A \succ B$ ,  $B \succ C$ , llavors,  $A \succ C$
- ii). si  $A \succ B$ ,  $B \sim C$ , llavors  $A \succ C$
- iii). si  $A \sim B$ ,  $B \succ C$ , llavors  $A \succ C$

Amb aquestes relacions es poden establir ordres de preferència complets (amb possibilitat d'indiferència) sobre combinacions de béns. Si l'individu pot ordenar totes les combinacions d'eleccions, es diu que les preferències són completes o totals (s'entén que s'hi apliquen totes les relacions definides) sobre el conjunt d'elements existent. Experimentalment es pot observar que no sempre es satisfan totes les relacions en els ordres de preferències, ja que existeixen diversos problemes en la forma de decisió i limitacions físiques i conceptuals en la percepció humana (pel tractament d'aquest problemes en situacions de risc, vegeu l'apartat 1.3.3.). No obstant això, força autors consideren que tenir preferències completes i transitives

és un requisit per considerar l'existència de comportament racional. Però també es pot definir com a condició mínima de racionalitat només l'existència de consistència interna en l'aplicació d'un principi d'elecció (referent al segon «moment»), sense considerar la forma com s'ordenen les preferències<sup>84</sup>.

Sobre aquestes relacions hi ha una condició al segon «moment» d'una elecció que molts autors defensen que s'ha d'acomplir perquè l'elecció pugui ser considerada racional: la consistència interna<sup>85</sup>. Aquesta condició implica que totes les preferències han de ser homogènies, o sigui, comparables. Però potser més apropiadament, caldria caracteritzar la consistència com un requeriment per excloure les eleccions que moltes modelitzacions d'«home racional» consideren irracionals, argumentant que apareixerien moltes formes de maximitzar una elecció respecte a una relació binària (sense disposar de cap criteri de determinació). Donats dos elements, A i B, la consistència mínima en una elecció entre els dos elements s'entén formalment com<sup>86</sup>:

Si  $\{A\} = C(\{A,B\})$ , llavors, per tots els conjunts possibles  $S = \{A,B,\dots\}$ ;  
 $\{B\} \neq C(S)$ .

Aquest axioma, conegut també com l'axioma d'independència (a les situacions amb informació incompleta), s'imposa per justificar que la incorporació

---

<sup>84</sup> Entre les dues posicions hi ha possibilitats intermèdies per analitzar els ordres de preferències, que són formalitzables mitjançant altres relacions, com la quasi-transitivitat (una aproximació a aquestes possibilitats es pot trobar a S. Barberà, "Alguns models de comportament racional en economia", a R. Marimon i X. Calsamiglia, *Invitació a la teoria econòmica*, Ariel, Barcelona, 1991).

<sup>85</sup> Així ho descriu A. Sen, "Rationality of behavior is identified with a requirement that choices from different subsets should correspond to each others in a cogent and systematic way" (a "Rational Behavior", art. cit., 1987, pàg. 69).

<sup>86</sup> Segons la definició de R. Sugden, a "Why Be Consistent?. A Critical Analysis of Consistency Requirements in Choice Theory", *Economica*, vol. 52, 1985. Aquest requeriment significa que si s'escull l'element A en un conjunt format per dos elements, A i B, llavors, en qualsevol altre conjunt en el qual també hi siguin presents A i B (junt a altres elements), un individu, per a ser consistent, no escolliria l'element B.

d'elements externs en una elecció no fa alterar una elecció prèvia (B no és escollit en cap conjunt on siguin A,B, així com altres elements). La justificació d'aquest axioma com un component de l'elecció racional ha estat discutit, tant experimentalment com normativament, perquè es difícil defensar que sigui una propietat intuïtiva de la racionalitat (sobretot degut a la negació de la influència de tercers elements en una elecció).

La idea del comportament maximitzador, entès com un principi d'elecció dels individus, s'associa generalment a la visió clàssica de la racionalitat instrumental. Aquesta concepció implica que el subjecte compara, sota els mateixos estàndards, diversos resultats, per valorar-los amb els mateixos criteris, escollint finalment el curs d'acció que el porta a un resultat «òptim» en relació a uns objectius definits prèviament. O sigui, que amb els criteris establerts, cal que no existeixi un altre comportament possible que aconseguixi millors resultats<sup>87</sup>. Tot i que la teoria econòmica ha defensat tradicionalment la unicitat de la funció de maximització com a variable explicativa; el principi d'elecció, perquè una elecció sigui considerada racional, no ha de ser necessàriament per maximitzar un resultat, ja que entre altres principis possibles, pot ser minimitzar -el malestar, o el dany ecològic, per exemple-, o per igualar (la riquesa, el fruit d'un treball conjunt, etc.). També es podrien aplicar altres principis aliens a la noció d'optimització, com l'elecció de valors mitjans. A més, cal recordar que tots els principis que s'adopten poden ser

---

<sup>87</sup> La noció d'optimització és un dels instruments centrals de la teoria econòmica, essent també compartit per altres camps científics. Aquesta noció té el seu origen en les aplicacions matemàtiques desenvolupades per la física del segle XVII, destacant la creació del càlcul infinitesimal per Leibniz i Newton. Vid. P. Koslowski, "Economic Principle, Maximizing, and the Co-ordination of Individual in Economics and Philosophy", a P. Koslowski (Ed.), *Economics and Philosophy*, J.C.B. Mohr, Tubinga, 1985. Potser la crítica més important que es pot fer de l'aplicació d'aquest principi al comportament dels individus humans parteix de destacar la seva incapacitat neuropsicològica d'acomplir els requisits que implica en moltes situacions d'elecció. Com destaca H. Simon, "*Human short-term memory can hold only a half dozen chunks, an act of recognition takes nearly a second, and the simplest human reactions are measured in tens and hundreds of milliseconds, rather than microseconds, nonoseconds, or picoseconds. These limits are among the most important invariants of intelligence*" (a "Invariants of Human Behavior", *Annual Review of Psychology*, vol. 41, 1990, pàg. 17). Les possibilitats d'incorporar les conseqüències d'aquestes limitacions en els models que empenen els supòsits de maximització -mitjançant probabilitats, introduint restriccions, etc.-, són discutides en els següents capítols, en connexió amb les propostes de variació dels models de la teoria de jocs.

independents en relació al propi interès, o en relació a altres ordenacions de preferències<sup>88</sup>.

No obstant la independència formal entre els dos «moments» d'un procés d'elecció, per caracteritzar el comportament racional, a la teoria econòmica s'han proposat dues grans vies per postular la connexió entre una elecció internament consistent i el conjunt d'elements amb relacions internes específiques sobre el qual s'efectua l'elecció. Per un costat, hi ha les propostes de vincle lògic, en les que s'entén que les preferències -les relacions entre el conjunt d'elecció- s'estableixen a priori mitjançant axiomes als quals s'afegeix un criteri sobre l'orientació d'interès (i regeix un principi d'optimització per la funció d'elecció). D'altre costat, les propostes que defensen un nexa causal entre elecció i preferències consideren que un ordre de preferències es pot deduir de les eleccions, i que aquestes s'han de determinar empíricament, de forma externa -com el seu efecte observable<sup>89</sup>. *Cap de les dues vies no està exclosa de problemes. Mentre la via de connexió lògica presenta problemes d'aplicació externa, la via causal té problemes de consistència interna. Ambdós problemes s'analitzen seguidament.*

Sobre la via «lògica», un element conflictiu és el supòsit de l'existència d'ordres de preferències complets, malgrat que la seva formació no és assumida per l'explicació teòrica. Amb aquest punt de vista, la racionalitat del conjunt de creences dels individus no és un tema de discussió, perquè tan sols es vol un model d'home que relacioni consistentment les seves preferències -ordenades completament- amb

---

<sup>88</sup> Un article clàssic en aquest sentit és el de R. Scott, "Avarice, Altruism and Second Party Preferences", *Quarterly Journal of Economics*, vol. 81, 1972.

<sup>89</sup> Com argumenta A. Hadari, "The parameters defining economic behavior should be correlated as causes and effects, and so lead to the discovery of causal economic mechanisms. Preferences are independent variables, accessible to empirical determination ex post to consumer behaviour; thus preference explanations are probabilistic causal explanations" (a "What are Preference Explanations?...", art. cit., 1987, pàg. 344).

les seves eleccions i accions<sup>90</sup>. Alternativament, les incoherències entre preferències i eleccions també poden ser enteses com fallides en la «gestió» de les preferències (incompliments de les relacions d'ordre esmentades), cosa que serveix per a proposar explicacions sobre freqüents comportaments individuals que s'aparten del model normatiu, alterant segons els temes els supòsits que s'apliquen sobre la racionalitat individual<sup>91</sup>. Ara bé, alguns autors suggereixen que -seguint aquesta via «lògica»- les preferències també s'haurien de formar racionalment, amb la qual cosa s'arribaria a la perspectiva de les teories àmplies de la racionalitat (vid. apt. 1.1.1.):

*"Rational action, then, involves three optimizing operations: finding the best action, for given beliefs and desires; forming the best grounded belief, for given evidence; and collecting the right amount of evidence, for given desires and prior beliefs. Here, desires are the unmoved movers..."<sup>92</sup>.*

Aquesta referència serveix per assenyalar les dificultats de la via «lògica» per explicar l'aparició de preferències que no segueixen les pautes d'una ordenació completa de preferències: quin sentit té proposar axiomàticament altres formes d'ordenar preferències? En el fons, o bé es traspassa el problema de la formació de preferències a l'anàlisi de les creences, o bé no es pot trobar un criteri definitiu per

---

<sup>90</sup> Altres autors critiquen la manca de realisme de la representació de les preferències en un estat estacionari, tot i entenen que és més realista considerar un procés de formació continuada al llarg del temps, on l'individu altera el seu ordre en experimentar amb el consum. Un aspecte a tenir en compte són els treballs dels psicòlegs sobre les motivacions per impulsos o emocions, que estan al darrera de les eleccions entre béns (per exemple, G. Ainslie, "Rationality and the Emotions, a Picoeconomic Approach", Social Science Information, vol. 24, 1985 o J. Elster, "Sadder but Wiser? Rationality and the Emotions", Social Science Information, vol. 24, 1985). Per alguns economistes, aquestes crítiques no contradiuen necessàriament la teoria econòmica, ja que es pot considerar que un bé és una combinació de característiques vinculades amb els impulsos fonamentals dels individus. Definint un ordre de preferències sobre aquestes característiques, junt a les creences sobre la presència d'aquestes característiques en cada bé per part de l'individu, elaboren un ordre de preferències sobre béns (vid. K. Lancaster, Consumer Demand: A New Approach, Columbia University Press, Nova York, 1971).

<sup>91</sup> Vegeu especialment els treballs de J. Elster, Ulysses and the Sirens. Studies in Rationality and Irrationality, (ed. revisada) Cambridge University Press, Cambridge, 1984; i Sour Grapes... op.cit., 1983.

<sup>92</sup> J. Elster, "When Rationality Fails", a Solomonic... op. cit., 1989, pàg. 4.

justificar una o altra proposta. Per exposar amb més detall aquest problema hom pot fer atenció als processos de canvi de preferències que són explicats per modificacions de creences i desitjos<sup>93</sup>, siguin aquestes afectades pels contextos, o bé alterades per la pròpia voluntat<sup>94</sup>. En l'alteració de les preferències per la pròpia voluntat, hom pot introduir la qüestió: és la voluntat, amb les característiques pròpies de la racionalitat, en imposar canvis de creences i desitjos que al seu torn modifiquen les preferències, la que provoca fallides en el comportament racional? Díficilment es pot deixar d'afirmar que aquests casos són possibles, i que quan es produeixen s'han d'integrar a l'anàlisi descriptiva. Entre els processos més destacats de canvi de preferències que violen algunes característiques l'elecció racional -quan els desitjos distorsionen les preferències-, es pot mencionar la debilitat de la voluntat ('Weakness of will'): quan (considerant tots els factors) X és preferit a Y, però s'escull Y en realitat, degut a que encara que X es preferit per l'interès de la persona a llarg terme, Y es un impuls a curt terme (deixar de fumar es millor a llarg terme, però impulsos a curt terme ho impedeixen sovint)<sup>95</sup>. També l'excés de voluntat ('Excess of will'), quan si es fa Z intencionalment perquè es produeixi W, llavors W no s'obté; i l'únic camí es fer Z sense el propòsit d'obtenir W (el cas de proposar-se dormir és un exemple paradigmàtic d'aquesta situació)<sup>96</sup>. En aquesta darrera situació, es podria destacar que és més convenient no adoptar el comportament racional, o que sentit comú i racionalitat entren en contradicció<sup>97</sup>.

---

<sup>93</sup> Per exemple, C. von Weizsäcker proposa, per aplicar la teoria de consum, l'endogenització de les preferències (enteses con el gustos -desitjos) amb un model que fa aquestes deponents del consum i nivell de benestar passat (cf. "Notes on Endogenous Change of Tastes", Journal of Economic Theory, vol. 3, 1971).

<sup>94</sup> El paper de la voluntat en els canvis de preferències és tractat per J. Elster a Ulysses and the Sirens...op.cit., 1984, especialment en la discussió del mecanisme de lligar-se un mateix.

<sup>95</sup> J. Elster, "Weakness of Will and the Free-Rider Problem", Economics and Philosophy, vol. 1, 1985.

<sup>96</sup> J. Elster, Sour Grapes...op.cit., cap. segon, 1983.

<sup>97</sup> La teoria de la dissonància cognitiva explica la reformulació de desitjos i creences per reduir les tensions psíquiques derivades de la formulació d'objectius que esdevenen (pel temps o per major informació) inabastables. D'altra banda, en condicions d'incertesa (vid. apt. 1.3.) les distorsions degudes als efectes de presentació o a les dificultats de mantenir les propietats d'inferència racional en l'observació,

Per discutir els problemes de la via «causal» de connexió entre elecció i preferències, es presenta com exemple un problema de la teoria de la «preferència revelada», que l'economia neoclàssica va desenvolupar com un intent de fonamentar una teoria de l'«elecció racional» no vinculada a supòsits que inevitablement impliquessin discutir la formació de preferències, evitant també el recurs a la utilització de la noció d'utilitat. La teoria de la «preferència revelada» pretenia ser una forma alternativa de justificar les preferències a la teoria de la «demanda del consumidor»<sup>98</sup>. La teoria parteix del supòsit que els individus apliquen el principi d'optimització per relacionar les seves preferències amb les seves eleccions: s'elegeix sempre el conjunt de béns que té la posició òptima en un ordre de preferències, o sigui, la corba d'indiferència que té per tangent la funció de restricció pressupostària (punt en que el consumidor òptimitza la seva funció d'utilitat -que representa el seus interessos)<sup>99</sup>. El nucli de la teoria de la preferència revelada és l'axioma feble de la preferència revelada: "Suposant que A i B són dues combinacions diferents de béns, si A es revela com preferida a B, B no es revela mai com a preferida a A" (la preferència es revela en el moment de produir-se l'elecció). Així, la combinació B només podrà ser escollida quan la combinació A no sigui assequible, per no produir violacions d'aquest axioma. Diversos autors han proposat demostracions en les que sí s'assumeixen els supòsits de maximització (en

---

també poden portar a eleccions inconsistentes (vid. E. Aronson, *El animal Social*, Alianza, Madrid, 1979). Per un exemple d'aplicació, vid. G. Akerlof i W. Dickens, "The Economic Consequences of Cognitive Disonance", *American Economic Review*, vol. 72, 1982.

<sup>98</sup> Un dels fonaments de la microeconomia moderna, que explica com intervenen els individus en el mercat per adquirir productes. Vid. P. Samuelson, "Consumption Theory in Terms of Revealed Preference", *Economica*, vol. 15, 1948. Cal destacar com a precedent d'aquest enfoc el treball de J. Hicks i D. Allen "A Reconsideration of the Theory of Value", a *Economica New Series (Econometrica)*, vol. 1, 1934.

<sup>99</sup> La teoria suposa que l'individu és capaç d'evaluar correctament totes les alternatives disponibles en termes de les seves preferències. La corba de demanda (respecta un bé) del consumidor es construeix suposant tots els preus i factors constants, excepte el preu d'un bé X, i observant per cada variació del preu d'X, amb quina quantitat (elecció) el consumidor maximitza el seu ordre de preferències (aconsegueix un ordre més elevat), donada la seva restricció pressupostària. Per una aproximació moderna a la teoria de la preferència revelada, vid. H. Varian, *Análisis Microeconómico*, Ed. Bosch, Barcelona, 1986. Vid. A. Mas-Colell, "On Revealed Preference Analyses", *Review of Economic Studies*, vol. 44, 1977, per una relació entre la perspectiva de la maximització de preferències amb l'anàlisi de la preferència revelada.

una restricció pressupostària), si hi ha insaciabilitat i si es compleix la llei de decreixement de la relació marginal de substitució, llavors les eleccions compleixen amb l'axioma dèbil de la preferència revelada<sup>100</sup>. L'axioma fort de la preferència revelada és l'extensió de l'axioma feble pel cas de 'n' béns.

La crítica d'A. Sen<sup>101</sup> al concepte de preferència revelada destaca fonamentalment la seva dependència del concepte d'utilitat, cosa que diversos dels fundadors d'aquesta teoria varen negar<sup>102</sup>, i que presentaren com una nova perspectiva deslligada de supòsits psicològics. A. Sen argumenta que el concepte de «preferència revelada» no és el mateix que el de preferència, ja que el primer incorpora axiomàticament el supòsit de consistència en les eleccions, sense justificar-ho. I cal afegir que la inclusió de la consistència prové d'una intuïció sobre les eleccions, basada en idees implícites sobre el comportament racional, que han estat força qüestionades experimentalment<sup>103</sup>. D'aquesta interpretació crítica proposada per A. Sen se'n pot derivar que la consistència en les eleccions no serveix per explicar completament les preferències: *si l'ordenació de les preferències es dedueix automàticament de les eleccions, no hi ha forma de comprovar realment que les preferències segueixen els criteris que defineixen una ordenació concreta prèviament suposada.*

D'aquest problema se'n deriva una pregunta doble pel manteniment del vincle «causal» entre elecció i preferència (el supòsit de què la preferència causa l'elecció d'una determinada forma, i per tant es pot deduir la preferència si es coneix l'elecció). Per un costat, caldria discutir perquè les eleccions reals no són sempre

---

<sup>100</sup> H. Varian, Microeconomia intermedia, Editorial Bosch, Barcelona, 1987, cap. 7.

<sup>101</sup> A. Sen, "Behaviour and the Concept of Preference", Economica, vol. 40, 1973 (també a Choice, Welfare...op. cit., 1983).

<sup>102</sup> Per exemple, J. Hicks, A Revision of Demand Theory, Clarendon Press, Oxford, 1956.

<sup>103</sup> Vid. apt. 1.3.3., i també R. Sugden, "Why be Consistent...", art. cit., 1985.



consistents (coneixent l'ordre de preferències i el principi d'elecció); i d'altre, si en aparèixer eleccions inconsistentes es podria deduir que s'han produït modificacions dels ordres de preferències. En aquest darrer cas, les relacions de preferència que causen les eleccions observades -potser només aparentment inconsistentes- no poden ser determinades seguint la teoria de la preferència rellevada, donat que en les eleccions hi poden haver influències d'altres preferències que no hagin derivat en eleccions (no revelades). En el mateix sentit, també es podria considerar que les eleccions s'estableixen a partir d'altres relacions causals amb les preferències, com mecanismes d'inferència formats a través de complexos problemes psicològics<sup>104</sup> i, per tant, sense integrar les relacions de consistència de la funció d'elecció. No obstant això, en aquestes circumstàncies, no es disposa de cap criteri per aplicar -la manca de consistència pot ser deguda a molts factors- una determinada relació causal que a partir de les eleccions expliqui les preferències. En tot cas, aquesta relació hauria de ser a l'inrevés.

---

<sup>104</sup> A. Sen, "Rational Fools...", art.cit., 1977.

### 1.3. Risc i racionalitat: la Hipòtesi de la Utilitat Esperada.

#### 1.3.1. Gènesi i evolució del concepte d'utilitat esperada.

Aquest apartat, que presenta un tractament més formal que els anteriors, està dedicat a discutir els problemes de les eleccions racionals en contextos d'informació incompleta; on per tant, s'ha d'assumir un cert risc sobre el resultat de les eleccions. La discussió, que s'articula entorn la noció d'utilitat, serveix com a exemple dels problemes sobre racionalitat i preferències esmentats fins el moment, així com d'alguns dels intents que -des de dins del propi paradigma- es realitzen per a solucionar-los. Una indicació de la importància del concepte d'utilitat la proveeix N. Georgescu-Roegen, quan senyala en iniciar el seu article sobre la utilitat a l'Enciclopèdia Internacional de les Ciències Socials que

*"In a broad perspective the history of economics emerges as a struggle with the problem of value. The importance of the concept known in modern economics by the name of utility derives precisely from the great light it shed upon this problem from the outset"*<sup>105</sup>

A partir d'una antiga idea dels economistes clàssics anglesos, sobre l'existència de mercaderies que tenien un valor determinat en gran part per la seva escassetat<sup>106</sup>, es va produir a partir de 1870 una revolució al camp de la teoria econòmica. Aquesta, anomenada la 'revolució marginalista', va desenvolupar una aproximació als problemes econòmics que tenia com a peça clau la demanda de bens escassos, front a l'anàlisi dels problemes de la producció que caracteritzava als

---

<sup>105</sup> N. Georgescu-Roegen, "Utility", a D. Sills (Ed.), International Encyclopedia of the Social Sciences, vol. 16, Macmillan i Free Press, Nova York, 1968, pàg. 236.

<sup>106</sup> Vegeu D. Ricardo, Els principis d'economia política i tributació, edicions 62, Barcelona, 1984, pàg. 34.

economistes clàssics. Aquesta nova perspectiva, en centrar-se als intercanvis entre individus, tenia una necessitat molt més gran que la perspectiva clàssica de disposar d'una teoria de l'acció. Sense necessitat de detallar les interpretacions, debats i polèmiques per trobar un fonament teòric del comportament individual, hom pot destacar que el propòsit constant ha estat desenvolupar una teoria sobre la forma en què els individus racionals realitzen les seves decisions, en un entorn donat de béns escassos i possibilitats d'acció alternatives. Per aquest propòsit, era imprescindible definir una escala de mesura, sobre la qual aplicar posteriorment criteris de decisió. Per a tal fi, es va recuperar i refonamentar la concepció de satisfacció individual desenvolupada pels utilitaristes a principis del segle XIX: la utilitat<sup>107</sup>. L'ideal programàtic de la teoria de la utilitat en J. Bentham, força diferent als actuals, considerava que cada individu podia mesurar de forma cardinal la utilitat de cada mercaderia<sup>108</sup>; que es el mateix que dir que existia un estàndard d'utilitat comú a tots els individus<sup>109</sup>. Precisament, la discussió d'aquest principi de Bentham va constituir el gran debat sobre la formulació de la noció d'utilitat durant la segona meitat del segle XIX, que va conduir al desenvolupament del principi d'«utilitat marginal».

Cap als anys 1870s, els marginalistes (especialment Menger, Jevons i Walras, que arribaren a posicions similars de forma independent) definiren la «utilitat

---

<sup>107</sup> La revolució marginalista va rellançar el concepte, poc definit, perquè es simplificava molt la teoria de la demanda. La idea que la demanda es modifica segons els preus no es va imposar fins ci.1870, i això permet explicar perquè amb anterioritat no tenia sentit pels economistes una teoria subjectiva del valor. Sobre aquesta transformació de la teoria econòmica, vegeu L. Pasinetti, Cambio estructural y crecimiento económico, Pirámide, Madrid, 1985, pàgs. 20-37.

<sup>108</sup> En paraules de Bentham, "*Sum up all the values of all the pleasures on the one side, and those of all the pains in the other. The balance, if it be on the side of pleasure, will give the good tendency of the act upon the whole, with respect to the interest of that individual person; if on the side of pain, the bad tendency of it upon the whole*" (cf. J. Bentham, "An Introduction to the Principles of Morals and Legislation", extret de A. Page, (Ed.), Utility Theory. A Book of Readings, Wiley, Nova York, 1968).

<sup>109</sup> Com a tal, s'entenia que aquest estàndard estava fora dels individus, i que aquests només l'apreciaven "*By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good or happiness (...) or (...) to prevent the happening of mischief, pain, evil or unhappiness to the party whose interest is considered..*" (cf. J. Bentham, Ibidem., pàgs. 3-4).

marginal» ressuscitant i modificant el concepte d'utilitat de J. Bentham -poc utilitzat als cinquanta anys anteriors- per mesurar la satisfacció derivada del consum d'un bé. Entenien el concepte d'utilitat amb un sentit subjectiu, que podia variar de persona a persona en relació al mateix bé, i no com la mesura d'una qualitat intrínseca del bé<sup>110</sup>. Així, la subjectivitat en la valoració de la utilitat, semblava l'aproximació més plausible per entendre la relació entre persones i mercaderies, propera a la concepció de «valor d'ús», i d'alguna forma, contraposada a la concepció de valor com a «valor treball». Aquesta concepció els va proporcionar un punt de partida per a la seva teoria de l'adopció de decisions per part del consumidor. La clau de l'aproximació era el concepte d'«increment d'utilitat» -infinitesimal- que els permetia postular -i tractar matemàticament- la seva proposició teòrica bàsica: la «lei de la utilitat marginal decreixent» que assenyalava que a partir d'un cert nivell de consum de qualsevol bé, disminuïa la «utilitat marginal» que proveïa aquest<sup>111</sup>. A partir d'aquesta llei, es definiren unes funcions d'utilitat com a corbes de consum (Demanda) i es suposà que s'arribava a l'equilibri quan les relacions entre utilitats marginals de diferents béns eren iguals a les relacions entre preus<sup>112</sup>.

El problema d'aquest enfocament era que, tot i estar centrat en una escala de mesura, presentava força dificultats per mesurar realment la utilitat marginal decreixent, o la utilitat total que aportaven els béns. Per resoldre aquest problema, al llarg de força anys molts economistes han proposat diversos procediments de mesura sobre la utilitat, intentant relativitzar la dependència amb un índex de

---

<sup>110</sup> Vid. N. Georgescu-Roegen, "Utility", art. cit., 1968.

<sup>111</sup> No els importava tant com a Bentham definir que era la utilitat, sinó assumir l'existència d'una mesura de la intensitat de la utilitat (en relació a la quantitat de mercaderia), i poder aplicar aquesta mesura a una escala absoluta. Sobre el desenvolupament de la teoria de la Utilitat, G. Stigler, "The Development of Utility Theory", *Journal of Political Economy*, vol. 58, 1950, reimprés a A. Page, (Ed.), *Utility... op. cit.*, 1968; i N. Georgescu-Roegen, "Utility", art. cit., 1968.

<sup>112</sup> També es varen definir les corbes d'indiferència, i es polemitzà sobre la possibilitat o no de sumar utilitats individuals (Cf. G.J. Stigler, "The Development...", art. cit., 1950, pàgs. 77-91). Un element comú de la teoria econòmica ha estat no basar-se en el concepte d'utilitat en termes absoluts, sinó en la utilitat marginal decreixent, que considera la utilitat en termes relatius, amb el supòsit de que cada unitat addicional del mateix bé augmenta la seva utilitat per l'individu consumidor en una magnitud decreixent.

satisfacció individual<sup>113</sup>. Amb el temps, es va perfeccionar molt l'instrumental analític de la teoria de la «utilitat marginal», millorant la presentació de l'esquema bàsic de l'enfocament, fins arribar a les formulacions dels anys cinquantes i seixantes sobre la teoria de la decisió sota risc, ja plenament formalitzada i assimilada com a base d'amplis camps de la microeconomia<sup>114</sup>. En aquest context es va desenvolupar la teoria moderna de la utilitat, centrada en l'anàlisi de l'estructura de preferències del consumidor i la forma com aquesta estructura influeix sobre les seves eleccions, sense necessitat d'una mesura cardinal directa de la utilitat, sinó tan sols d'una ordenació d'aquesta (sense una escala de mesura absoluta).

Aquesta teoria moderna de la utilitat estableix uns axiomes sobre les relacions de preferències, que al terreny de les eleccions racionals en situacions de risc ha conduït a la formulació de la Hipòtesi de la Utilitat Esperada. Aquesta hipòtesi proposa una explicació axiomàtica de com es formen les preferències sobre cursos alternatius d'acció amb resultats incerts<sup>115</sup>. Abans d'entrar a les arrels conceptuals de la Hipòtesi, cal assenyalar la particular pretensió de la mateixa, a voltes descriptiva i a voltes prescriptiva. La pròpia ambigüïtat que genera el seu camp d'aplicació -que la pròpia teoria, i les seves diverses versions no defineixen- ha

---

<sup>113</sup> La crítica de I. Fisher sobre el confusionisme de les diverses concepcions de la utilitat destacava la necessitat d'un nou concepte que aportés claredat a l'ús que se'n feia a la teoria econòmica. I. Fisher, "Is 'Utility' the Most Suitable Term for the Concept it is Used to Denote?", American Economic Review, vol. 8, 1918, reimprès a A. Page, (ed.), Utility...op. cit., 1968. Vid. també l'apt. 1.3.2.

<sup>114</sup> Una panoràmica de la situació a finals dels anys seixantes és exposada per P. Fishburn, a "Utility Theory", Management Science, vol. 14, 1968. L'autor defineix el conjunt de teories de la utilitat com aquelles que, definint unes relacions de preferència dels individus, realitzen un conjunt d'assumpcions -internament consistent- sobre com els individus assumeixen aquestes relacions. La teoria de la decisió sota risc es refereix a l'anàlisi de l'adopció de decisions per part d'individus que no disposen de tot la informació necessària per valorar els possibles resultats de les seves eleccions. Però es considera que aquests individus sí disposen de valors probabilístics sobre l'ocurrència d'aquests resultats. Quan aquestes probabilitat tampoc són disponibles, es considera que els individus decideixen en condicions d'incertesa.

<sup>115</sup> En aquest sentit, aquestes formulacions axiomàtiques segueixen les propostes de vincle lògic entre racionalitat i preferències (vid. apt. 1.2.2.), i pateixen -per tant- els problemes ja esmentats d'aplicació externa. Per una presentació formal de la Hipòtesi, vegeu l'annex 1.3.A..

portat a una àmplia discussió entorn les seves propietats, sense que moltes vegades s'explicités clarament quin era l'àmbit d'aplicació pel qual es definia una proposta. Com a teoria descriptiva, el que pretén la Hipòtesi de la Utilitat Esperada és explicar com els individus estableixen preferències i fan eleccions en situacions de risc; mentre que com teoria prescritiva, la teoria mostra com s'han de prendre les decisions en situacions de risc per obtenir els millors resultats.

A partir d'unes paradoxes clàssiques (com la paradoxa de Sant Petersburg, resolta per D. Bernoulli) sobre el comportament dels individus en situacions de risc<sup>116</sup>, durant els anys quarantes es varen desenvolupar les bases d'allò que seria la teoria axiomàtica de l'elecció racional front al risc, a partir del treball de J. von Neumann i O. Morgenstern sobre teoria de jocs. Posteriorment, durant els anys cinquantes es va portar a terme una tasca de refinament -més matemàtica que conceptual- sobre aquesta formulació teòrica. Durant els anys seixantes, ja semblava un camp molt estable, amb uns fonaments ben axiomatitzats i amb una creixent aplicació en diferents camps de la ciència econòmica. Especialment a partir dels anys setantes aquests fonaments es varen començar a posar en dubte, com a fruit d'una abundant producció de treballs d'economistes i psicòlegs que assenyalaven l'escassa plausibilitat empírica dels axiomes de la Hipòtesi de la Utilitat Esperada, mostrant que aquests no reflectien el comportament humà en situacions de risc (l'exposició de la Hipòtesi i dels seus problemes es realitza als apts. 1.3.2. i 1.3.3.)<sup>117</sup>. Les crítiques varen ser tan fonamentals i motivades per diversos aspectes que, actualment ja és majoritàriament acceptada entre els especialistes la necessitat de procedir a

---

<sup>116</sup> L'article original "Exposition of a New Theory on the Measurement of Risk" (1728) es reproduït (en traducció del llatí a l'anglès) a *Econometrica*, vol. 22, 1954. La paradoxa era que en contraposició a allò que la lògica matemàtica assenyalava, la gent decidia normalment no seguir l'esperança matemàtica d'un joc d'atzar. D. Bernoulli va proposar per caracteritzar l'elecció humana, ponderar l'esperança matemàtica per una funció logarítmica que reflectís l'aversion al risc de les persones. Per a altres paradoxes clàssiques, veure E. Kauder, *A History of Marginal Utility Theory*, Princeton University Press, Princeton, 1965.

<sup>117</sup> És remarcable l'existència d'algunes crítiques anteriors, com la de M. Allais ("Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine", *Econometrica*, vol. 21, 1953), contemporània a l'emergència de la Hipòtesi de la Utilitat Esperada.

alguna reforma dels axiomes, afectant inclús a les aproximacions prescriptives i normatives. Els anys vuitantes han estat anys d'una forta polèmica, accentuada pel fet que bona part de la teoria econòmica s'havia desenvolupat sota aquests axiomes i, per tant, la modificació de la teories d'elecció sota risc implica qüestionar -i alterar- part dels edificis teòrics de l'economia neoclàssica. La teoria de la decisió sota risc és avui un camp d'intensa circulació de propostes, que abasten gaire bé totes les modificacions lògicament possibles<sup>118</sup>.

La discussió continua durant aquesta dècada, ja amb més detall i concreció, sobre les implicacions i possibilitats de cada proposta, així com sobre els seus defectes. *Probablement no apareixerà una proposta triomfadora que resolgui el problema de les ordenacions de preferències implausibles, sinó que es desenvoluparà un estoig de models per aplicar en situacions i camps teòrics diferents; trencant amb la perspectiva d'un nucli originari i unitari que intentava donar un nexa a tota la teoria social de base individualista i racional.* La modificació dels axiomes que constitueixen la Hipòtesi de la Utilitat Esperada està provocant implicacions molt importants per tot el plantejament metodològic del paradigma. La renúncia a una unicitat axiomàtica planteja nous problemes per la heurística de l'«elecció racional» basada en la connexió lògica entre preferències i eleccions. Apareix per un costat una menor capacitat de predicció (en reduir axiomes o la seva capacitat de determinació), que provoca una parcialitat en els nous plantejaments; i d'altre costat, tot i la modificació dels axiomes, continua sempre present la discussió sobre la capacitat d'aquests com a predictors del comportament observat, un problema

---

<sup>118</sup> Com a panoràmica general i reflexió sobre la situació actual, vegeu P. Fishburn, "Expected Utility: An Anniversary and a New Era", *Journal of Risk and Uncertainty*, vol. 1, 1988. Fishburn destaca que front a l'abundància en els anys setantes d'evidència empírica contra la Hipòtesi de la Utilitat Esperada, a la dècada dels vuitantes s'ha produït una gran abundància de propostes normatives. Acaba assenyalant que les línies generals de les propostes possibles ja són visibles i, per tant, que el que cal esperar a les properes dècades són nombrosos refinaments, aplicacions i anàlisis de les propostes desenvolupades.

general de l'enfocament (vid. apt. 1.1.2.)<sup>119</sup>. No obstant, per delimitar les seves expectatives, cal analitzar amb més detall aquests problemes, cosa que es proposa en els següents apartats.

### 1.3.2. La Hipòtesi de la Utilitat Esperada i el seu vincle amb la noció de probabilitat.

L'origen directe de la Hipòtesi de la Utilitat Esperada és l'obra de J. von Neumann i O. Morgenstern, *Theory of Games and Economic Behaviour*<sup>120</sup>. El seu propòsit era desenvolupar la teoria matemàtica dels jocs d'estratègia, i per tant varen simplificar el problema de la utilitat, formulant una hipòtesi senzilla sobre com es formaven les preferències en situacions de risc. Degut a que aquesta qüestió no era l'objectiu central de la seva recerca, no formularen una definició totalment completa del procés de decisió en la seva proposta<sup>121</sup>. Malgrat tot, la proposta va implicar un nou significat pel concepte d'utilitat, renovat en relació a les polèmiques existents anteriorment sobre la conveniència de considerar l'ordinalitat o bé la cardinalitat en la seva mesura. Constituint el nucli d'una nova línia teòrica, es va convertir en la hipòtesi estàndard de la teoria econòmica per a fonamentar explicativament les decisions sota risc.

---

<sup>119</sup> Tot i que ja queda fora de l'àmbit de la present tesi, es pot indicar la possibilitat de què al camp prescriptiu es produeixi una situació similar; i segons els fins que es defensin, s'adoptin diverses formes de valoració, que amb uns vincles de justificació força complexos normativament, portin a una o altra modelització per a la resolució del problema plantejat. Vegeu E. McClennen, *Rationality and Dynamic Choice. Foundational Explorations*, Cambridge University Press, Cambridge, 1990.

<sup>120</sup> Editat per Princeton University Press, Princeton, 1944. Inicialment el llibre s'havia de titular *General Theory of Rational Behavior* (Cf. O. Morgenstern, "The Collaboration between O. Morgenstern and J. von Neumann on the Theory of Games", *Journal of Economic Literature*, vol. 14, 1976).

<sup>121</sup> En les edicions posteriors del llibre, dels anys 1947 i 1953, els autors afegiren un apèndix precisant algunes qüestions més. No obstant, mai varen definir específicament un dels axiomes més polèmics, el d'independència de les alternatives irrelevants, i això va portar a una gran polèmica posterior sobre la interpretació de la seva teoria.



La dificultat que existia per analitzar aquest tipus de decisions era com valorar la diferència d'utilitat que es produïa entre les distintes alternatives possibles: saber quan la possibilitat de guanyar una alternativa A compensava el risc de què el resultat fos l'alternativa B. La dificultat d'aquest càlcul era que no hi havia cap mercaderia ideal amb la qual poder fer una transformació lineal, que portés a una mesura lineal de la utilitat sobre probabilitats. Von Neumann i Morgenstern no estaven interessats en ressuscitar una noció d'utilitat cardinal, que tantes polèmiques havia portat als decennis anteriors, però volien algun tipus de mesura que permetés tractar numèricament les utilitats, tot i que rebutjaven la «mesurabilitat» de la utilitat i de les comparacions «interpersonals» d'utilitat<sup>122</sup>. Això els va portar a construir una nova teoria, la qual tenia poca vinculació -apart del mot 'utilitat'- amb les teories anteriors sobre les fonamentacions de les preferències<sup>123</sup>. El punt clau de la nova teoria, que incorporava l'existència de risc en les decisions, és la introducció d'una forma especial d'utilitat numèrica:

*"It can be shown that under conditions on which the indifference curve analysis is based very little extra effort is needed to reach a numerical utility"*<sup>124</sup>

---

<sup>122</sup> Una combinació de béns ofereix més utilitat a un consumidor que una altra combinació, perquè és més preferida, sense saber en quina mesura ho és, de més preferida. Perquè la utilitat fos comparable, tindria que existir un origen i una unitat de mesura aplicable a tots els individus.

<sup>123</sup> Tant en relació al nou enfoc ordinalista, com al tradicional enfoc cardinalista. Cal destacar la seva posició respecte a la discutida qüestió de la comparabilitat entre diferències de preferències, introduïda per W. Pareto com una nova via de cardinalitat, i formalitzada per R. Frisch ("A Problem of Pure Economics", a J. S. Chipman et al. (eds.), *Preferences, Utility and Demand*, Hancourt Brace Jovanovich, Nova York, 1971, edició original en francès de 1926). S'ha argumentat que la teoria de von Neumann i Morgenstern recull implícita aquesta noció, tot i no fer-ne cap referència explícita, però com assenyala P. Fishburn, "In fact, if there is any legitimate meaning to a utility difference within the von Neumann-Morgenstern theory without the imposition of further axioms, it is only as an algebraic difference defined from their utility function that has no particular relationship to strength of preference, preference-difference comparisons, or notions of 'a psychological entity measurable in its own right'" (cf. P. Fishburn, "Retrospective on the Utility Theory of von Neumann and Morgenstern", *Journal of Risk and Uncertainty*, vol. 2, 1989 -pàgs. 135-6).

<sup>124</sup> J. von Neumann i O. Morgenstern, *Theory of Games...op.cit.*, 1953<sup>3</sup>, (pàg. 17, apt. 3.3.1.).

Aquesta mesura de la utilitat establia la comparació entre combinacions de successos, i no només comparant successos, com s'havia fet fins aquell moment<sup>125</sup>. Les combinacions de successos es podien entendre com distribucions de probabilitat a través de jocs o previsions, vinculant l'adopció del risc a la idea de loteria (un cop ponderades les probabilitats, es deixava a l'atzar la decisió final). La novetat de l'aproximació de von Neumann-Morgenstern fou precisament la consideració de preferències entre tals entitats<sup>126</sup>, cosa que els va permetre fer comparacions numèriques en la construcció de les seves funcions d'utilitat, tot i no mesurar el nivell de diferència entre les preferències. A partir del supòsit de què qualsevol succés imaginable podia ésser combinat amb probabilitats, von Neumann i Morgenstern proposaren una axiomatització conjunta de les preferències i les probabilitats. Però el concepte tradicional de probabilitat subjectiva davant el risc no els servia, ja que no és constant ni comparable interpersonalment; mentre que sí ho és el concepte de probabilitat com a freqüència a llarg terme<sup>127</sup>.

Aquesta relació entre preferències i probabilitats és basicament la que aplicaren els economistes al llarg dels anys cinquantes com a fonament per una

---

<sup>125</sup> És destacable l'expressió de J. von Neumann quan, junt amb O. Morgenstern varen definir els axiomes de la utilitat esperada: "Ja hat denn das niemand gesehen?" (cf. O. Morgenstern, "The Colaboration of...", art. cit., 1976). Existeix un precedent, F.P. Ramsey, que en 1931 va fer una proposta similar (a "Truth and Probability", dins de R. B. Braithwaite (ed.), F.P. Ramsey, The Foundations of Mathematics and Other Essays, Humanities Press, Nova York, 1950). La proposta de Ramsey no va ser recollida pels economistes, i tanpoc es cita a l'obra de von Neumann i Morgenstern.

<sup>126</sup> J. von Neumann i O. Morgestern, Theory of Games..op. cit., 1953, (pàgs. 18-19, apt. 3.3.2.): "The introduction of numerical measures can be achieved even more directly if use is made of all possible probabilities. Indeed: Consider three events, C, A, B, for which the order of the individual's preferences is the one stated. Let  $\sigma$  be a real number between 0 and 1, such that A is exactly equally desirable with the combined event consisting of a chance of probability  $1 - \sigma$  for B and the remaining chance of probability  $\sigma$  for C. Then we suggest the use of  $\sigma$  as a numerical estimate for the ratio of the preference of A over B to that of C over B", l'elaboració exacta d'aquestes idees, com assenyalen von Neumann i Morgenstern, requereix un tractament axiomàtic (vegeu l'annex 1.3.A.).

<sup>127</sup> Com assenyala M. Allais, "The concept of objective probability can be given an operational expression in that it relates to the concept of empirical frequency, which itself can be subjected to experimental measurement" (a "The Foundations of the Theory of Utility and Risk. Some Central Points of the Discussions at the Oslo Conference", dins O. Hagen i F. Wenstoep (Eds.), Progress in Utility and Risk Theory, Reidel, Dordrecht, 1984, pàg. 12).

teoria de la decisió sota risc, prenent el nom d'Hipòtesi de la Utilitat Esperada. La Hipòtesi vincula les preferències d'individus «racional» sobre alternatives que prometen diversos premis, amb les probabilitats conegudes de cada alternativa (loteries). Així, l'adopció de decisions és dependent -suposant una connexió consistent amb un principi d'elecció- de les preferències sobre les distribucions de probabilitat que es poden aplicar al conjunt de resultats possibles. Quantificant el risc assumit, s'arriba al concepte d'«utilitat mesurable» (o «nova utilitat cardinal»), que es pot utilitzar per explicar el comportament dels individus en realitzar eleccions amb un component aleatori. Al llarg dels anys cinquantes i seixantes aquesta idea va ser molt desenvolupada formalment, proposant-se diverses versions que aclarien les suposicions implícites existents en la proposta de von Neumann i Morgenstern<sup>128</sup>. La polèmica més destacada que va aparèixer en aquell moment voltava entorn l'axioma d'independència<sup>129</sup>, defensat per L. Savage i no definit per von Neumann i Morgenstern explícitament en el seu sistema. Segons Savage i altres autors, l'axioma es podia deduir implícitament de la proposta de von Neumann i Morgenstern. Savage va utilitzar precisament aquest principi com a peça clau d'una teoria de la utilitat esperada que incorporava probabilitats subjectives.

La teoria de Savage suposa l'existència d'ordres de preferències en els individus i assigna índexs d'utilitat a distribucions de probabilitat lineals de forma que conservin l'ordre de preferències de l'individu sobre tals distribucions. L'existència dels índexs d'utilitat i de les probabilitats subjectives es dedueix dels axiomes de la seva teoria, entre els quals forma part el polèmic axioma de la

---

<sup>128</sup> Hom pot destacar, entre d'altres, els treballs de D. Ellsberg, "Classic and Current Notions of 'Mesurable Utility'", *Economic Journal*, vol. 64, 1954 (reimpres a A. Page (ed.), *Utility... op.cit.*, 1968); R. Strotz, "Cardinal Utility", *American Economic Review*, vol. 43, 1953; L. Savage, *Foundations of Statistics*, Wiley, Nova York, 1954; D. Luce i H. Raiffa, *Games and Decisions*, Wiley, Nova York, 1957.

<sup>129</sup> La independència de les preferències és un supòsit restrictiu que implica que no afecten a l'ordre de preferències d'un consumidor les eleccions dels altres consumidors o l'aparició d'altres béns, i només té en compte les combinacions de béns que se li assignen.

independència<sup>130</sup>. Savage suposa que guiat per la seva utilitat, l'individu pot elegir entre distribucions de probabilitat alternatives. Aquesta coincidència porta a la propietat de la utilitat esperada: la utilitat d'una loteria és la utilitat esperada dels seus premis<sup>131</sup>. Es pot destacar especialment que mentre von Neumann i Morgenstern només comptaven amb l'esperança matemàtica de la utilitat cardinal, Savage proposava també considerar la seva distribució sobre la mitjana<sup>132</sup>, i rebutjava, a més, la referència al concepte d'«utilitat cardinal», perquè considerava que implicava un índex de satisfacció de caire «metafísic»<sup>133</sup>.

Per aprofundir en la tensió de moltes discussions d'aquest àmbit, entre la voluntat descriptiva i l'orientació normativa o prescriptiva<sup>134</sup>, es presenta la «paradoxa» d'Allais, un dels referents clàssics de la crítica a la Hipòtesi de la Utilitat Esperada. La qüestió que planteja és si l'axioma d'independència de Savage, que

---

<sup>130</sup> M. Allais critica aquesta teoria destacant que "a basic question is how to determine the subjective probabilities  $p_i$ , and the proponents of Savage's theory in reality provide no operational method for doing this" (cf. M. Allais, "The Foundations..", cit., 1984, pàg. 13).

<sup>131</sup> H. Varian, Análisis Microeconómico, Ed. Bosch, Barcelona, 1987, pàg. 183. En tot cas, com destaca M. Allais, els resultats de la teoria de Savage només poden ésser considerats com regles normatives pel comportament racional, i no com un intent explicatiu del comportament racional.

<sup>132</sup> Vegeu l'annex 1.3.A., i també M. Allais, "The So-called Allais Paradox and Rational Decisions Under Uncertainty", a M. Allais i O. Hagen (Eds.), Expected Utility Hypotheses and the Allais Paradox, Reidel, Dordrecht, 1979, pàg. 598-99.

<sup>133</sup> "It seems mystical, however, to talk about moral worth apart from probability and, having done so, doubly mystical to postulate that this undefined quantity serves as utility", L. Savage, Foundations of Stadiistics, op. cit., 1954, pàg. 94 (citada per M. Allais, "The Foundations ..", cit., 1984, pàg. 91).

<sup>134</sup> Aquesta darrera perspectiva és defensada com "the art and science of decision analysis require many skills not readily classifiable as either normative and descriptive, but which nonetheless have a legitimate and important role in both research and practice", per D. Bell, H. Raiffa i A. Tversky, (a "Descriptive, Normative and Prescriptive Interactions in Decision-Making", a Bell i altres, Decision-Making...op.cit., 1988, pàg. 30). P. Schoemaker hi afegeix una quarta dimensió o perspectiva sobre el model de la utilitat esperada, la "postdictiva", que entén que tot comportament humà observat és òptim, i ha d'ésser, per tant, modelat en la seva forma adequada (l'aparició de suboptimalitats és tractada com la necessitat de cercar nous factors que -suposant el comportament optimitzador permanent- expliquin les anomalies del model anterior). Aquesta forma d'enfocar la racionalitat ja ha estat tractada a l'apartat anterior, i no es discuteix aquí; tan sols s'indica la vinculació dels seus fonaments amb la Hipòtesi de la Utilitat Esperada, tot i que aquesta no porta únicament a aquesta perspectiva (P. Shoemaker, "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations", Journal of Economic Literature, vol. 20, 1982, pàg. 539).

implica el supòsit de linealitat en l'aplicació les probabilitats en l'ordenació de les preferències, és sistemàticament violat o no quan les persones realitzen eleccions convencionals. Allais va elaborar un experiment, el qual proposava dues possibilitats d'elecció, per comprovar si s'acomplia experimentalment l'axioma d'independència:

*Probabilitats de guany en elegir entre dues alternatives, A o B:*

(a): 100% guanyar 100\$	(b): 10% guanyar 500\$
	89% guanyar 100\$
	1% guanyar 0\$

*Probabilitats de guany en elegir entre dues alternatives, C o D:*

(c): 10% guanyar 500\$	(d): 11% guanyar 100\$
90% guanyar 0\$	89% guanyar 0\$

Segons el postulat d'independència de Savage, si a la primera elecció hom prefereix (a) sobre (b), en la segona hauria de preferir (d) sobre (c), ja que sinó l'aversió al risc seria variable, contrariant aquest postulat. Això es degut a que l'axioma d'independència assenyala que tercers factors no afecten una decisió donada. Això no obstant, l'explicació del canvi de preferència observat experimentalment -(c) sobre (d) en la segona elecció- és que, encara que la diferència d'esperança matemàtica és la mateixa (100-139; 11-50), la proporció canvia molt, i aquest factor porta a molts individus a reduir la seva aversió al risc en una decisió, amb la qual cosa es trenca la linealitat de la funció d'utilitat davant el risc que presuposa l'axioma d'independència<sup>135</sup>. En aquesta situació, Allais argumenta que es produeix una sobrevaloració dels resultats que s'obtenen amb certesa relativa en relació a resultats que sols són probables (efecte «certesa»). Referint-se a l'axioma d'independència, Allais senyala:

<sup>135</sup> M. Allais, "Le comportement...", art. cit., 1953, pàg. 527.

*"Si tant de personnes admettent si facilement cet axiome, c'est en fait parce qu'elles n'en aperçoivent pas toutes les implications dont certaines, loin d'être rationnelles, peuvent être au contraire dans certaines situations psychologiques (...) parfaitement irrationnelles"*<sup>136</sup>.

La crítica d'Allais a l'axioma es vincula a la seva concepció de la teoria de les decisions en risc, en el sentit que no considera aquest axioma com una condició necessària per definir la racionalitat, sinó una condició excessivament restrictiva per caracteritzar la racionalitat de les decisions en situacions de risc. Allais critica la Hipòtesi de la Utilitat Esperada perquè només funciona quan el individu no és sensible a la dispersió dels valors psicològics (sols en aquest cas es pot mesurar la utilitat cardinal amb la fórmula de Bernoulli). També critica l'adopció de la probabilitat objectiva com a fonament de la funció d'utilitat, destacant que la possibilitat d'aplicar la llei dels grans números té un àmbit limitat, i que moltes decisions es produeixen una sola vegada<sup>137</sup>. Com a alternativa, Allais pressuposava que algun índex d'utilitat existia en la ment dels individus (establint una vinculació de la teoria de la decisió amb la psicologia del risc); però la seva posició va canviar parcialment el 1977:

*"My starting point is, on the one hand, the hypothesis of the existence of objective probability defined with reference to observed frequency, independently of all considerations of random choice; and, on the other hand, the hypothesis of the existence of psychological value (or cardinal utility) defined directly with reference to the intensity of preference for certainly available sums, again independently of all considerations of random choice"*<sup>138</sup>

---

<sup>136</sup> M. Allais, "Le comportement...", art. cit., 1953, pàgs. 527-8.

<sup>137</sup> M. Allais, "Le comportement...", art. cit., 1953.

<sup>138</sup> M. Allais, "The So-Called Allais...", art. cit., 1979, pàg. 548. Aquí cal destacar que Allais separa dues qüestions que no eren prou delimitades a l'article de 1953. Per un costat rebutja la probabilitat subjectiva, degut a la impossibilitat d'ésser obtinguda experimentalment, i d'altre costat, associa en el seu

Malgrat tot, aquestes dues línies de crítica a la teoria de von Neumann i Morgenstern no portaren a Allais a la proposta de Savage. Si Savage i els principals autors de l'«escola americana» proposaren la probabilitat subjectiva i l'axioma d'independència per resoldre la dificultat de determinar les probabilitats objectives i els problemes de la dispersió de valors d'elecció<sup>139</sup>, Allais va criticar la proposta de Savage i va formular una aproximació pròpia<sup>140</sup>, prenent com a punt de partida la crítica esmentada:

*"On est arrivé à la conception qu'il fallait tenir compte non seulement de la moyenne pondérée suivant leurs probabilités des valeurs psychologiques (...), mais également de l'ensemble de la distribution des probabilités..."*<sup>141</sup>

Així, Allais proposava una teoria general que inclou la formulació de von Neumann-Morgenstern com un cas particular<sup>142</sup>, en el qual la "utilitat esperada" pot ésser inferida directament de l'elecció entre loteries, gràcies a la linialitat que assegura l'axioma d'independència. En la proposta d'Allais, s'incorpora un índex cardinal

---

índex d'utilitat psicològica les distorsions que originàriament suposava implicades en la probabilitat subjectiva, resultant aquest índex d'elaboració més complex. Com diu M. Machina, *"Many of the differences, and most of the misunderstanding, between Allais and his critics stem from this fundamental difference over whether preference rankings or psychological values are the underlying 'real' generators of choice"* (a "Rational' Decision Making versus 'Rational' Decision Modelling?", Journal of Mathematical Psychology, vol. 24, 1981, pàg. 168).

<sup>139</sup> Com destaca A. Tversky, Savage reflecteix una interpretació monetària de les conseqüències quan defensa l'axioma d'independència (que Allais no comparteix), i en aquest sentit, *"Savage's analysis is therapeutic rather than prescriptive: It suggests a state of mind which leads to a particular resolution of the problem, but it contains no reason why that particular state of mind should be adopted"* (cf. A. Tversky, "A Critique of Expected Utility Theory: Descriptive and Normative Considerations", Erkenntnis, vol. 9, 1975, pàg. 172).

<sup>140</sup> Per a una exposició de la pròpia aproximació, i una revisió posterior, M. Allais, "The So-Called Allais...", art. cit., 1979.

<sup>141</sup> M. Allais, "Le comportement..." art. cit., 1953, pàg. 513.

<sup>142</sup> Allais destaca que *"it is possible to prove that in the case of repetitive games for which the probability of ruin is low, the value of  $r$  is relatively negligible and that as a result, the von Neumann-Morgenstern neo-Bernoullian formulation becomes approximately valid"* (cf. M. Allais, "The foundations..." cit., 1984, pàg. 37).

$\{s_i\}$ , que recull la distribució de probabilitat (variància, moda,..) del conjunt de valors psicològics dels resultats<sup>143</sup>. No obstant, la proposta d'Allais, com assenyala M. Machina, era relativament ambigua:

*"It is true that Allais offers many 'psychological' factors of choices involving risk without always suggesting how strong, and in which direction, he expects them to operate"*<sup>144</sup>.

Tota aquesta discussió és important per dos motius. El primer és que les crítiques, així com les propostes inicials de M. Allais, es van produir just en el moment quan s'estava desenvolupant l'axiomàtica de la Hipòtesi de la Utilitat Esperada, i de forma associada, la normativització del concepte de racionalitat que incorporava. A partir dels propis experiments d'Allais<sup>145</sup>, es va desenvolupar una àmplia evidència experimental de que les desviacions de la linealitat en l'aversion al risc són sistemàtiques i previsibles<sup>146</sup>. D'aquesta forma, les crítiques que sorgiren amb posterioritat a aquest i a altres axiomes de la Hipòtesi de la Utilitat Esperada tenien ja un punt de partida. Això ha facilitat que, sobretot en els anys vuitanta, es desenvolupin diverses formulacions alternatives, majoritàriament generalitzacions de

---

<sup>143</sup> Com destaca M. Machina, per M. Allais "the  $\{s_i\}$  index exist independently of and logically prior to risk preferences, and since it is the  $s_i$ 's rather than the actual monetary values  $\$i$  which measure the true psychological benefit of the outcomes, a risk averter would naturally choose to take into account the dispersion as well as the mean of the  $s_i$ 's in ranking lotteries", (Cf. M. Machina, "Rational..", art. cit., 1981, pàg. 168). O. Hagen intentà construir formalment aquest índex d'utilitat, incorporant diferents moments estadístics a la seva formulació, mantenint-se dins un model de von Neumann-Morgenstern sense l'axioma d'independència (Cf. O. Hagen, "Towards a Positive Theory of Preferences under Risk", a M. Allais i O. Hagen, (Eds.) *Expected Utility...*, op. cit., 1979).

<sup>144</sup> M. Machina, "Rational'..", art. cit., 1981, pàg. 170.

<sup>145</sup> M. Allais, "Selected Findings of the 1952 Experiment", M. Allais i O. Hagen (eds.), *Expected Utility...*, op. cit., 1979.

<sup>146</sup> Per exemple, O. Hagen, "Towards a Positive Theory...", art. cit., 1979; C. Hong i S. Waller, "Empirical Tests of Weighted Utility Theory", *Journal of Mathematical Psychology*, vol. 29, 1986.



la teoria de la utilitat esperada<sup>147</sup>. El segon motiu és que les posicions d'Allais es centren en els aspectes que s'han qüestionat anteriorment sobre la racionalitat econòmica fonamentada en l'aplicació consistent d'un principi d'elecció sobre els ordres de preferències (vid. apt. 1.2.2.). Així, la necessitat d'un índex o un principi psicològic que mostri com els individus valoren les probabilitats (un indicador de la seva satisfacció individual) apareix com un requisit per fonamentar en darrer terme les preferències.

Davant aquests arguments, els defensors de la Hipòtesi presenten la següent qüestió: on està la fallida, en els axiomes de la teoria, o en la capacitat dels subjectes per analitzar els complexos problemes d'una elecció?. La seva resposta és que la presentació de les dades («framing») distorsiona de forma determinant moltes decisions, però la teoria és essencialment correcta<sup>148</sup>. Aquestes dues posicions tenen aspectes irreconciliables, malgrat compartir força elements conceptuals. En aquest treball interessa especialment analitzar les dificultats de la primera posició, que defensa la necessitat d'establir un nexa entre el càlcul racional i la satisfacció individual. No hi ha cap necessitat convincent de mantenir un concepte ideal de racionalitat per desenvolupar seguidament una fonamentació adequada d'una teoria descriptiva del comportament racional individual<sup>149</sup>, que tingui capacitat de predir gran part de les eleccions que s'observen.

---

<sup>147</sup> Per a una revisió general, veure P. Shoemaker, "The Expected Utility Model...", art. cit., 1982; i també M. Machina, "Choice Under Uncertainty: Problems Solved and Unsolved", *Economic Perspectives*, vol. 1, 1987.

<sup>148</sup> Aquesta darrera és la tesi de H. Bierman (cf. "The Allais Paradox: A Framing Perspective", *Behavioral Science*, vol. 34, 1989), quan destaca que l'error no està en la teoria de la utilitat sinó en la manca d'habilitat dels subjectes per analitzar els "complexos" problemes de probabilitat sense assistència. En el seu experiment, quan en la prova de decisió s'acompanyava una nota explicativa sobre l'estructura lògica de l'elecció, els resultats eren majoritàriament consistents.

<sup>149</sup> Com diu P. Anand, "To some extent the axioms have been advanced as canons of rationality by brute assertion rather than cogent reason (...) surely rationality has something to do with the deliberative process and is not just defined by certain preference patters.." (cf. "Are the Preference Axioms Really Rational?", a *Theory and Decision*, vol. 23, 1987, pàgs. 210-11).

### 1.3.3. Noves discussions de la Hipòtesi de la Utilitat Esperada: els intents de relaxació axiomàtica.

El significat de la paraula 'racionalitat' en l'elecció sota risc ha estat el debat més llarg, confús i controvertit de tota la teoria moderna de la decisió. Com ja s'ha assenyalat en l'apartat anterior, la discussió sobre els axiomes d'aquesta teoria té els seus orígens arrel de la pròpia formulació de la Hipòtesi, a principis dels anys cinquantes. Després d'uns vint anys llargs d'acceptació i desenvolupament, aquesta discussió torna a estar en el centre dels debats sobre la fonamentació de la teoria econòmica<sup>150</sup>, destacant-se que la violació de la consistència en les eleccions sota risc pot ser deguda a més factors que en la teoria de la utilitat amb informació completa (vid. apt. 1.2.2.). Seguidament es presenta una anàlisi dels problemes dels tres axiomes bàsics (ordre-completesa, independència i continuïtat) que assumeix la forma canònica de la hipòtesi de la Utilitat Esperada<sup>151</sup>.

#### 1. Ordre-Completesa

(1. ordre):  $>$  sobre P és un ordre dèbil.

Cal afegir la completesa a aquesta expressió per mostrar com es consideren les relacions de preferència en la Hipòtesi de la Utilitat Esperada. La completesa és un axioma fonamental (més que la independència) de la Hipòtesi de la Utilitat Esperada, mantenint-se com a base en la major part de les propostes de modificacions de la Hipòtesi<sup>152</sup>. Donat un conjunt amb dos elements, 'a' i 'b', tenim que:

---

<sup>150</sup> Vegeu per exemple, K. Arrow, "Risk Perception in Psychology and Economics", *Economic Inquiry*, vol. 20, 1982; i J. Hirshleifer i J. Riley, "The Analytics of Uncertainty and Information -An Expository Survey", *Journal of Economic Literature*, vol. 17, 1979.

<sup>151</sup> Per un tractament formal dels axiomes, vid. l'annex 1.3.A.

<sup>152</sup> M. Machina, "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty", *Journal of Economic Literature*, vol. 27, 1989. Una excepció és l'article de R. Aumann, "Utility Theory without the Completeness Axiom", *Econometrica*, vol. 32, 1962 (vid. apt. 2.2.4.).

$a > b$ , o  $b > a$ , o  $a \equiv b$ , per tot  $a, b \in u$ .

Produïda una elecció entre dos elements, o bé 'a' és preferit a 'b', o 'b' és preferit a 'a', o bé 'a' i 'b' són indiferents. Quan hi ha més de dos elements, a l'axioma d'ordre de les alternatives s'hi afegeix generalment la condició de transitivitat. La violació d'aquest axioma es produeix quan, per exemple, simultàniament 'a' es preferir a 'b', i 'b' és preferit a 'a', sigui perquè són alternatives incomparables, sigui perquè reflecteix un comportament inconsistent. Una de les formes més freqüents en que es produeix aquesta situació és la que s'anomena «framing», o efectes de representació<sup>153</sup>.

Els efectes de «framing» en les eleccions són desconcertants per aquest axioma. El fenomen s'observa quan mitjans alternatius de representar eleccions probabilísticament equivalents porten a diferències sistemàtiques en l'elecció<sup>154</sup>. Així, variacions en l'estructura de l'elecció produeixen violacions sistemàtiques de l'estabilitat de les eleccions, amb el qual es qüestiona la condició de la invariància (diferents representacions d'una mateixa elecció han de portar a la mateixa preferència); i porten a què per situacions d'elecció probabilísticament equivalents, sigui alhora 'a' > 'b' i 'b' > 'a'. La literatura sobre marketing és plena d'exemples sobre el «framing» a la vida real. Un cas es la venda d'habitatges cars amb un nombre de terminis més que proporcional al dels habitatges barats. Aquests efectes són utilitzats molt sovint intuïtivament per finalitats comercials. La solució per evitar aquestes violacions seria imposar una representació canònica dels dilemes

---

<sup>153</sup> L'efecte es pot produir tot i existir informació completa. D'altre costat, el sentiment d'ambivalència de preferir alhora 'a' > 'b' i 'b' > 'a' és freqüent que es produeixi quan manca informació sobre l'espai de decisió o sobre l'objecte de l'elecció. Una altra qüestió seria quan el nombre d'elements inclosos en una elecció fos molt ampli i difícilment l'individu tingués capacitat cognitiva suficient per ordenar comparativament totes les combinacions del conjunt d'elecció, restant llavors preferències indeterminades. (Cf. P. Anand, "Are the Preference...", art. cit., 1987).

<sup>154</sup> Els treballs sobre els efectes del «framing» han estat elaborats especialment per A. Tversky i D. Kahneman, "Rational Choice and the Framing of Decisions", a R. Hogarth i M. Reder (Eds.), Rational Choice, University of Chicago Press, Chicago, 1987.

d'elecció; però per descomptat, no és així com la gent decideix, i la invariància falla contínuament<sup>155</sup>. La pregunta central per analitzar els efectes de «framing» és la qüestió de la seva estabilitat, per a poder introduir -en cas d'una certa estabilitat que permeti fer prediccions- els elements que el determinen en la funció d'utilitat de l'agent.

## 2. Independència

(2. independència):  $p > q \Rightarrow \delta p + (1-\delta)r > \delta q + (1-\delta)r$

És l'axioma que ha atret més l'atenció en les discussions sobre la Hipòtesi de la Utilitat Esperada. Això es degut a què aporta a la teoria el seu contingut empíric imposant una restricció sobre la forma de la funció de preferència, en definir que els subjectes entenen les probabilitats com a lineals<sup>156</sup>. En primer lloc, els experiments mostren que difícilment una funció d'utilitat lineal pot integrar les diverses formes de comportament front el risc. En la teoria de la utilitat esperada, l'aversion al risc és equivalent a la convexitat de la funció d'utilitat (i la concavitat és igual a l'atracció pel risc). Un dels problemes més importants és que experimentalment s'ha observat que de forma sistemàtica un mateix individu, pot mostrar un comportament advers al risc en una elecció, i un comportament favorable

---

<sup>155</sup> A. Tversky i D. Kahneman, "Rational Choice and the Framing of Decisions", a D. Bell et al., Decision-Making..op.cit., 1988 (versió revisada de l'article citat a la nota anterior). Tversky i Kahneman afegeixen que "*unfortunately, the mental machinery that transforms precepts and sentences into standard forms does not automatically apply to the process of choice*" (pàg. 172).

<sup>156</sup> E. McClennen, "Sure-Thing Doubts", a B. Stigum i F. Wenstoeep (Eds.), Foundations of Utility and Risk Theory with Applications, Reidel, Dordrecht, 1983; per McClennen l'única defensa poderosa de l'axioma d'independència com a principi normatiu és tractar aquest com un principi de domini, basant-se en una intuïció segura sobre el contingut de la racionalitat. Entenent l'elecció racional com a maximització de les preferències de l'agent en funció de les conseqüències; tot i així, la conclusió de McClennen és ambigua sobre les possibilitats de defensa de l'axioma en la perspectiva normativa: "*If the sure-thing or independence axiom has yet to be linked in any decisive fashion to behavior which can be taken as paradigmatically rational then we would be well advised not to judge irrational those patterns which fail to conform this principle*" (pàg. 130).

al risc en una altra elecció<sup>157</sup>.

Dos efectes observats que violen l'axioma d'independència són l'«efecte isolació» i l'«efecte d'assegurança probabilística». L'«efecte isolació» mostra que per simplificar les eleccions, la gent tendeix a descompondre els problemes; com que la descomposició es pot fer de diverses formes (per exemple, seqüencial o comparativa), això pot portar a problemes d'inconsistència en les preferències (perquè en una mateixa elecció, segons la descomposició efectuada, s'escullin diferents alternatives). L'«efecte de l'assegurança probabilística», com una altra forma de no linealitat (no es manté la convexitat de la funció d'utilitat -aversió al risc-), apareix quan en dues situacions diferents (la primera amb menys risc que la segona) els individus tenen el mateix cost per una assegurança contingent (amb la qual es redueix el risc a zero a la primera situació) que per una assegurança probabilística (es redueix una part del risc a la segona situació); així, en la primera situació escullen fer l'assegurança, mentre que a la segona no ho han. Tot i que el cost sigui igual, i que el nivell de risc que es pot reduir amb l'assegurança és el mateix (o bé que la reducció de la pèrdua no protegida sigui en proporció la mateixa), la primera decisió s'adopta perquè és més preferida una absència total de risc que una absència parcial<sup>158</sup>. En altres paraules, és freqüent pagar per reduir el risc del 25% a zero, però és estrany fer-ho per reduir el risc del 75% al 50%, tot i que el cost sigui el mateix.

### 3. Continuïtat

(3. Continuïtat):  $p > q > r \Rightarrow \sigma p + (1-\sigma)r > q > \beta p + (1-\beta)r$ , per algun  $\sigma$  i  $\beta$

---

<sup>157</sup> D. Kahneman i A. Tversky desenvolupen la idea de M. Allais sobre l'efecte certesa, i destaquen que *"In the positive domain, the certainty effect contributes to a risk averse preference for a sure gain over a larger gain that is merely probable. In the negative domain, the same effect leads to a risk seeking preference for a loss that is merely probable over a smaller loss that is certain. The same psychological principle -the overweighting of certainty- favors risk aversion in the domain of gains and risk seeking in the domain of losses."* (a "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, vol. 47, 1979, pàgs. 268-9).

<sup>158</sup> D. Kahneman i A. Tversky, "Prospect...", art. cit., 1979.

només entre 0 i 1.

Experimentalment s'han observat freqüents violacions no estocàstiques de l'axioma de la continuïtat. El fenomen implica també indirectament la violació de la condició de transitivitat de les preferències, almenys per funcions lineals<sup>159</sup>. L'observació del fenomen de reversió de la continuïtat en les preferències és força estesa entre els psicòlegs a través de recerques experimentals<sup>160</sup>. Aquest apareix quan un individu valora indiferentment una loteria A i una determinada suma fixa 'a', prefereix la loteria A a la loteria B, i és indiferent entre la loteria B i una suma 'b' superior a 'a'; però en el moment de l'elecció prefereix 'b' a una menor suma 'a'. La interpretació dels psicòlegs és que no existeix un mecanisme comú que generi a la vegada el comportament d'elecció i el de la valoració, sinó que són processos distints, sotmesos a diverses influències<sup>161</sup>. En tot cas però, el que cal distingir és específicament si es tracta de fallides en la consistència, que afecten als processos d'elecció; o bé si cal analitzar com és que les preferències no segueixen l'axioma de la continuïtat.

En aquest sentit, per defensar una possible teoria de la racionalitat sense la continuïtat, hom està forçat a mostrar que l'assignació d'un simple nombre univalorescala és inapropiat. Argumentar que una persona amb preferències discontinües és irracional implica acceptar abans la condició de què només hi ha un ordre de preferències -apart de la consistència en l'elecció-. En el cas de no acceptar aquesta condició, es planteja el problema del conjunt d'elecció; que pot variar en cada

---

<sup>159</sup> A. Tversky, "Intransitivity of Preferences", Psychological Review, vol. 76, 1969.

<sup>160</sup> Com exemple d'una revisió d'aquestes recerques per part dels iniciadors d'aquesta discussió, vegeu P. Slovic, B. Fischhoff i S. Lichtenstein, "Preference Reversals: A Broader Perspective", American Economic Review, vol. 73, 1983.

<sup>161</sup> M. Machina, "Choice under..", art. cit., 1987. En un sentit similar destaquen A. Tversky, P. Slovic i D. Kahneman la dificultat de construir "A theory of choice that is both normatively acceptable and descriptively adequate" (a "The Causes of Preference Reversal", American Economic Review, vol. 80, 1990, pàg. 215).

elecció, sigui introduint elements nous, sigui incorporant diversos atributs als objectes d'elecció (cosa que porta a problemes de decisions múltiples)<sup>162</sup>.

Com a producte de les refutacions experimentals dels axiomes i dels diversos problemes de fonamentació teòrica -alguns dels quals ja han estat esmentats- s'han produït en els darrers anys multitud de noves propostes per «reformar» la Hipòtesi de la Utilitat Esperada. Aquestes propostes tenen en comú la relaxació d'un o altre dels axiomes que conformen la Hipòtesi, i intenten ésser ambivalents respecte al seu caire descriptiu o normatiu, encara tendeixen a prendre una determinada orientació entre els dos extrems. En tot cas, són més sensibles a la justificació de la seva plausibilitat, sobretot quan tenen com a punt de partida la voluntat d'introduir capacitat interpretativa sobre el comportament observat. Com destaca M. Machina, generalment els nous models intenten assolir tres objectius<sup>163</sup>: i) ajustar millor les dades que el model de la Utilitat Esperada, ii) intentar mantenir en tot el possible l'elegant simplicitat de la Hipòtesi de la Utilitat Esperada, iii) evitar la inconsistència dinàmica que poden produir les generalitzacions del model, quan abandonen la linealitat de les funcions d'utilitat. Entre les diverses teories que proposen una alternativa a la Hipòtesi de la Utilitat Esperada per fonamentar l'anàlisi de les decisions sota risc, cal destacar per la diversitat dels seus pressupòsits, les propostes de la «Prospect Theory», la «Generalized Utility Theory», i la «Regret Theory»<sup>164</sup>.

---

<sup>162</sup> P. Anand, "Are the Preference...", art. cit., 1987. Una possibilitat interessant pels problemes de la consistència derivats de la discontinuïtat seria l'aplicació de funcions de preferències basades en la teoria dels conjunts borrosos, que consideren amb un certa probabilitat la inclusió o no d'una alternativa en un conjunt donat, sobre el tema, vid. J. Kacprzyk i M. Roubens (Eds.), Non-Conventional Preference Relations in Decision-Making, Springer Verlag, Berlin-Heidelberg, 1988.

<sup>163</sup> M. Machina, "Dynamic Consistency...", art. cit., 1989. En termes més generals -en incloure propostes que rebutgen la Hipòtesi de l'Utilitat Esperada- Elster defensa que els grans objectius són: "to explain the observed violations of expected utility theory (...) by providing an account that (a) is simple and intuitively plausible, (b) explains all observed deviations from expected utility theory and (c) predicts no unobserved deviations" (cf. Solomonic... op. cit., 1989, pàg. 31).

<sup>164</sup> Per a una visió general de les diverses propostes existents, vegeu M. Machina, "Choice under...", art. cit., 1987; P. Shoemaker, "The Expected...", art. cit., 1982 i P. Fishburn, "Expected Utility...", art. cit., 1988.

La «Generalized Utility Theory», proposada per M. Machina<sup>165</sup>, i la «Prospect Theory», proposada per Kahneman i Tversky<sup>166</sup> treuen l'axioma d'independència del model de la Utilitat Esperada i proposen diferents funcions no lineals de les preferències individuals, que intenten ésser una resposta a l'observació de la no linealitat en les probabilitats, seguint una proposta original de M. Allais. El primer proposa un índex quadràtic general sobre les probabilitats, desenvolupant una successió de Taylor fins el segon membre<sup>167</sup>, com un principi d'adopció de probabilitats més complexe, i més plausible com a hipòtesi del comportament individual. Els segons proposen la construcció d'un índex psicològic (ponderacions de decisió) no entès com una derivació de la teoria de les probabilitats, i no dependent per tant dels seus axiomes, però sí dependent de referents empírics per a la seva aplicació<sup>168</sup>. Tversky i Kahneman defensen la seva proposta assenyalant que solament la «Prospect Theory», no necessita els axiomes de domini (ordre) i

---

<sup>165</sup> M. Machina, "Expected Utility' Analysis Without the Independence Axiom", *Econometrica*, vol. 50, 1982; M. Machina, "Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom", a B. Stigum i F. Wenstoeop (Eds.), *Foundations of Utility...*, op. cit., 1983.

<sup>166</sup> D. Kahneman i A. Tversky, "Prospect Theory...", art. cit., 1979.

<sup>167</sup> La crítica a la incoherència seqüencial de les eleccions és la més freqüent en els models que relaxen l'axioma d'independència (per exemple T. Seidenfeld, "Decision Theory without 'Independence' or without 'Ordering'", *Economics and Philosophy*, vol. 4, 1988). En un dels seus darrers articles, M. Machina intenta respondre a aquestes crítiques senyalant que els supòsits de crítica són inapropiats per ésser aplicats als "maximitzadors" dels models sense axioma d'independència, especialment el supòsit de que els agents són conseqüencialistes. En el seu model, Machina assumeix que els agents no han d'ésser conseqüencialistes: "*The resolution of uncertainty that takes such an agent to a particular decision node does not lead to a new opportunity set of probability distributions and associated preference functions values, but rather to that subset of the original opportunity set (...) that corresponds to the choices still available, with bars used to denote those risks that have already been borne*" ("Dynamic Consistency...", art. cit., 1989, pàg. 1469). Machina afirma que els agents no han d'acceptar que els elements del conjunt d'oportunitat original que eren més preferits es mantinguin sempre en els nous subconjunts de decisions seqüencials, sense possibilitat de reconsiderar les seves preferències a la meitat d'un procés seqüencial d'elecció. Per a tal fi, Machina defensa l'existència de preferències no-separables (intertemporalment o sobre qualsevol altra dimensió) destacant que les decisions anteriors afecten a les valoracions de les noves eleccions al llarg d'un procés de decisió. Així, en les decisions amb preferències no-separables es necessiten més requeriments d'informació (p.e. estat emocional) que amb preferències separables, i l'únic que es pot fer és restringir els efectes d'interacció a aquells que es considerin més importants.

<sup>168</sup> Es suposa que  $\pi$  (índex de decisió) es una funció monotònica de  $p$  (probabilitats), amb una escala normalitzada entre [0,1]. Llavors, les baixes probabilitats estan sobrevalorades, i les probabilitats moderades i altes són subvalorades, essent aquest darrer efecte més pronunciat que el primer. A. Tversky i D. Kahneman, "Rational Choice..", art. cit., 1988.



d'independència, ja que no es proposen explicar com es realitzen les eleccions racionals, sinó tan sols les eleccions en situacions de risc, sense concisions de racionalitat:

*"Because invariance and dominance are normatively essential and descriptively invalid, a theory of rational decision cannot provide an adequate description of choice behavior [...] Prospect theory (...) being unabashedly descriptive and in making no-normative claims. It is designed to explaining preferences, whether or not they can be rationalized"*<sup>169</sup>

La «Regret Theory», proposada paral·lelament per D. Bell i per G. Loomes i R. Sugden<sup>170</sup> estableix un model de decisió sota risc sense l'axioma de consistència, cosa que implica el no manteniment -o debilitació- de la transitivitat, però sí el manteniment de l'axioma d'independència. La proposta de treure l'axioma de transitivitat ja havia estat formulada anteriorment<sup>171</sup>, però l'especificat de la «Regret Theory» consisteix en canviar la funció d'utilitat de von Neumann-Morgenstern  $V(x)$  per una funció d'atracció/rebuig  $R(x,y)$  que representa el nivell de satisfacció o insatisfacció que experimenta un individu al rebre el resultat 'x' quan el resultat alternatiu era 'y'. La funció és ponderada per dos conjunts de probabilitats, i s'escull un d'ells segons quina sigui l'expectativa més elevada de la seva utilitat<sup>172</sup>, de forma que l'elecció sota risc es converteix en una elecció interactiva:

---

<sup>169</sup> A. Tversky i D. Kahneman, "Rational Choice..", cit., 1988, pàgs. 165 i 185.

<sup>170</sup> D. Bell, "Regret in Decision-Making under Uncertainty", Operations Research, vol. 20, 1982; G. Loomes i R. Sugden, "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty", Economic Journal, vol. 92, 1982.

<sup>171</sup> Per exemple I. Levy, "On Indeterminate Probabilities", Journal of Philosophy, vol. 71, 1974.

<sup>172</sup> M. Machina, "Choice under..", art. cit., 1987; R. Sugden, "Why be Consistent..", art. cit., 1985. La consistència es pot eliminar en l'aplicació d'un principi d'elecció perquè es substitueix per un mecanisme més complex d'adopció de decisions, que tampoc requereix la transitivitat.

*"In regret theory (...) actions are not evaluated independently of one another (...) the notion of value that determines choice is that of modified utility, and the modified utility of an action depends not only on the nature of the action itself, but also on the nature of the other action in the feasible set"<sup>173</sup>*

El problema dels models de la «regret theory» és que el seu correcte funcionament depèn de què es compleixi en la funció d'utilitat l'axioma d'independència<sup>174</sup>, apart de l'axioma de completesa, axiomes que -com ja s'ha assenyalat- han estat molt criticats experimentalment.

Els resultats exposats en aquesta secció són una mostra dels problemes existents per fonamentar una teoria de la decisió que pugui ser aplicada com a hipòtesi de racionalitat per a explicar el comportament observat. Com a síntesi de l'anàlisi efectuada, s'ofereix un conjunt de consideracions per a tenir presents en els propers capítols. Aquestes consideracions es refereixen successivament als diferents nivells en que s'ha efectuat l'anàlisi -heurístic, metodològic i teòric- dels models interpretatius sobre els processos de decisió racional en situacions de risc:

*i) Tot i acceptar la necessitat de formular models de decisió basats en la hipòtesi de racionalitat, cal insistir que resten problemes de plausibilitat interpretativa en els models que mantenen una estructura axiomàtica-deductiva, especialment quan pretenen una aplicabilitat general.*

*ii) La construcció d'un índex de satisfacció per mesurar la utilitat no ha de dependre d'observacions empíriques, si hom desitja disposar d'una teoria de la decisió racional que parteixi de principis heurístics sobre el comportament humà. Ara bé, per aplicar a moltes situacions, aquest índex ha de ser força més complex que una*

---

<sup>173</sup> R. Sugden, "Why be Consistent..", art. cit., 1985, pàg. 175.

<sup>174</sup> Ch. Holt, "Preference Reversals and the Independence Axiom", American Economic Review, vol. 76, 1986.

*simple funció línia.*

*iii) Un resultat significatiu que es desprèn del conjunt de propostes examinades és que la hipòtesi de racionalitat no implica un conjunt axiòmatic únic, sinó que pot ser entesa com un concepte que engloba un espectre de processos de decisió amb força variacions internes, segons els axiomes que es mantinguin i com es combinin. L'èxit de cada formulació concreta pot dependre en molt bona part de factors externs a la interpretació, com la teoria fàctica amb què es vinculi.*

*iv) Apareixen situacions d'incompatibilitat en els models de decisió racional quan es comencen a reduir o modificar alguns axiomes. Per exemple, tot i existir alguns problemes secundaris, és possible eliminar l'axioma d'independència, o bé el de consistència, però no els dos al mateix temps. És necessari salvaguardar sempre, si hom vol mantenir una mínima i dèbil condició de racionalitat, l'axioma d'ordre (extensionalitat); tot i què també s'ha posat en dubte el seu compliment en moltes situacions experimentals d'elecció.*

L'estat en que es troba la part més nuclear de l'«elecció racional» (els axiomes sobre les preferències en situacions de risc) porta a suposar que intentar fonamentar exclusivament els models explicatius de la teoria econòmica a partir de la formulació clàssica de la Hipòtesi de la Utilitat Esperada es una dificultat creixent. D'altra banda, cal senyalar que trobant-se la teoria de jocs formalment vinculada amb la Hipòtesi pel seu origen comú en l'obra de von Neumann-Morgenstern - encara que sense ésser-ne dependent-, apareixen els mateixos problemes quan hom discuteix la interacció estratègica d'actors racionals. Una altra possibilitat seria desconnectar la racionalitat de la teoria de jocs: en aquest cas els problemes són distints; però aquesta ja és una qüestió a tractar en el següent capítol. Des de les ambigües indicacions de M. Allais sobre la necessitat de construir un índex d'utilitat força més complex, fins a les propostes de Tversky i Kahneman sobre un índex psicològic de caire observacional, passant pels suggeriments de Machina sobre la

definició de conseqüències per aplicar a les preferències no separables; allò que es planteja, encara de forma temptativa, és la necessitat d'incorporar més interpretació teòrica, no sempre necessàriament de forma axiomàtica, sobre els models de decisió racional. En aquest nous models, les dificultats de construcció empírica -reduïdes- han de ser compensades per la major plausibilitat de proposicions fàctiques sobre els contextos en els que es formulen les decisions, integrant teòricament les dues vies en l'anàlisi dels mecanismes de decisió. En aquesta situació, es podrien incorporar valors, normes, junt a les propietats de la racionalitat:

*"..Utility theory provides only a partial analysis of the problem of rational choice. A comprehensive analyses of rational choice and risk should face the interpretation problem as well as the problem of the legitimacy of values. This analysis is likely to be explicative, or even therapeutic, rather than normative in value"<sup>175</sup>*

Tanmateix, per més que es facin força complexes, els models d'home sempre seran grans simplificacions rudimentàries de la gran complexitat i diversitat existent entre els individus humans. A més, aquests sempre hauran de dependre d'alguna referència psicològica, encara que aquesta sigui força llunyana. Però si hom vol derivar propietats socials emergents a partir de l'acció individual, ha d'acceptar la tensió d'aquestes limitacions.

---

<sup>175</sup> A. Tversky, "A Critique of Expected Utility Theory...", art. cit., 1975, pàg. 173.

### 1.3.A. Annex: La formalització de la Hipòtesi de la Utilitat Esperada.

Entre les diverses formalitzacions alternatives sobre la Hipòtesi de la Utilitat Esperada de J. von Neumann i O. Morgenstern, realitzades per diversos autors, s'adopta la realitzada per N. Jensen<sup>176</sup>, amb algunes petites alteracions assenyalades en propostes d'altres autors.

Sigui  $X$  un conjunt no buit, i sigui  $P$  un conjunt convex de distribucions de probabilitat sobre  $X$ , llavors  $\delta p + (1-\delta)q$  està a  $P$  quan  $p, q \in P$  i  $0 \leq \delta \leq 1$ . Es pot entendre  $\delta p + (1-\delta)q$  com una combinació de probabilitat de  $p$  i  $q$  amb les respectives ponderacions  $\delta$  i  $(1-\delta)$ <sup>177</sup>. Els axiomes que defineix Jensen són els següents, per a tot  $p, q, r \in P$  i per tot  $0 < \delta < 1$ :

1. (ordre):  $>$  sobre  $P$  és un ordre dèbil.
2. (independència):  $p > q \Rightarrow \delta p + (1-\delta)r > \delta q + (1-\delta)r$ <sup>178</sup>

En aquesta situació s'imposa un axioma de continuïtat: existeix algun  $\delta \in [0, 1]$  tal que,  $\{\delta p + (1-\delta)q\}$  i  $r$ ; amb el qual s'arriba al teorema  $0 < \delta < 1$ ; cosa que permet construir una funció d'utilitat, que es basa en una funció contínua de probabilitats associada a una funció ordinal de la jerarquia de preferències<sup>179</sup>.

---

<sup>176</sup> N. Jensen, "An Introduction to Bernoullian Utility Theory. I. Utility Functions" i "An introduction to Bernoullian Utility Theory. II. Interpretation, Evaluation and Application: a Critical Survey", *Swedish Journal of Economics*, vol. 69, 1967.

<sup>177</sup> Una definició alternativa, presentada per G. Owen (a *Game Theory*, op. cit., 1982, pàg. 117) es: "Siguin  $p, q$  dos successos, i  $0 \leq \delta \leq 1$ ; llavors per  $\delta p + (1-\delta)q$  entendrem una loteria que té dos resultats possibles, amb les probabilitats  $\delta$  i  $(1-\delta)$  respectivament".

<sup>178</sup> La defensa d'aquest axioma és sintetitzada per M. Machina: "The argument for the 'rationality' of this prescription is straightforward: the choice among the latter pair of prospects is equivalent in terms of final probabilities to being presented with a coin which has a  $(1-\delta)$  chance of landing tails (in which case you will win the lottery ( $r$ )) and being asked before the flip whether you would prefer to win the lottery ( $p$ ) or ( $q$ ) in the even of a head. Now, either the coin will land tails, in which case your choice won't have mattered, or else it will land heads, in which case you are in effect back to a decision between ( $p$ ) and ( $q$ ) and you should clearly make the same choice as you did before" (canvi nomenclatura pròpia) (Cf. M. Machina, "Rational' Decision Making...", art. cit., 1981, pàg. 166-7).

<sup>179</sup> Quan l'ordre de preferències és un continu, es suposa una funció de probabilitat contínua entre 0 i 1, és possible convertir-la en una funció numèrica d'utilitat (suposant que les combinacions preferides continguin la seva frontera), però com a propietat, no ha d'existir necessàriament la continuïtat en totes les eleccions.

3. (Continuïtat):  $p > q > r \Rightarrow \sigma p + (1-\sigma)r > q > \beta p + (1-\beta)r$ , per algun  $\sigma$  i  $\beta$  només entre 0 i 1.

Els axiomes afirmen que les preferències estan ordenades correctament, que combinacions similars convexes amb una tercera distribució (r a 2.) mantenen les preferències, i que una distribució entre altres dues en preferència (q a 3.) està també entre combinacions convexes no trivials de les dues primeres.

Llavors, podem presentar el nucli central de la teoria de von Neumann i Morgenstern, el seu teorema sobre la utilitat, que diu que donada  $>$  sobre P convexa, els axiomes 1., 2. i 3., prenen, en el cas i només en el cas que hi hagi una funció real 'u' sobre P tal que, per tot  $p, q \in P$  i tot  $0 < \delta < 1$ ,

$$p > q \Leftrightarrow u(p) > u(q)^{180}$$

$$u(\delta p + (1-\delta)q) = \delta u(p) + (1-\delta)u(q) \text{ [linealitat]}$$

Com que ens movem amb valors entre 0 i 1, això és igual a:

$$\delta \cdot 1 + (1-\delta) \cdot 0 = \delta$$

significa que l'índex d'utilitat mesurable és l'esperança matemàtica de la utilitat dels

<sup>180</sup> S'assigna una utilitat  $\delta_i$  (algun valor entre 0 i 1) al pagament monetari  $x_i$  (o  $q$ ) tal que l'individu és indiferent entre rebre amb seguretat aquest pagament o participar en el joc  $x_i$  (o  $p$ ) (valor 0) amb probabilitat  $\delta$  i  $x_n$  (o  $r$ ) (valor 1) amb probabilitat  $(1-\delta)$ . Fent el mateix per tot els valors de  $x_1$ , de  $x_2$  fins  $x_{n-1}$ , obtenim les "utilitats mesurables"  $(\delta_2, \dots, \delta_{n-1})$  que corresponen als diferents pagaments  $x_2$  fins  $x_{n-1}$ . De forma general, això pot expressar-se com:

$$U(\text{joc}) = \delta U(x_n) + (1-\delta)U(x_1)$$

Aquí es pot destacar la diferència entre la proposta de von Neumann-Morgenstern i la de Savage. Mentre que la utilitat  $\delta_i$  és equivalent a la probabilitat de freqüència (objectiva) per von Neumann i Morgenstern, per Savage es refereix a la probabilitat subjectiva (i per tant, Savage necessita incloure a l'índex d'utilitat mesurable, a més de l'esperança matemàtica, la seva variància). En tot cas, en la definició no es considera el procés mitjançant el qual es formen les probabilitats subjectives. Per von Neumann i Morgenstern, la subjectivitat de l'individu només està present en la seva relació d'indiferència entre situacions de risc (i no hauria de ser necessàriament sempre constant -vid. punt 1.3.3.). En aquest punt, l'axioma d'independència garanteix una actitud constant davant el risc en cada decisió (i això ho necessita més clarament Savage que von Neumann i Morgenstern, ja que la variació en el seu model prové de la variació -subjectiva- en la fixació de les probabilitats).

pagaments del joc.

Aquestes condicions sobre 'u' impliquen que 'u' és única fins a una transformació lineal positiva, amb el qual tenim per cada distribució finita de X (definint u sobre X per  $u(x) = u(p)$  quan  $p(x) = 1$ ),

$$u(p) = \sum p(x) u(x), \text{ per tot } p \in P$$

La teoria moderna de la utilitat rebutja el concepte d'utilitat marginalista i suposa que el consumidor té unes preferències donades sobre combinacions de béns (es suposa una elecció simultània), observables (no reduïbles a una mesura de satisfacció)<sup>181</sup>.

---

<sup>181</sup> Com exemple, es pot destacar que en aquest context la llei de decreixement de la relació marginal de substitució reflexa la taxa de substituïbilitat entre un parell (o més) de béns donades les preferències del consumidor (quant d'X per Y accepta un individu per mantenir-se indiferent). La relació és decreixent (com més X, menys costa canviar per Y) i

CAPÍTOL SEGON:

Fonaments de la teoria de jocs:

Una «gramàtica» per a les teories de  
l'acció?



## 2.1. Les relacions de la teoria de jocs amb la racionalitat.

### 2.1.1. Una perspectiva sobre els conceptes de la teoria de jocs.

Deis problemes de fonamentació d'una teoria de la decisió racional -esmentats al capítol anterior-, en aquest segon capítol es passa a l'anàlisi dels problemes que genera un mecanisme formal -la teoria de jocs- que serveix per establir proposicions sobre els resultats agregats del comportament a partir de les decisions individuals. La teoria de jocs és un instrument matemàtic, entre altres de possibles, que facilita l'anàlisi de les decisions individuals interdependents, no paramètriques.

La formulació de decisions que tenen en compte estratègicament les decisions dels altres per establir les seves pròpies es contraposa a les decisions paramètriques, en les quals només es consideren les restriccions i les inferències de la natura (la qual no té capacitat de resposta directa a les nostres accions). El nucli essencial de la teoria de jocs és una estructura isomòrfica, que serveix per a analitzar les situacions d'interdependència i oferir una explicació basada en termes d'anàlisi de l'equilibri. L'equilibri en una situació d'interdependència es produeix quan cap individu pot millorar la seva satisfacció canviant de decisió. Com a síntesi del concepte d'interdependència definit per la teoria de jocs, s'han formulat tres proposicions que recullen el seu sentit essencial: a) La recompensa de cada agent depèn de l'elecció de tots els agents, b) La recompensa de cada agent depèn de la recompensa de tots els agents, i c) L'elecció de cada agent depèn de l'elecció de tots els agents<sup>1</sup>. En termes generals, donades les preferències dels agents, pot ser que

---

<sup>1</sup> Alguns autors defensen la necessitat de modificar aquesta concepció d'interdependència, per incloure els aspectes d'identitat col·lectiva en la definició dels càlculs estratègics dels individus, apart dels aspectes de recompensa i elecció generalment considerats. Entre aquests, es pot mencionar la proclama de R. L. Johnson ("Symbolic Action and the Limits of Strategic Rationality", Political Power and Social Theory, vol. 7, 1988) sobre la incorporació de la mútua dependència com a quart axioma; l'aproximació alternativa a la teoria de jocs de J. Greenberg ("An Application of the Theory of Social Situations to Repeated

un joc no tingui cap punt d'equilibri, que en tingui un, o més d'un. Quan n'hi ha només un, aquest punt és converteix específicament en la solució del joc. Per proveir les solucions, generalment es suposa l'existència d'uns interessos dels jugadors, que serveixen per ordenar les seves preferències. Al mateix temps, també s'adopta la hipòtesi de racionalitat en la «gestió» de les preferències: però aquestes només són les hipòtesis de l'enfocament de l'«elecció racional». Bàsicament, el que la teoria de jocs necessita es un mecanisme generador de decisions estratègiques, segueixi aquest les pautes de la racionalitat o no.

Durant els anys vint, el teorema del minimax va donar inici a la teoria de jocs. Aquest teorema trencava amb la creença anterior de què en qualsevol interacció es podia produir una sèrie infinita de moviments d'un individu envers l'altre per millorar els seus resultats, en no existir una solució d'equilibri per aquestes situacions<sup>2</sup>. El teorema demostrava que sempre existia un conjunt d'eleccions, una per cada individu implicat en la interacció, que portava als participants en la interacció a adoptar decisions estables, ja que cap podia millorar la seva situació desviant-se de la solució d'equilibri, si els altres tampoc ho feien<sup>3</sup>. Un joc té una

---

Games", Journal of Economic Theory, vol. 49, 1989); i fins a cert punt, la proposta de J. Coleman (a "Simulation Games and the Development of Social Theory", Simulation and Games, vol. 20, 1989) sobre els intercanvis no visibles com a fonament d'una part de la teoria social. No obstant, segons la perspectiva defensada, no és necessari introduir transformacions externes, perquè conceptes similars poden ésser integrats internament.

<sup>2</sup> En aquest sentit, l'aportació per la qual es reconeix a John von Neumann com el principal fundador de la teoria de jocs és la seva demostració de què existia un punt d'equilibri en la interacció entre individus racionals a través del teorema del minimax (cf. J. von Neumann, "Zur Theorie der Gesellschaftsspiele", Mathematische Annalen, vol. 100, 1928). Abans de l'article de von Neumann, ja s'havien publicat diversos treballs (Zermelo, Borel) amb diversos teoremes sobre teoria de jocs, però no existia un concepte de solució aplicable de forma general, demostrant que es podia aplicar per cada joc de suma zero de dues persones amb un nombre finit d'estratègies pures per cada jugador (admetent estratègies mixtes). Per una discussió sobre l'article de von Neumann, veure la ressenya de G. Thompson, "John von Neumann", a The New Palgrave. A Dictionary of Economics, vol. 3, 1987.

<sup>3</sup> Així destaca T. Schelling l'equilibri d'una interacció "It has the property that if we both make the corresponding choices, each expecting to other to do so, each has behaved correctly in accordance with his expectations and each has confirmed the other's expectations. (...) There is no other cell in the matrix that can improve the outcome for one player without worsening it for the other" (a "What is Game theory?" (1967), dins Choice and Consequence, Harvard University Press, Cambridge, 1984, pàg. 232).

solució trivial -paramètrica- si cada jugador té una estratègia immillorable independentment de l'estratègia de l'altre jugador (estratègia dominant). Quan no es produeix aquesta situació, el teorema del minimax afirma que cada agent té assegurada en el punt d'equilibri la recompensa mínima dins el conjunt de resultats màxims possibles, considerant que el contrari prengui la pitjor estratègia pels nostres interessos<sup>4</sup>. Per provar aquest teorema, von Neumann va tenir que introduir el concepte de les estratègies mixtes (ponderar amb probabilitats cada possible resultat, i escollir-ne un per atzar dins de cada conjunt) i una proposta nova d'utilitat cardinal, per ordenar les preferències entre distints conjunts ponderats de resultats. Cal destacar que aquest procediment és també el fonament conceptual de la Hipòtesi de la Utilitat Esperada (vid. apt. 1.3.1.). Von Neumann va demostrar que per tots els jocs de suma nul·la existia sempre un o més punts d'equilibri, en el cas que fos possible emprar estratègies mixtes<sup>5</sup>. Una derivació del teorema del minimax va ser l'equilibri estratègic de Nash per a jocs no cooperatius de n-persones, molt més àmpliament utilitzat en l'actualitat degut a la seva major generalitat com solució pels jocs no-cooperatius amb suma variable<sup>6</sup>.

---

<sup>4</sup> Quan en un joc de suma fixa existeix un equilibri amb estratègies pures, coincideixen l'estratègia maximin i la minimax. El maximin i el minimax són dos procediments estratègics per maximitzar resultats, amb orientacions distintes davant el risc (aversió en el primer cas, atracció en el segon). El primer consisteix en buscar el valor mínim de cada interacció possible, i escollir el màxim entre aquests mínims; el mínimax calcula el valor màxim de les interaccions, i escull el mínim entre aquests.

<sup>5</sup> Els jocs de suma nul·la són aquells en què, el que un individu perd, ho guanya l'altre, però el resultat global del joc és constant. Entre moltes altres referències, una demostració modernitzada i pedagògica del teorema es pot trobar a A. Colnan, Game Theory and Experimental Games, Pergamon Press, Oxford, 1982 (pàgs. 273-280) i també a J. Ricart, "Una introducció a la teoria de jocs", Cuadernos Económicos de ICE, vol. 40, 1988, pàg. 31.

<sup>6</sup> J. Nash, "Non-Cooperative Games", Annals of Mathematics, vol. 54, 1951. El concepte d'equilibri de Nash s'aplica als jocs mixtes de conflicte-cooperació, i ha estat la base d'un gran nombre de refinaments als darrers decennis per trobar solucions més perfectes (o ajustades) i evitar la indeterminació que produeix l'existència de molts equilibris. Per a una aproximació recent a una part d'aquestes discussions, vegeu les revisions que fan R. Myerson, "Refinements of the Nash Equilibrium Concept", International Journal of Game Theory, vol. 7, 1978; F. Vega, "Amenazas increíbles, percepciones insostenibles: refinamientos del equilibrio de Nash en juegos dinámicos", Cuadernos Económicos de ICE, vol. 40, 1988; i D. Kreps, "Nash Equilibrium", The New Palgrave. A Dictionary of Economics, vol. 2, 1987. Una acurada conceptualització de les diverses solucions proposades per aclarir les interaccions estratègiques es pot trobar a R. Aumann, "What is Game Theory Trying to Accomplish?", a K.J. Arrow i S. Honkapohja (Eds.), Frontiers of Economics, Blackwell, Oxford, 1985.

Tot i que la teoria de jocs és una teoria matemàtica, basada en un isomorfisme sense contingut teòric fàctic, aquesta ha contribuït a desenvolupar una perspectiva analítica a les ciències socials, amb els seus problemes específics. Això és degut a que la teoria es va crear directament per ésser aplicada a fenòmens polítics, socials o econòmics, i encara que no es tracta pròpiament d'una teoria social (l'estructura matemàtica de la teoria de jocs ha estat aplicada a altres ciències, com la biologia) permet definir amb precisió conceptes com amenaça o conflicte, que normalment encobreixen un ampli ventall de significats<sup>7</sup>.

Els individus presents en una situació d'interacció s'anomenen jugadors. Els models més simples assumeixen l'existència de dos jugadors, tot i que no hi ha greus problemes formals per generalitzar els models a 'N' jugadors. L'aplicació empírica dels models amb un nombre elevat de jugadors ofereix el problema de que els resultats possibles incrementen exponencialment. Això redueix el gran poder d'abstracció que ofereixen els jocs amb pocs jugadors, a no ser que es defineixin supòsits força restrictius. Una discussió important s'ocupa de les característiques dels subjectes que efectuen decisions estratègiques. Normalment, s'assumeix que són individus, però en moltes anàlisis s'han considerat com a jugadors a organitzacions, Estats o altres formes d'agents col·lectius. En aquest sentit, es presenta el problema de la consistència interna de les eleccions de cada actor, però no té perquè impossibilitar l'anàlisi<sup>8</sup>. Un altre tema vinculat amb els jugadors es refereix als

---

<sup>7</sup> Els principals elements de la teoria de jocs constitueixen quelcom més que simples instruments d'una teoria matemàtica. També defineixen una perspectiva interpretativa de la societat -o dels fenòmens concrets que s'analitzen en aquesta-, partint de l'acció individual com element clau de les explicacions que es formulen. Com diu T. Schelling, aquesta proveeix un esquema interpretatiu: "*game theory is more than a 'theory', more than a set of theorems and solutions; it is a framework for analysis*" (T. Schelling, "What is Game Theory", art. cit., 1984, pàg. 221).

<sup>8</sup> No cal pensar que com més 'compost' sigui un jugador menor és el seu nivell de racionalitat. Tot depèn bàsicament de les característiques de l'anàlisi que es porti a terme, el seu nivell d'abstracció, etc.. Dins d'un individu és possible l'existència d'un conflicte de preferències, que d'alguna forma el propi individu pot intentar conduir (com tracten T. Schelling, "Self-Command in Practice, in Policy, and in a Theory of Rational Choice", *American Economic Review*, vol. 74, 1984; i J. Elster, "Weakness of Will...", art. cit., 1986); i també dins d'una organització (H. Wiesenhal, "Unsicherheit und Multiple-Self-Identität. Eine Spekulation über die Voraussetzungen strategischen Handelns", Max-Planck-Institut für Gesellschaftsforschung, Discussion Paper 90/2, Colonia, 1990), o a l'Estat, entre les seves diverses

supòsits informacionals que cal considerar. En principi, la teoria de jocs assumeix que existeix informació completa. Això implica que cada jugador té coneixement de les seves preferències i de les dels seus contrincants, així com de tots els possibles resultats i, per tant, aquest pot deduir les estratègies que adoptaran els altres jugadors. Recentment s'han desenvolupat també models per tractar altres condicions d'informació, que són tractats més endavant.

Cada jugador té un conjunt d'estratègies possibles. Una estratègia és pot considerar com un moviment o una acció<sup>9</sup>, en forma de jugada -o successió de jugades- que inclou la previsió sobre la reacció del contrari (o la reacció pròpia a la provocació del contrari):

*"A complete plan for that player to play the game, as function of what he observes during the course of play, about the play of others and about chance occurrences affecting the game"*<sup>10</sup>.

S'assumeix que l'estratègia pròpia s'escull considerant les estratègies que el contrari té la possibilitat de realitzar, amb la intenció d'obtenir el millor resultat possible entre els definits a l'ordenació de preferències. Es considera que la forma estratègica d'un joc és una funció que associa a cada conjunt d'estratègies un conjunt de

---

administracions (G. Allison, Essence of Decision, Little, Brown and Co., Boston, 1971). Els conflictes de preferències i les limitacions que apareixen a cada nivell de decisió no impliquen necessàriament que incrementin de forma més que exponencial les dificultats de formalització en relació a un jugador homogeni. Es pot interpretar que és precisament la natura de cada aplicació o interpretació la que determina la capacitat (o incapacitat) explicativa de la simplificació que s'efectua.

<sup>9</sup> Cada ocasió concreta de joc s'anomena una **jugada**, "*which is described by specifying, out of all possible, allowable moves, the sequence of choices actually made by the players*" (O. Morgenstein, "Game Theory. Theoretical Aspects", International Encyclopaedia of Social Sciences, vol. 6, 1968, pàg. 63).

<sup>10</sup> R. Aumann, "Game Theory", The New Palgrave. A Dictionary of Economics, vol. 2, 1987, pàg. 460.

resultats<sup>11</sup>.

La hipotètica confrontació entre totes les estratègies de cada actor (o jugador) produeix un conjunt de possibles resultats, que són valorats pels jugadors mitjançant assignacions d'utilitat o escales ordinals de preferències. En una situació de risc, es suposa que l'agent coneix les probabilitats de tots aquests resultats, o bé els assigna probabilitats subjectives (vid. apt. 1.3.2.). El problema de la comparança entre els resultats porta a dues discussions paral·leles. El primer, sobre la comparança entre diferents resultats per un mateix jugador. Per justificar que un individu prefereix més un resultat que un altre, a la teoria de jocs s'han de construir ordenacions de preferències que segueixin una orientació d'interès. Es poden emprar preferències ordinals, però donat que els valors numèrics són necessaris per buscar els punts d'equilibri quan apareix el risc, cal introduir una mesura continua de les preferències, la qual cosa incorpora diverses dificultats, ja discutides anteriorment<sup>12</sup>. El segon problema tracta de la comparativitat entre jugadors per un resultat. La utilitat «no transferible», en el sentit de no poder fer comparacions entre ratis d'increments d'utilitat, sembla tenir entre els especialistes de teoria de jocs una certa homogeneïtat en considerar què és més desitjable perquè és més realista, però és molt més complexa formalment<sup>13</sup>.

Una primera distinció sobre les característiques bàsiques dels jocs s'estableix

---

<sup>11</sup> La utilització de les estratègies mixtes, desenvolupades per aconseguir solucions generals dels models de jocs, tot i que s'utilitzen per provar els teoremes sobre l'equilibri, ha estat força criticada des del punt de vista de la seva plausibilitat empírica, perquè cal pressuposar que els subjectes l'adoptarien com a forma de decisió en les seves interaccions (Cf. F. Scharpf, "Games Real Actors Could Play: The Problem of Complete Information", Max-Planck-Institut für Gesellschaftsforschung, Discussion Paper 89/9, Colonia, 1989). Scharpf reconeix que no acceptar les estratègies mixtes implica rebutjar moltes solucions possibles, però indica que "*what we lose in analytical power, we probably gain in greater realism*" (pàg. 13).

<sup>12</sup> Vid. la secció 1.3. Vegeu també T. Schelling, "What is Game Theory?", art. cit., 1984, pàg. 238. Resta la possibilitat de proposar altres mecanismes de decisió, tal com es discuteix en propers apartats.

<sup>13</sup> R. Aumann (a "Game Theory", art. cit., 1987, pàg. 466) destaca la importància dels models d'utilitat transferible com pas previ indispensable per poder bastir posteriorment els models més complexos d'utilitat no-transferible.

entre els jocs cooperatius i els no-cooperatius. S'assumeix que als jocs cooperatius es possible abans de l'inici del joc lligar mútuament als jugadors sobre el conjunt de possibles resultats, generalment gràcies a les decisions adoptades en una comunicació prèvia entre aquests. En aquest procés, es suposa també que no hi ha distorsions dels propis interessos. La millor solució possible a obtenir en termes de benefici pels participants és un acord mutu sobre totes les eleccions estratègiques. Quan hi ha més de dos jugadors poden aparèixer formes de coalició entre aquests, per obtenir millors resultats. Entorn de les possibilitats de coalició existeixen diversos conceptes de solució que intenten aproximar prediccions sobre les coalicions més probables que es formaran en produir-se una interacció. D'altra banda, els jocs no-cooperatius presenten una relació de no compromís abans del procés de joc. Es pot produir comunicació, però sense estar garantida l'absència de traïció en cas d'arribar-se a un acord previ. Així, en iniciar-se el joc, cada agent desconeix l'elecció real dels contraris, fins el moment en què és realitza la jugada. El teorema del minimax i l'equilibri de Nash, així com els refinaments posteriors d'aquestes dues solucions, són els procediments analítics que s'empren per descobrir l'existència d'equilibris en la interacció no-cooperativa que s'analitza<sup>14</sup>.

Una altra forma de classificar els jocs és a partir de la suma de pagaments o resultats globals. La categoria més general són els jocs de suma variable, on els resultats globals -no constants- i la seva distribució depenen de les decisions estratègiques dels jugadors<sup>15</sup>. Com un cas més simple, existeixen els jocs amb un resultat constant, en els quals la interacció produeix un resultat -negatiu o positiu-, i apareix un conflicte per la repartició d'aquest resultat (per exemple, un joc esportiu com els escacs, o bé es guanya, o bé es perd, però no poden guanyar els dos

---

<sup>14</sup> Aquests tipus d'equilibri es basen en la noció d'acord mitjançant el qual s'autoreforquen, en el sentit que és el millor interès de cada jugador acceptar l'equilibri-acord que s'indica, ja que els altres jugadors també ho veuen (racionalment) igual.

<sup>15</sup> Com diu J. Elster "*there are not many real-life examples of constant-sum games. The struggle over the division of the total can be more or less costly*", ("Introduction", a J. Elster, (Ed.), Rational..., op. cit., 1986, pàg. 8).

contrincants) . Dins aquests, hi ha aquells que tenen un resultat global nul -jocs de pur conflicte-, en el sentit que la interacció es produeix per a obtenir la major part possible d'un resultat fix (el joc en sí no produeix cap resultat -repartir un pastel, o distribuir una tasca de neteja). La seva estructura és simple: tot allò que un jugador guanya, ho perd l'altre.

### 2.1.2. L'evolució de la teoria de jocs en el context de les ciències socials.

El desenvolupament de la teoria de jocs es va produir durant els anys quarantes, a partir de l'obra de J. von Neumann i O. Morgenstern, que varen introduir aquesta teoria com una nova àrea d'investigació científica, tant al camp de les matemàtiques com al de les ciències socials<sup>16</sup>. Des dels anys quarantes s'ha desenvolupat un ampli camp de teories matemàtiques especialitzades en buscar condicions per a la consecució d'equilibris, definir normes per a la realització d'acords estables, proposar i provar teoremes sobre les accions dels jugadors racionals, etc.<sup>17</sup>. La línia de recerca majoritària en teoria de jocs ha estat protagonitzada per matemàtics i, per tant, caracteritzada per l'exigència de simplicitat i abstracció que aquests persegueixen sempre a les seves formulacions.

El refinament formal -i conceptual- de la teoria de jocs ha tingut un intens creixement als darrers decennis, mentre que el desenvolupament de les seves

---

<sup>16</sup> Així ho reconeix R. Aumann, en considerar que "*the book broke fundamental new ground in several directions. These include the notion of a cooperative game, its coalitional form, and its von Neumann-Morgenstern stable sets. Though axiomatic expected utility theory had been developed earlier (...), the account of it given in this book is what made it 'catch on'.*" ("Game Theory", art. cit., 1987, pàg. 463).

<sup>17</sup> El desenvolupament va ser continuat des de principis de segle (començant pel teorema de Zermelo, 1912), amb un salt qualitatiu durant els anys quaranta-cinquantes amb la monumental obra de J. Von Neumann i O. Morgenstern i la formulació d'alguns dels teoremes més importants (per exemple, el teorema de Nash, l'any 1950, o el de Shapley, l'any 1951) que va donar consistència a la teoria. El salt quantitatiu es va produir durant els anys setantes, en augmentar molt el nombre d'investigadors dedicats a aquest camp, cosa que el va convertir en un camp de recerca matemàtica important (això ho reflecteix per exemple l'aparició l'any 1972 de la revista International Journal of Game Theory) i també va començar a desenvolupar-se la recerca empírica aplicada.



aplicacions a les ciències socials, tot i ésser proclamat des dels anys cinquantes<sup>18</sup>, ha rebut generalment escassa atenció fins anys recents<sup>19</sup>. D'altra banda, la seva aplicació a la teoria econòmica ha tingut fases d'evolució molt ràpida i fort entusiasme, i altres de pràctica paralització de la recerca<sup>20</sup>. En general -tot i existir excepcions importants-, s'ha mantingut en aquest camp la tendència a aplicar directament les teories matemàtiques que s'anaven produint per interpretar la realitat econòmica<sup>21</sup>. En compartir la hipòtesi de racionalitat, ha estat senzill incorporar resultats provinents de la teoria de jocs als models econòmics, i inclús, en els darrers anys, produir teoria «matemàtica-econòmica» en alguns dels camps teòrics més deficitaris de la teoria econòmica<sup>22</sup>. Cal mencionar que no totes les línies de recerca

---

<sup>18</sup> Especialment per T. Schelling, com expressava clarament l'any 1967: "*There is an unusual dichotomy between the subtle, elegant, mathematical accomplishments of game theorists and the interests of social scientists*" (a "What is Game Theory?" art. cit., 1984, pàg. 241).

<sup>19</sup> El text de T. Schelling *The Strategy of Conflict* (Harvard University Press, Cambridge, 1960), tot i molt difús (per exemple, traduït al castellà l'any 1964), durant els anys següents va ésser poc treballat en els seus aspectes metodològics pels científics socials (amb la relativa excepció del especialistes en relacions internacionals). Un curiós exemple en aquest sentit és la menció que en fa R. Boudon, (a *L'analyse mathématique des faits sociaux*, Plon, Paris, 1970<sup>2</sup>) de la teoria de jocs "*comme des théories spéculatives analysant des situations délibérément idéalisées (...) dont les applications à l'analyse sociologique inductive sont, à ce jour, pratiquement inexistantes*" (pàgs. 16), tot justificant l'exclusió que fa de la teoria de jocs en el seu tractat.

<sup>20</sup> Per exemple, després d'un fort entusiasme sobre les possibilitats d'aplicació a les ciències econòmiques i polítiques durant els anys cinquantes, com ho reflectia l'obra actualment clàssica de D. Luce i H. Raiffa (*Games and Decisions*, op. cit., 1957), que incorporava tot el potencial de propostes de solucions desenvolupades en els anys immediatament anteriors, suggerint exemples d'aplicació i interpretació, el desenvolupament d'aplicacions va restar limitat a unes petites àrees de la teoria econòmica per reduïts nuclis d'especialistes. Ha estat posteriorment, durant els anys setantes, i sobretot durant la dècada dels vuitantes, quan s'han elaborat moltes aplicacions de la teoria de jocs a la teoria i la política econòmica, especialment a la teoria econòmica neoclàssica. Per a una panoràmica recent, vegeu els llibres de J. Friedman (*Game Theory with Applications to Economics*, Oxford University Press, Nova York, 1990<sup>2</sup>) i de E. Rasmusen, *Games and Information*, Basil Blackwell, Oxford, 1989). Però també s'ha desenvolupat a altres camps, com la ciència política (P. Ordeshook, *Game Theory and Political Theory. An Introduction*, Cambridge University Press, Cambridge, 1986) o l'economia política (M. Shubik, *A Game Theoretical Approach to Political Economy*, MIT Press, Cambridge, 1984).

<sup>21</sup> Així ho argumenta O. Morgenstern: "*The Theory of Games is a mathematical discipline designed to tread rigorously the question of optimal behavior of participants in games of strategy and to determine the resulting equilibria*" (a "Game Theory...", art. cit., 1968, pàg. 62).

<sup>22</sup> Tant el creixent nombre d'articles dedicats a la teoria de jocs en les revistes més prestigioses de la teoria econòmica (el *Journal of Economic Theory*, *Mathematics of Operations Research* o *Econometrica*, entre d'altres), com l'aparició l'any 1988 de la revista *Games and Economic Behavior* és una mostra de l'interès creixent dels teòrics de l'economia per aquesta qüestió.

sobre l'aplicació de la teoria de jocs han seguit aquesta evolució. En aquest sentit, l'obra de T. Schelling apareix com un dels punts de referència bàsics d'altres línies d'aplicació. Aquest autor considera que la hipòtesi de racionalitat pot servir per a interpretar l'acció, però ni molt menys l'única hipòtesi possible a preveure en termes teòrics, com ho il·lustra en la seva formulació dels punts focals<sup>23</sup>.

En els darrers decennis s'han desenvolupat moltes àrees de treball dins la teoria de jocs. Algunes ja havien estat obertes durant els anys cinquantes, mentre que altres s'han iniciat més recentment. Tan sols per a situar el terreny de la discussió i sense cap intent d'exhaustivitat, es comenten sintèticament algunes d'aquestes àrees<sup>24</sup>. Al terreny dels jocs no-cooperatius, cal destacar el desenvolupament de models amb jocs d'informació incompleta a partir dels anys seixantes, seguint la proposta d'Harsanyi de substituir els jocs d'informació incompleta per jocs amb informació completa però imperfecta (que pot tenir errors), per afrontar el freqüent problema empíric de la informació. Es suposa l'existència de jugadors que traslladen els desconeixements mutus en desconeixement sobre les funcions d'utilitat. Els jugadors realitzen estimacions sobre les estratègies del contrari que desconeixen, i apliquen les probabilitats per realitzar l'estimació dels resultats<sup>25</sup>. Aquest tipus de jocs ha seguit dues línies principals de

---

<sup>23</sup> T. Schelling defineix els «punts focals» senyalant que quan els jugadors no troben en la configuració matemàtica del joc un punt d'equilibri, que actuï com a indicatiu per a coordinar les seves accions (perquè no existeix, o és excessivament complicat, o perquè hi ha massa punts d'equilibri), els jugadors cerquen algun signe («punt focal») per a coordinar les seves estratègies fora de la configuració matemàtica. Schelling proposa l'aparició de diverses possibilitats: l'aparició de mecanismes d'insinuació, la consideració dels costos de comunicació o la recerca de pistes discriminatòries al terreny de la interacció; apareix en conjunt un ampli camp d'interpretació i teorització (Cf. "Reinterpretación de un concepto de solución para juegos 'no-cooperativos'", a *La Estrategia del Conflicto*, Tecnos, Madrid, 1964).

<sup>24</sup> Actualment, per la seva amplitud i diversitat, ja és gairebé impossible, inclús pels propis especialistes matemàtics en teoria de jocs, estar al corrent de totes les línies d'investigació que es treballen, constituint un dels camps formals en més ràpida expansió a les ciències socials. Així ho comenta R. Aumann, destacant que "*as in other disciplines, it is becoming difficult to keep abreast of the important developments*" ("Game Theory", art. cit., 1987, pàg. 476).

<sup>25</sup> J. Harsanyi, "Games with Incomplete Information Played by Bayesian Players", *Management Science*, vol. 13, 1967-68. Per evitar una situació d'incertesa, en el sentit que es perd la capacitat d'assignar utilitats, perquè no tots els possibles resultats són coneguts (no totes les estratègies del contrari

desenvolupament. La primera, els jocs repetits amb informació incompleta, analitza com les modificacions en la informació que disposa cada jugador en diversos moments del joc (mitjançant «aprenentatge», o la formació de «reputacions», per exemple) porta a transformacions importants dels equilibris que es poden assolir<sup>26</sup>. La segona, els processos de comunicació en els jocs estàtics amb informació incompleta: es proposen mecanismes de comunicació i coordinació (anomenats «principi de revelació») que serveixin per a que els jugadors no tinguin cap incentiu per mentir<sup>27</sup>, donat que la informació sobre preferències i resultats és de domini privat -en cas d'existir- i contínuament els agents han de realitzar estimacions sobre aquesta informació<sup>28</sup>.

Una altra línia de recerca important sobre els jocs no cooperatius són jocs repetits (finites) i superjocs (que es repeteixen infinitament), que analitzen les relacions entre conflicte i cooperació en contextos dinàmics; aquests jocs tenen punts de connexió amb els jocs d'informació incompleta, encara que el camp és força més ampli<sup>29</sup>. L'aparició de nous equilibris és un tema clau de discussió en aquests jocs,

---

són conegudes). Harsanyi suposa l'existència d'equilibris bayesians, a partir de la hipòtesi que els jugadors, en relació a les funcions d'utilitat que desconeixen, apliquen el teorema de Bayes sobre distribució de probabilitat, per maximitzar posteriorment la utilitat esperada. Per una aproximació actualitzada a aquest camp de la teoria de jocs, vegeu J. Ricart, "Juegos con información incompleta", Cuadernos Económicos de ICE, vol. 40, 1988; i també H.-M. Wallmeier, "Games and Incomplete Information -A Survey-", Institute of Mathematical Economics, Working Paper # 161, Bielefeld, 1986.

<sup>26</sup> Sobre aquesta qüestió, vegeu D. Kreps, P. Milgrom, P. Roberts i R. Wilson, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma", Journal of Economic Theory, vol. 27, 1982; i S. Hart, "Nonzero-Sum Two-Person Repeated Games with Incomplete Information", Mathematics of Operations Research, vol. 10, 1985.

<sup>27</sup> Com ho il·lustra el cas exposat per G. Akerlof sobre el mercat de cotxes de segona mà, on el venedor coneix molt millor l'estat del cotxe que el possible comprador, i aprofita aquesta informació per fixar arbitràriament un preu superior (a "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism", Quarterly Journal of Economics, vol. 84, 1970).

<sup>28</sup> Sobre el tema, vegeu D. Kreps i R. Wilson, "Sequential Equilibria", Econometrica, vol. 50, 1982; i R. Myerson, "Two-Person Bargaining Problems with Incomplete Information", Econometrica, vol. 52, 1984 (vid. apt. 2.2.4.).

<sup>29</sup> Una panoràmica general es pot trobar a R. Marimon, "Introducción a los juegos dinámicos", Cuadernos Económicos de I.C.E., vol. 40, 1988.

així com les condicions estratègiques sota les quals els equilibris són possibles. La discussió es centra intensament en dos temes: el teorema bàsic sobre la cooperació, el «Folk Theorem»<sup>30</sup>, i l'aplicació d'un model clàssic de la teoria de jocs -el dilema del presoner- als cas dels jocs repetits i superjocs<sup>31</sup>, amb l'aparició de sorprenents resultats sobre seqüències estratègiques<sup>32</sup>.

Finalment, una tercera gran àrea a destacar és la que tracta de les actituds cooperatives en els jocs d'estratègia. Hi ha dos enfocaments, l'anàlisi de les situacions de negociació prèvies al desenvolupament del joc; i la discussió de propostes normatives que poden conduir a solucions arbitrals. Bàsicament, la teoria dels jocs cooperatius analitza les possibilitats de formació de coalicions, suposant que les coalicions competitives, un cop formades, es mantenen amb l'objectiu d'assolir el màxim resultat. Això és degut a què aquesta branca de la teoria de jocs parteix de la hipòtesi que els jugadors tenen instruments que els permeten arribar a

---

<sup>30</sup> Aquest teorema -formulat anònimament a la fi dels anys cinquantes (per això s'anomena «popular»)- postula que quan el joc és repetit indefinidament, el resultat d'equilibri afavoreix la cooperació, ja que els acords es reforcen per un mecanisme que pena els que es desvien de la pauta d'equilibri en els diferents moments del joc. En jocs finits la discussió es centra sobre els efectes de la taxa de descompte del futur que adopten els participants en el joc. Per a un exemple d'aquesta discussió, vegeu entre d'altres D. Fudenberg i E. Maskin, "Folk Theorems for Repeated Games with Discounting and Incomplete Information", *Econometrica*, vol. 54, 1986; S. Sorin, "On Repeated Games of Complete Information" *Mathematics of Operations Research*, vol. 11, 1986. i W. Güth, W. Leiniger i G. Stephan, "On Supergames and Folk Theorems. A Conceptual Discussion", *Diskussionsschriften* 127, Heidelberg Universität, 1989.

<sup>31</sup> R. Axelrod ( a La evolución de la cooperación, Alianza, Madrid, 1986) mostra, en una competició de jocs iteratius del dilema de presoner finit, com l'estratègia seqüencial més simple, que no té necessitat d'un gran càlcul de mecanismes instrumentals, és la que aconsegueix triomfar davant estratègies molt més sofisticades (però adaptades a situacions molt concretes).

<sup>32</sup> En aquestes teories s'introdueixen supòsits sobre l'aprenentatge al llarg dels jocs, introduint la selecció de les estratègies per mecanismes de prova i error, i considerant que els actors poden distingir entre resultats més o menys satisfactoris. Altres temes són investigats en els jocs repetits: els mecanismes markovians -les accions dels agents tenen efecte directa sobre el futur desenvolupament del joc, transformant les seves regles-, el disseny de mecanismes -caracterització normativa de les solucions més eficients dins el marc de les factibles-, l'anàlisi de les accions no observables -induir a partir de les distribucions de resultats del joc com es desvien alguns agents de les pautes d'equilibri- (Cf. R. Marimon, "Introducción a los juegos..", art. cit., 1988).

compromisos entre ells<sup>33</sup>. Als jocs cooperatius es poden aplicar altres propostes d'equilibri -o espais d'equilibris possibles- distintes de les aplicables als jocs no-cooperatius<sup>34</sup>. D'altra banda, una important branca teòrica molt connectada amb els jocs cooperatius és la teoria de la negociació (bargaining). La teoria recull els procediments de negociació per arribar a un punt de solució, quan existeix un conflicte d'interessos i moltes solucions possibles<sup>35</sup>.

Sobre la relació entre la racionalitat i la teoria de jocs, es pot afirmar que la teoria de jocs fa ús dels pressupòsits de la Hipòtesi de la Utilitat Esperada en la major part dels seus desenvolupaments teòrics. *La Hipòtesi es emprada en la*

---

<sup>33</sup> A. Mas Colell, "Algunos comentarios sobre la teoria cooperativa de los juegos", Cuadernos Económicos de ICE, vol. 40, 1988, pàgs. 143-5. Mas Colell destaca també que en un sentit general, la teoria cooperativa dels jocs serveix per delimitar un espai d'anàlisi dins la teoria més general i bàsica de la dels jocs no-cooperatius, i que el seu valor és degut a què aconduïx a precisar millor les possibilitats de trobar equilibris determinats. Per a una aproximació general als jocs cooperatius, vegeu E. Moulin, Game Theory in the Social Sciences, New York University Press, Nova York, 1986.

<sup>34</sup> Cal destacar tres conceptes d'equilibri en els jocs cooperatius: el *core*, formulat per M. Shubik a partir d'una proposta d'Edgeworth; el *conjunt estable* proposat per von Neumann i Morgenstern; i la mesura del *valor* d'un joc, definida per S. Shapley. El *core* implica la idea de predicar l'existència de resultats definits per la seva superioritat sobre la resta (en termes de preferències de tots els agents implicats en la interacció). El *conjunt estable* defineix un espai de resultats no dominat per cap altre resultat extern possible. Finalment, el concepte de *valor* d'un joc, que és el que té una gamma més àmplia d'aplicacions en ciències socials, construeix un índex del poder de cada jugador, basat en la capacitat de cada jugador per formar part de les coalicions possibles en el joc, en relació amb el nombre de coalicions que no poden formar part, o a les quals la seva presència no té cap pes significatiu per modificar les decisions. L'índex parteix de l'observació que el poder d'un jugador -o una coalició- augmenta de forma no lineal quan s'incrementa el nombre de vots que controla (vid. L.S. Shapley i M. Shubik, "A Method for Evaluating the Distribution of Power in a Committee System", American Political Science Review, vol. 48, 1954). Per una aproximació conceptual sobre l'«estat de l'art» de les diferents propostes d'equilibri realitzades, vegeu R. Aumann, "What is Game Theory..." art. cit., 1985.

<sup>35</sup> Els vincles de les teories de la negociació -un tema clàssic de l'economia- amb la teoria de jocs parteixen de la proposta de Nash per definir un punt d'equilibri normatiu, dins l'espai de possibles acords a realitzar. Els axiomes que Nash defineix són els d'eficiència, simetria, invariància lineal i independència d'alternatives irrelevantes (cf. J. Nash, "Two Person Cooperative Games", Econometrica, vol. 21, 1953). Les discussions sobre «bargaining» dins la teoria de jocs, s'han centrat posteriorment en la proposta de Nash i la presentació d'alternatives -generalment normatives-. Per a una aproximació, veure J. Harsanyi, "Bargaining", a The New Palgrave. A Dictionary of Economics, vol. 1, 1987 i C. Ponsati, "Juegos de negociación", a Cuadernos Económicos de ICE, vol. 40, 1988. Una visió més àmplia de la teoria de la negociació, integrant els temes de la teoria de jocs però no esgotant-se en ells, es pot trobar en la revisió que fa J. Elster en el capítol 2 del seu llibre The Cement of... op. cit., 1989. Els temes de la formulació d'incentius, amenaces, promeses en la negociació tan sols han estat parcialment integrats en la teoria de jocs, i fora d'aquesta presenten una gran varietat de matisos.

*formulació d'una teoria general de l'equilibri en jocs no cooperatius mitjançant el teorema del minimax i el teorema de Nash, que inclouen l'ordenació de preferències sobre estratègies mixtes. Els teoremes es poden entendre com a propostes per integrar les eleccions racionals en els models d'interacció de la teoria de jocs. Però formalment no són sempre necessaris, ja que qualsevol funció matemàtica amb una distribució de probabilitat pot determinar uns resultats anàlegs al context de la teoria de jocs*<sup>36</sup>. En tot cas, el problema es traslladaria a l'hora de determinar els equilibris. Aquesta qüestió es tractada en detall als propers apartats, seguint l'intent de separar l'utilització normativa de la racionalitat de l'aplicació de models de decisió racional per a predir el comportament observat. Com els fonaments de la teoria moderna de la decisió racional sota risc són fruit dels elements teòrics que proposaven en la seva formulació de la teoria de jocs von Neumann i Morgenstern<sup>37</sup>, cal entendre que aquest origen comú ha determinat de forma sensible la continuïtat dels lligams entre els dos desenvolupaments teòrics. Curiosament, s'han mantingut els lligams, però ja des dels seus inicis els investigadors en teoria de jocs s'han preocupat relativament poc dels fonaments de la teoria. A més, la teoria de la decisió racional ha seguit també un desenvolupament propi. Aquest desenvolupament no s'ha vinculat específicament a la teoria de jocs, sinó que deixant a banda la pròpia discussió interna, s'ha aplicat a les necessitats de fonamentació d'amplis camps de la microeconomia, en principi no relacionats necessàriament amb la teoria de jocs (com la teoria de la informació, temes d'organització industrial, o aspectes de la teoria de l'equilibri general). L'evolució mencionada ha portat a què es suposi el manteniment de la vinculació entre les dues teories, sense qüestionar excessivament la continuïtat dels seus vincles; així, les interpretacions estàndard suposen que la teoria de jocs assumeix els postulats de la

---

<sup>36</sup> B. Peleg, "The Independence of Game Theory of Utility Theory", Bulletin of the American Mathematical Society, vol. 72, 1966. Aquesta qüestió es discuteix a l'apt. 2.2.4..

<sup>37</sup> Però els autors matisaren que "We wish to concentrate on one problem -which is not that of the measurement of utilities and of preferences- and we shall therefore attempt to simplify all other characteristics as far as reasonably possible", (a Theory of Games... op.cit., 1953<sup>3</sup>, pàg. ).

teoria de la decisió racional tal com s'han presentat al capítol anterior<sup>38</sup>. Actualment, en trobar-se la formulació de la Hipòtesi de la Utilitat Esperada en una situació de crisi important, amb una multitud de formulacions alternatives que qüestionen diversos dels seus axiomes (vid. apt. 1.3.3.), aquest tipus d'iniciatives «transformadores» comencen a ser també aplicades a la teoria de jocs.

La Hipòtesi de la Utilitat Esperada, o possibles variacions d'aquesta, no són sempre absolutament necessàries per a aplicar la teoria de jocs, inclús en el cas de mantenir la hipòtesi de racionalitat per explicar com formulen les estratègies els jugadors. Si una decisió s'expressa amb preferències ordinals i sobre temes definits i concrets, poden modelitzar-se sense necessitat d'una ordenació cardinal de les preferències. Abanda que, per trobar un equilibri, no sempre es necessari aplicar la probabilitat, si aquest existeix en estratègies pures<sup>39</sup>. Diversos intents de superar aquests problemes constitueixen una de les discussions importants de la resta del capítol. D'altra banda, resta una pregunta prèvia a totes aquestes argumentacions que no ha estat discutida perquè és la més vinculada a la fonamentació de la teoria de jocs. *Què passa amb el concepte d'equilibri<sup>40</sup>? Es pot rebutjar la utilització de la hipòtesi de racionalitat mantenint la d'equilibri? Quines serien les conseqüències?* Per afrontar aquestes qüestions cal discutir l'estatus metodològic de la teoria de jocs dins les ciències socials.

---

<sup>38</sup> Prenent una mostra dels tractats més rellevants de la darrera dècada sobre la teoria de jocs (per exemple, G. Owen, *Game Theory.. op. cit.*, 1982; G. Moulin, *Game Theory for...op. cit.*, 1986; i H. Hamburger, *Games as Models of Social Phenomena*, Freeman, San Francisco, 1979), tots presenten un capítol molt similar, en el qual es presenten de forma resumida els postulats formals de la teoria estàndard de la utilitat, sense comentar gairebé la discussió existent i els problemes de la teoria de la utilitat.

<sup>39</sup> Altra vegada és T. Schelling qui formula amb més claredat la qüestió: "*The need for numerical values arises only in the presence of uncertainties of this sort (and only when the number of alternative outcomes is at least three), when one has to place his bets in a probabilistic environment*" (a "What is Game Theory?", art. cit., 1984, pàg. 237).

<sup>40</sup> Aumann considera els conceptes de solució (punts d'equilibri) com a indicadors, més que com a prediccions ("Game Theory", art. cit., 1987, pàg. 464). Per Aumann cada concepte representa una diferent aproximació o punt de vista als resultats possibles, sense defensar l'existència d'una funció que determini un resultat exacte, ja que el resultat depèn de quins elements es tinguin en compte per calcular les solucions dels jocs (excepte en els jocs de suma zero i dues persones, en els quals sí apareix una clara solució racional).

## 2.2. La vinculació entre acció i teoria de jocs.

### 2.2.1. Ciència social i teoria de jocs.

Plantejar en termes generals les relacions entre les ciències socials i la teoria de jocs pot semblar relativament innecessari, ja que hom podria argumentar que només cal adoptar el que sigui interessant d'un instrumental matemàtic per a bastir interpretacions substantives sobre els processos socials. Malgrat tot, la pròpia potencialitat de la teoria de jocs per aproximar-se a tot tipus d'interdepències, porta a què els seus àmbits i formes d'aplicació siguin molt diferents, introduint certa confusió sobre els diferents plans d'aplicabilitat de la teoria, les relacions entre aquests plans i els objectius d'explicació que proposen les distintes formes d'instrumentalització. Per aclarir algunes d'aquestes confusions, en aquest apartat es proposa una sistematització dels 'estils' d'aplicació de la teoria de jocs a les ciències socials.

Les relacions entre teoria de jocs i ciència social s'estableixen a través de dos plans distints, i l'encreuament de les posicions oposades en cada pla il·lustra quatre possibles àrees de treball, cada una amb pressupòsits propis i amb uns objectius de recerca distints. Com a primer pla, es discuteix la inclusió de distintes formes de comportament individual en el marc analític de la teoria de jocs (amb els models de decisió racional entesos com a casos particulars)<sup>41</sup>. En el segon pla, s'analitzen els objectius que es pretenen amb la utilització de la teoria de jocs, independentment del model d'home escollit: defensar una aplicació exclusiva de la teoria de jocs per a la construcció de teories fàctiques a les ciències socials; o bé acceptar la compatibilitat amb elements externs a aquesta teoria instrumental en l'elaboració d'interpretacions.

---

<sup>41</sup> La defensa de les normes socials fora d'una interpretació racionalitzadora és per exemple l'argument recent de J. Elster, quan destaca les limitacions que porta l'enfocament racionalitzador (vid. entre altres J. Elster, "Social Norms and...", art. cit., 1990).



I. La primera àrea de treball es pot definir a través de la coincidència entre la utilització exclusiva del model de decisió racional convencional<sup>42</sup> amb l'objectiu d'elaborar explicacions del comportament humà basades només en la teoria de jocs. Des d'aquesta perspectiva, apareixen moltes limitacions a interpretacions plausibles, degut a l'escàs realisme dels supòsits utilitzats (com una capacitat de càlcul sobrehumana)<sup>43</sup>. Per exemple, es pot produir una situació de dilema del presoner quan es produeix foc en un lloc tancat i ple de gent que no es coneix entre si: quan no existeix la confiança, en la recerca d'una salvació individual la gent produeix el caos i augmenten molt més els accidents<sup>44</sup>. La dificultat es integrar en el model els factors emocionals que aconduïxen a la confiança, per explicar com els individus poden assolir racionalment el millor resultat en aquesta situació. Davant d'aquestes dificultats, un exemple de la defensa clàssica d'aquesta perspectiva és degut a R. Aumann, qui planteja una refinada argumentació sobre l'aplicació de la teoria de jocs:

*"There is apparently some kind of generalized invisible hand at work. While in any given situation an individual may well act irrationally, there seems to be a cumulative effect of numbers and time and learning that pushes people 'in general' in the direction of rational decision making"*<sup>45</sup>.

Implícitament, l'argument adopta la clàssica versió extrema de la falsedat que defensen alguns economistes (vid. apt. 1.2.2.), els quals afegeixen que no existeixen

---

<sup>42</sup> L'obra de J. Harsanyi, Rational Behavior and Bargaining Equilibrium Selection in Games and Social Situations, (Cambridge University Press, Cambridge, 1977) és probablement l'intent més desenvolupat per exposar coherentment aquest conjunt de relacions d'una forma explícita.

<sup>43</sup> Com assenyala recentment el propi H. Raiffa, un dels més destacats especialistes durant els anys cinquantes i seixantes, *"The theory of games focuses its attention on problems where the protagonist in a dispute are superrational, where the 'rules of the game' are so well understood by the 'players' that each can think about what the others are thinking about what he is thinking, ad infinitum"* (a The Art and Science of Negotiation, Belknap Press, Cambridge, 1982, pàg. 2).

<sup>44</sup> M. Granovetter, "Economic Action and...", art. cit., 1985.

<sup>45</sup> R. Aumann, "What is Game Theory...", art. cit., 1985, pàg. 36.

altres alternatives millors<sup>46</sup>. Seguint la lògica d'aquesta perspectiva, interessada en identificar solucions d'equilibri únic per tot tipus de jocs, els matemàtics especialitzats en la teoria de jocs han desenvolupat poderosos algorismes de càlcul, els quals difícilment són accessibles als actors de les interaccions del món real (ja que les solucions s'han d'imputar als hipotètics càlculs racionals dels jugadors). Si bé aquesta perspectiva presenta greus dificultats per ésser aplicada a l'anàlisi dels processos socials, presenta un gran interès en molts camps on es fa ús de la racionalitat des d'un punt de vista normatiu. En aquest sentit, cal mencionar les possibilitats d'aplicació com eina de clarificació conceptual en camps com l'ètica i la filosofia política<sup>47</sup>, o inclús l'estratègia militar<sup>48</sup>, entre d'altres.

II. Una segona perspectiva accepta la modificació d'aspectes significatius del model de decisió racional o la incorporació d'altres mecanismes de decisió, mantenint l'objectiu d'explicar els resultats agregats del comportament a partir exclusivament de la teoria de jocs. Apareixen molt diverses possibilitats de realitzar modificacions a la hipòtesi de racionalitat, segons quins siguin els axiomes que es critiquin i les propostes alternatives que es formulin (vid. apt. 1.2.2. i 1.3.3.). Cal esmentar dues de les principals dificultats d'aquesta perspectiva. La primera, és l'absència de criteris per escollir una hipòtesi de racionalitat que serveixi per explicar el comportament observat, entre les diverses combinacions d'axiomes possibles. D'altra banda, tampoc pot ser una hipòtesi necessitada sempre d'una definició completament externa, experimental (vid. apt. 1.1.4.). La segona, és el

---

<sup>46</sup> Evidentment, seguit el judici d'Elster, aquest és l'autèntic argument de pes, però que no pot ésser exposat pels seus defensors degut a què només és manté gràcies a la debilitat de les opcions rivals: *"Neoclassical theory will be dethroned if and when satisficing theory and psychology join forces to produce a simple and robust explanation of aspiration levels, or sociological theory comes up with a simple and robust theory of the relation between social norms and instrumental rationality. Until this happens, the continued dominance of neoclassical theory is ensured by the fact that one can't beat something with nothing"* (a "Introduction", J. Elster, (ed.), Rational...op. cit., 1986, pàgs. 26-27).

<sup>47</sup> T. Schelling, "Some Thoughts on the Relevance of Game Theory to the Analysis of Ethical Systems", Journal of Conflict Resolution, vol. 12, 1968. Vegeu també A. Domènech, De la Ètica a...op. cit., 1989; i F. Ovejero, "Diez años de marxismo analítico", Arbor, octubre 1989.

<sup>48</sup> D.G. Haywood, "Military Decision and Game Theory", Operations Research, vol. 2, 1954.

dubte de si tot i aconseguint l'aplicació d'una teoria de la decisió adient, serà possible la reconstrucció de la teoria social i econòmica en termes d'accions i interdependències d'individus explicades exclusivament amb models de teoria de jocs.

És interessant destacar l'ampliació que A. Ryll realitza d'aquesta perspectiva, en entendre la teoria de jocs com un instrument útil per a la construcció de teories socials, gràcies a la provisió d'un concepte d'equilibri (solució que permet la predicció) en les estructures d'interdependència que representen el comportament dels individus<sup>49</sup>. Com un cas extrem, aquest autor accepta la possibilitat d'abandonar qualsevol formulació sobre la decisió racional a la teoria de jocs, i proposa el seguiment de normes socials com un mecanisme alternatiu de decisió. Formalment, la noció d'equilibri no cal que estigui sempre vinculada al supòsit de racionalitat individual. Es pot produir un equilibri a través d'estratègies estables, entenent que aquestes són normes dictades prèviament. Portant a la fi el seu argument, Ryll arriba a la conclusió següent:

*"Die Spieltheorie stellt ein besonderes Instrument der Gesellschaftsforschung dar. Sie ist nicht als Theorie menschlichen Verhaltens zu betrachten, sondern sie ist ein analytisches Werkzeug zur Formulierung einer logischen Isomorphie, die zur Präzisierung verschiedener (präskriptiver und deskriptiver) Theorien mit ganz unterschiedlichen Erklärungsobjekten dient."*<sup>50</sup>

Les afirmacions d'A. Ryll tenen algunes analogies amb els arguments emprats per

---

<sup>49</sup> Aquesta és la perspectiva d'A. Ryll, la qual destaca a més -amb l'exemple de l'aplicació de la teoria de jocs a la biologia- la independència de la teoria de jocs de la noció de racionalitat: "*Die Verwendung der Spieltheorie zur Modellierung von Evolutionsvorgängen zeigt die Möglichkeit, interaktive Verhaltensweisen ohne Bezugnahme auf etwaiges Rationalverhalten zu analysieren*" (a "*Die Spieltheorie als instrument der Gesellschaftsforschung*", Max-Planck-Institut für Gesellschaftsforschung, Discussion Paper 89/10, Colonia, 1989, pàg. 56).

<sup>50</sup> A. Ryll, "Die Spieltheorie...", art. cit., 1989, pàgs. 10-11.

a defensar el concepte de la preferència rellevada, que pretén separar les eleccions de la noció d'utilitat. En aquest sentit, tot i compartir la interpretació de Ryll sobre la independència entre la teoria de jocs i la teoria de l'«elecció racional», es considera que no es pot evitar la dependència de la teoria de jocs de l'establiment d'alguna hipòtesi sobre el comportament individual. Això es degut a que en les aplicacions de la teoria de jocs és necessari mantenir alguna interpretació sobre com es realitzen les decisions. Per moltes aplicacions a les ciències socials, s'entén que les decisions deriven de les preferències -encara que no exclusivament-, per la qual cosa sovint és necessari formular alguna teoria sobre com es generen les preferències. En aquest sentit, és aquesta teoria que es formula allò que determina en bona part les característiques dels equilibris que es poden assolir<sup>51</sup>.

III. Una tercera perspectiva consisteix en mantenir l'estructura bàsica de la teoria de jocs, que cerca explicacions basades en l'acció racional òptima -amb possibles modificacions de les orientacions d'interès-, però acceptant que per l'explicació de fenòmens col·lectius cal introduir complementàriament altres conceptualitzacions, externes a la teoria de jocs. Per construir teoria social s'argumenta que hom ha d'incloure elements fàctics com les restriccions a l'acció, els nivells d'informació i de comunicació, les regles institucionals, etc., així com la seva justificació en termes teòrics. No és possible explicar la formació i el funcionament d'aquest elements fàctics solament en termes dels requisits que la pròpia teoria imposa per facilitar la determinació dels equilibris<sup>52</sup>. Des d'aquesta perspectiva, no tots els resultats observats han de ser justificats necessàriament com a producte d'una acció racional,

---

<sup>51</sup> A. Ryll dona com exemple l'aplicació de la teoria de jocs a la biologia, però oblida que implícitament també hi ha un supòsit de comportament individual en l'anàlisi de la interacció, que en aquest cas és una clau genètica, cosa que pels individus podria ésser equivalent a una norma de comportament constant, almenys pel temps de l'anàlisi. Segons les comunitats d'origen les normes podrien ésser diferents, i produir-se llavors un procés de competició entre les diferents comunitats.

<sup>52</sup> Al marge de la capacitat analítica de la teoria de jocs, aquesta teoria també pot ajudar la investigació social sense necessitat d'establir una teoria social. Seria possible interpretar els resultats del càlcul interactiu com «tipus ideals» que fossin utilitzats per delimitar models descriptius; també podrien ésser aprofitables per a l'anàlisi de situacions històriques concretes, justificant externament els canvis de preferències (aquestes possibilitats reproduïxen els problemes esmentats a l'apt. 1.1.2.).

i poden ser entesos gràcies a factors que complementen els aspectes racionals del comportament (que són explicats amb la teoria de jocs). Tanmateix, com descriu M. Shubik referint-se a l'actual estat de la teoria de jocs, aquesta es pot entendre paradoxalment en el sentit que

*"one of the most important uses of the theory of games is in making precise the limitations on rational strategic behavior"*<sup>53</sup>.

Dins aquesta perspectiva, una interessant proposta de modelització de teoria de jocs amb la finalitat d'evadir el requisit d'informació completa és la que realitza F. Scharpf. Acceptant d'entrada el conjunt d'axiomes sobre racionalitat, aquest autor rebutja la noció d'estratègies mixtes per la seva escassa plausibilitat empírica<sup>54</sup>, i rebutja també -com possibilitat d'aplicació empírica- els recents refinaments sobre l'equilibri en jocs no-cooperatius, com els que aporta la recent teoria general de la selecció d'equilibri<sup>55</sup>, per la gran capacitat de càlcul que requereixen. Amb la resta de la teoria de jocs -especialment estratègies pures i equilibris no generalitzables- Scharpf creu que n'hi ha suficient per oferir força explicacions de fenòmens socials, si hom hi incorpora altres elements teòrics que expliquin l'orientació d'interès que segueixen les preferències (que defineixen la funció d'utilitat). L'acceptació que fa Scharpf de la teoria de jocs, tot la seva intenció de retenir els mínims elements possibles per a un treball empíric realment factible<sup>56</sup>, és defensada assenyalant que

---

<sup>53</sup> M. Shubik, "Comment", a K. Arrow i S. Honkapohja (Eds.), *Frontiers of...* op. cit., 1985, pàg. 97.

<sup>54</sup> F. Scharpf, "Games Real...", art. cit., 1989, pàg. 13. Scharpf es basa en els treballs experimentals de Colman (*Game Theory and Experimental...*, op. cit., 1982, pàgs. 77 i ss.) per argumentar que és absolutament infreqüent aquest tipus de comportament en les situacions reals d'interacció.

<sup>55</sup> J. Harsanyi i R. Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, 1988.

<sup>56</sup> Les seva estructura analítica es centre en què *"if profitable games are in fact played (and if Bayesian solution are empirically unlikely), actors must somehow be able to create or exploit conditions that allow them to interact as if they were playing with complete information (...) Such confidence-building, or empathy-creating, mechanisms could be either endogenous or exogenous to game-theoretical models"*. (Cf. F. Scharpf, "Games Real..", art. cit., 1989, pàg. 25).

pel moment no hi ha una altra alternativa plausible<sup>57</sup>. Per aquest motiu proposa aplicar les funcions d'utilitat relacionals, que permeten una gran varietat de fórmules per a caracteritzar les orientacions d'interès -definides externament a la teoria de jocs- (vid. apt. 2.2.2.).

IV. Finalment, com a quarta àrea de treball, hom pot utilitzar les possibilitats de transformació de la teoria de jocs, per a no seguir exclusivament la hipòtesi de racionalitat. La proposta és no acceptar aquestes possibilitats com a nou fonament exclusiu de la teoria social, sinó sols com un instrument explicatiu molt precís i útil. Al mateix temps, es tracta de no suposar sempre un propòsit d'optimització en el comportament dels individus, acceptant la incorporació d'altres elements explicatius sobre els processos de decisió i de formació de preferències. En un article de resposta al text d'Aumann mencionat anteriorment, R. Selten defensa la necessitat de desenvolupar una teoria descriptiva del comportament en jocs, i proposa la recerca d'alternatives al concepte de maximització per a aplicar a la teoria de jocs descriptiva (destacant la possibilitat d'aplicar conceptes derivats de la racionalitat limitada per desenvolupar una teoria de l'aprenentatge en jocs), perquè considera que

*"In view of the fact that human beings do not seem to maximize utility, one is tempted to jump to the conclusion that game theory has not relevance whatsoever for the description of economic reality. Of course, this conclusion would apply not only to game theory, but to most of microeconomics theory as well"*<sup>58</sup>.

---

<sup>57</sup> F. Scharpf (a "Games Real.", art. cit., 1989, pàg. 7) accepta l'estatus de la racionalitat en els models que proposa, arguentant la manca de generalitat dels models alternatius. Per aquest autor, els problemes d'aplicació empírica de la teoria de jocs es troben en els supòsits d'informació i en el nombre d'individus autònoms que hi participen. Scharpf segueix després amb l'estratègia d'aïllar alguns àmbits, acotar-los, i aplicar-hi altres assumpcions, per deixar als supòsits de racionalitat la «gestió» d'unes preferències ja donades. Precisament aquest és un exemple de la plasticitat instrumental d'aquestes teories de l'acció racional: dins una lògica pròpia de coherències i incoherències, es poden traspasar constructes i línies argumentals amb sentits o intencionalitats inicialment diferents per construir-ne d'altres.

<sup>58</sup> R. Selten, "Comment", a K. Arrow i S. Honkapohja (Eds.), Frontiers of... cit., 1985, pàg. 82.

Aquesta perspectiva, aplicada a la teoria de jocs, és força treballada i discutida actualment, defensant-se -com fa M. Shubik- la necessitat d'incorporar "*satisfactory concept of behavior with limited rationality*"<sup>59</sup> dins l'estructura de la teoria de jocs, encara que presenta greus problemes a l'hora de definir una nova noció de racionalitat (vid. apt. 1.1.4.). Seguint a T. Schelling, cal afegir que pot ser força absurda, donat l'estat actual de la teoria social, una gran sofisticació en els nous models de teoria de jocs desenvolupats per a l'anàlisi social:

*"For social scientist, what is rudimentary and conceptual about game theory will be, for a long time, the most valuable"*<sup>60</sup>.

Sintèticament, a les pàgines anteriors s'ha reproduït la tensió teòrica plantejada al primer capítol, derivada de la necessitat de simplicitat que tenen, d'una banda, la hipòtesi de la racionalitat, generalment vinculada a una noció d'elecció individual òptima; i d'altra banda, els intents, poc articulats fins al moment, de bastir i utilitzar models d'home amb conceptualitzacions més flexibles del comportament. Aquesta situació porta a una doble estratègia d'investigació en el context de la teoria de jocs. Per un costat, es discuteix la proposta de desenvolupar models d'interacció que no mantenen completament les conegudes propietats de la racionalitat (transitivitat, coherència, consistència) ja discutides al capítol anterior. D'altre costat, s'analitzen diversos intents per aplicar propostes basades en la racionalitat limitada, que rebutgen o integren en diferents graus el principi d'optimització. Finalment, cal destacar que s'utilitzen bàsicament dos criteris per

---

<sup>59</sup> M. Shubik, "Comment", art. cit., 1985, pàg. 98. En un altre context, M. Shubik afegeix que "*The immediate task of formal game theorists, social psychologists, political and military analysts is to begin to close the gap between the context free theories of games and risk behavior and the context specific problems of strategy, diplomacy, threat and war. This involves the construct of games where cultural bias, limited perception of agents, errors in systems and overloads play an identifiable operational role*" ("Games with Perceptive Commanders but with Indoctrinated or Less Perceptive Subordinates", a A. Diekmann i P. Mitter, (Eds.) *Paradoxical Effects of Social Behavior*, Physica Verlag, Heildelberg, 1986, pàg. 211).

<sup>60</sup> T. Schelling, "What is Game Theory?", art. cit., 1984, pàg. 241.

analitzar aquestes propostes: la plausibilitat dels supòsits de partida, i l'efectivitat que assolixen els mecanismes explicatius en cada cas. Utilitzar una teoria formal de la interacció basada en una anàlisi de l'equilibri interdependent, porta a discutir les conseqüències que per a la consistència interna de la teoria implica l'adopció de supòsits de comportament més plausibles descriptivament. La intencionalitat d'aquesta estratègia d'investigació es analitzar com es pot treure més benefici explicatiu -pel conjunt de la teoria social que es formuli- de la part de la teoria de jocs que es mantingui aprofitable.

Una de les qüestions que s'analitzen en aquesta investigació és com modificar les assumpcions de la teoria de jocs sobre l'entorn explicatiu que dona sentit a l'anàlisi de la interacció individual. Rebutjant les posicions que incorporen restriccions a l'optimització, s'obre un ampli camp de possibilitats i nivells de vinculació entre entorns de domini de normes, estructures físiques i biològiques, i formes de decisió humanes. Per oferir un exemple d'aquest problema es pot destacar el cas de l'anàlisi input-output en economia (una matriu que descriu el conjunt d'interdependències entre els diferents sectors econòmics en un moment donat), on s'ha d'afegir a més algun tipus de teoria macroeconòmica a la interpretació de la matriu d'interdependències per explicar com funciona l'economia. Normalment, es connecta amb teoremes keynesians sobre el creixement, que tenen unes determinades assumpcions sobre el comportament de decisió individual, però també seria factible vincular les matrius amb altres enfocaments<sup>61</sup>. Un altre exemple fa referència a la teoria de la reproducció sraffiana i les seves assumpcions sobre racionalitat com una variable més del model. Segons quin sigui l'estat que prengui la variable (encara que es suposa generalment que els empresaris maximitzen els seus beneficis) es poden produir diversos resultats en un model donat basat en les condicions tècniques, les

---

<sup>61</sup> Un tractament d'aquests problemes es pot trobar per exemple a W. Leontief, "Input-Output Data Base for Analysis of Technological Change", Economic Systems Research, vol. 1, 1989.



característiques dels béns, així com l'adaptació humana a aquests<sup>62</sup>. L'interès d'aquesta perspectiva en relació a la racionalitat és que el model no entra en crisi quan el comportament dels actors és clarament no-racional; al contrari, el model ens mostra els diferents resultats possibles.

Després d'aquesta aproximació general, hom pot afirmar que per la pròpia natura instrumental de la teoria de jocs, poden coexistir perspectives d'aplicació distintes. Això és degut a que aquesta teoria té una àmplia aplicabilitat, sense cap perversitat intrínseca, però amb la necessitat de ser utilitzada mitjançant adaptacions específiques, degut a la conceptualització que indispensablement ha de tenir associada, en estar buida de contingut fàctic<sup>63</sup>. En aquest sentit, es pot observar que enfocaments teòrics diferents poden fer ús de la mateixa metodologia, de la mateixa manera que poden existir discussions dins un mateix enfocament teòric sobre quina metodologia utilitzar i per quina finalitat<sup>64</sup>. El que cal recordar és que *dins la teoria de jocs es poden produir diverses línies de treball, alhora que diversos enfocaments teòrics substantius poden compartir una mateixa línia de treball. L'únic que produeix la teoria de jocs com instrumental és una certa unificació de procediments i mecanismes similars, que sense ser idèntics, presenten força trets anàlegs, perquè el seu significat el prenen de les teories que «representen» en cada cas* (per exemple, les nocions d'equilibri seran iguals formalment, però diferents significativa i conceptualment, per referir-se a xarxes de conceptes diferents). Per tant, es pot

---

<sup>62</sup> A. Sen, "Economic Methodology...", art. cit., 1989; F. Ovejero, La químera fèrtil. La posibilidad de una teoría de la historia, Ms., Barcelona, 1989; P. Sraffa, Producció de mercaderies per mitjà de mercaderies, edicions 62, Barcelona, 1985.

<sup>63</sup> En aquest sentit, M. Bunge, destaca que la teoria de jocs -de forma aïllada- és un joc numèric en el qual l'estimació dels valors inicials és totalment subjectiva: "Our main objection to game theory is that it is a pseudomathematical for involving utility functions that are nowhere mathematically defined (...) By conveniently altering the entries in the payoff matrix we can get to 'prove' almost anything". Per aquest motiu -i d'altres derivats de la mateixa apreciació-, Bunge conclou, potser precipitadament, que la teoria de jocs no té cap mena d'interès per la ciència política (a "Game Theory is Not a Useful Tool for the Political Scientist", Epistemologia, vol. 12, 1989, pàg. 206; vid. la secció 2.3. per una rèplica a Bunge).

<sup>64</sup> Un exemple pot ser el debat sobre les relacions entre el marxisme analític i la teoria de la racionalitat, esmentat a l'apt. 1.1.3.

afirmar que la teoria de jocs admet la diversitat, sense integrar-la, fent-la manejable. Sols en aquest sentit, i no en un sentit explicatiu, es pot dir que la teoria de jocs, més o menys formalitzada, constitueix una «gramàtica» per a les teories de l'acció.

### 2.2.2. Evidència psicològica i modelització interactiva de les orientacions d'interès.

Aquest apartat es dedica a comentar alguns resultats existents en l'experimentació sobre teoria de jocs, especialment l'aparició d'actituds cooperatives i competitives en diferents estructures d'interacció. Per un costat, s'analitzen les dificultats experimentals per validar els resultats que ofereix la teoria de jocs. D'altre costat, es discuteixen les propostes elaborades per a introduir distintes orientacions d'interès, sense alterar els supòsits de racionalitat, així com les implicacions d'aquestes modificacions.

Des de fa força anys, existeixen treballs de psicòlegs socials, que emprant models de la teoria de jocs com a punt de referència, mostren la diversitat d'actituds existents en la interacció social, en contra de la univocitat del model d'home de la teoria econòmica, que generalment sols adopta l'actitud de 'maximització del propi interès'<sup>65</sup>. Sense entrar a destacar les dificultats de contrastació dels experiments dels psicòlegs socials<sup>66</sup>, cal distingir les especificitats de l'experimentació en teoria de jocs de les qüestions referents als problemes generals de la decisió racional, ja

---

<sup>65</sup> Una de les primeres aproximacions va ser la de M. Flood ("Some Experimental Games", Management Science, vol. 5, 1958), que ja comentava "The utility concept, as it enters into game theory, can be criticized on the basis that suitable operational measures often cannot be found in real life applications" (pàg. 6). Per exemple, demostrant que en moltes situacions, els «guanys relatius» (o diferències entre jugadors) són més importants que la maximització dels guanys individuals per a la realització d'una elecció (Cf. D. Messick i W. Thorngate, "Relative Maximization in Experimental Games", Journal of Experimental Social Psychology, vol. 3, 1967).

<sup>66</sup> M. Plon, "On the Meaning of the Notion of Conflict and its Study in Social Psychology", European Journal of Social Psychology, vol. 4, 1974.

exposats en el capítol anterior. En aquest sentit, els jocs de suma variable són més coneguts i utilitzats en els experiments, essent emprats com analogies de les situacions socials en les que la cooperació pot millorar el resultat, malgrat que les condicions de l'estructura d'interacció no permeten assegurar aquesta. És interessant d'observar en el món real quants significats diferents tenen nocions tan simples com «amenança», «agressió», «acord» o «conflicte». Això dóna una oportunitat a la teoria de jocs per aplicar el seu aparell conceptual a precisar moltes qüestions importants de la teoria social.

Tant els psicòlegs socials des de fa força temps, com els economistes darrerament, s'han dedicat a realitzar experiments partint de models formals de la teoria de jocs<sup>67</sup>. Pels psicòlegs, la teoria de jocs els proveïa d'un instrumental analític força precís per analitzar els processos d'interacció entre individus. Més que en les conseqüències col·lectives agregades, aquests estaven interessats en comprendre la part que la racionalitat ocupa en la interacció entre grups de persones, assumint o no d'entrada el supòsit de maximització de l'interès propi<sup>68</sup>. D'altra banda, els economistes (i en menor mesura altres científics socials), estaven interessats en predir amb certesa comportaments agregats, a partir de models generalitzables de la teoria de jocs. Especialment durant la dècada dels anys vuitantes, s'han desenvolupat per part dels economistes experimentacions en tres àmbits de la teoria de jocs: la teoria de la negociació, els refinaments de l'equilibri de Nash en diverses etapes, i els jocs de coordinació d'una sola etapa. D'entre aquests, només es discuteix el segon àmbit, pel seu interès en relació a

---

<sup>67</sup> Una primera revisió es pot trobar a A. Rapoport i C. Orwant, "Experimental Games: A Review", Behavioral Science, vol. 7, 1962. Una visió actual del tema es pot trobar a A. Colman, Game Theory and Experimental... op. cit., 1982.

<sup>68</sup> Per exemple, es pot destacar l'afirmació de J. Rabbie, J. Schot i L. Visser, "*it is assumed that among competing action tendencies and available strategies, actions and alternatives will be chosen that promises, with a high probability of success to attain the most valued goals or profitables outcomes, whereby the gains clearly exceed the costs of achieving them*", (a "Social Identity Theory: A Conceptual and Empirical Critique from the Perspective of a Behavioral Interacción Model", European Journal of Social Psychology, vol 19, 1989, pàg. 177).

l'argumentació d'aquest treball.

Els resultats mostren que els subjectes actuen de vegades d'acord amb els arguments racionals de sofisticats equilibris arribant a les situacions previstes, però que no realitzen de forma conscient els refinats càlculs que suposa la teoria<sup>69</sup>. Moltes altres vegades els subjectes actuen de forma inconsistent amb la teoria, especialment quan s'aplica el teorema de Bayes<sup>70</sup>, o quan existeixen diverses possibilitats d'equilibri i no s'adopta el punt d'equilibri que es un òptim paretian<sup>71</sup>. Donades les dificultats per contrastar empíricament els resultats de la teoria de jocs formalment sofisticada, especialment quan es considera el supòsit que els jugadors adopten distribucions de probabilitats bayesianes, caldria adoptar supòsits de comportament sota incertesa més versemblants, tot i l'interès de preservar aquestes formulacions per les aplicacions normatives i prescriptives de la teoria de jocs<sup>72</sup>.

Deixant a banda la qüestió de l'aplicació de probabilitats (vid. apt. 2.2.5.),

---

<sup>69</sup> J. Brandts, "Economía experimental y teoría de juegos: un panorama", Cuadernos Económicos de ICE, vol. 40, 1988. Aquesta qüestió remet a la discussió general sobre la manca de plausibilitat en què es mou l'aplicació de la hipòtesi de racionalitat a la teoria econòmica (vid. apt. 1.1.3.).

<sup>70</sup> El teorema de Bayes s'empra en els jocs amb informació incompleta, o en diverses versions dels equilibris amb estratègies mixtes. C. Camerer i K. Weigelt, (a "Experimental Test of a Sequential Equilibrium Reputation Model", Econometrica, vol. 56, 1988) descobreixen la freqüent aplicació de conceptes de probabilitat no bayesianes, 'naives' en moltes situacions d'interacció. Vegeu la polèmica entorn l'adopció d'una visió subjectiva de la probabilitat per aplicar a la teoria de jocs descriptiva, que enllaça amb les propostes de Savage sobre la utilitat esperada (cf. J. Kadane i P. Larkey, "Subjective Probability and the Theory of Games", Management Science, vol. 28, 1982; la réplica d'Harsanyi i la contrarréplica dels autors en el mateix volum). J. Wilson (a "Subjective Probability and the Prisoner's Dilemma", Management Science, vol. 32, 1986) destaquen el problema de determinar les probabilitats subjectives, si hom abandona la regla bayesiana. El problema es pot vincular amb els arguments d'Allais i Machina sobre la necessitat de construir un índex més complex de la formació de preferències (vid. apt. 1.3.2.)

<sup>71</sup> R. Cooper, D. Dejong, R. Forsythe i T. Ross, "Selection Criteria in Coordination Games: Some Experimental Results", American Economic Review, vol. 80, 1990. Els autors analitzen jocs amb múltiples equilibris de Nash, equilibris que poden ésser ordenats en termes paretians, i descobreixen que tot i que els jugadors es situen en un dels possibles equilibris, no és normalment el pareto-dominant el que adopten, tot i considerar l'existència de jugadors altruistes.

<sup>72</sup> Sobre aquesta qüestió, vegeu J. Kadane i P. Larkey, "The Confusion of Is and Ought in Game Theoretic Contexts", Management Science, vol. 29, 1983 així com la resposta de Shubik al mateix volum.

una proposta interessant d'ampliar la capacitat explicativa de les funcions d'utilitat que s'empren en teoria de jocs, es realitza abandonant el supòsit unívoc de maximització de l'interès propi. Schulz i May<sup>73</sup> proposen un sistema de preferències per modelitzar un *continuum* d'interessos (absolut o en referència als altres jugadors), trencant així les dicotomies simples de tipus «egoista-altruista», «amic-enemic», etc. Mitjançant la creació d'escala d'orientacions socials, la seva proposta és fàcil d'operacionalitzar i força flexible, però es manté al camp de la teoria de la utilitat, emprant els supòsits estàndard de la racionalitat i el principi de maximització<sup>74</sup>. Un problema d'aquesta perspectiva és que només recull parcialment la interpretació de funcions d'utilitat «solidàries», perquè no postula com ha d'ésser l'estructura interna del grup per a possibilitar la satisfacció «solidària»<sup>75</sup>. Tradicionalment, la teoria de jocs pressuposa que els actors només tenen interès en la seva pròpia utilitat (com a derivació del concepte predominant en l'economia) i, per tant, que el seu «pagament» -quan es produeix el resultat del joc- és la utilitat de l'actor (encara que es derivi d'una funció altruista). No obstant això, aquesta perspectiva individualista, és un dels factors, entre d'altres, que complica els problemes d'aplicació empírica a la teoria de jocs. La qüestió és que en els jocs experimentals moltes investigacions -malgrat els pressupòsits de partida- han mostrat

---

<sup>73</sup> U. Schulz i T. May, "The Recording of Social Orientations..", art. cit., 1989.

<sup>74</sup> F. Scharpf, adopta aquest constructe per analitzar les decisions en les administracions públiques (a "Decision Rules, Decision Styles and Policy Choice", *Journal of Theoretical Politics*, vol. 1, 1989), acceptant la lògica de la teoria de jocs definida per Harsanyi (*Rational Behavior..op.cit.*, 1977, pàg. 116): els jugadors han de preferir els resultats que maximitzen els seus propis pagaments. Però contràriament als matemàtics de la teoria de jocs, Scharpf no accepta que els jugadors coneixen les estratègies i els pagaments dels altres, proposant suplir la necessitat d'informació completa per normes i identificació dels «estils» dels actors o situacions mitjançant la modelització d'orientacions socials d'interès (F. Scharpf, "Games Real...", art. cit., 1989, pàg. 25).

<sup>75</sup> K. Uhlander critica les propostes que intenten construir funcions d'utilitat no individuals ni centrades en l'interès propi, destacant que "*In practice, almost existing rational choice models of action (..) depend upon a decision maker's optimization of individually possessible goods*" ("Relational Goods' and Participation: Incorporating Sociability into a Theory of Rational Action", *Public Choice*, vol. 62, 1989). Uhlander proposa el concepte de «béns relacionals» que només poden ésser posseïts per mutu acord d'un conjunt d'individus, i que només augmenta la utilitat d'una persona si són consumits també per altres persones del conjunt determinat. Un exemple molt clar d'aquests tipus de béns són les llengües. Com més gent parli 'anglès' -per destacar-ne un-, més útil és saber aquest idioma.

que sovint els actors tenen en compte els «pagaments» dels altres en valorar la pròpia utilitat<sup>76</sup>.

Han estat principalment els psicòlegs socials que analitzen models de comportament interactiu els que, per les seves necessitats d'investigació empírica, han desenvolupat una estructura analítica de les interaccions entre individus a partir dels models formals que ofereix la teoria de jocs<sup>77</sup>. Un aspecte descobert és que no és possible inferir explícitament com un actor considera el «pagament» dels altres per la seva utilitat analitzant les seves pautes de decisió en simples jocs experimentals, sinó que cal buscar mètodes indirectes, com pot ser l'observació de les estructures de preferències en jocs ja coneguts -per evitar els problemes de la teoria de la preferència rellevada-<sup>78</sup>. Examinant casos on la recerca individual de l'interès propi no permet explicar les accions que es realitzen, apareixen fenòmens com les competicions esportives, on moltes vegades la victòria o derrota pròpia depèn especialment dels factors que controlen els altres jugadors, com la cohesió d'equip. En aquestes situacions, els membres tenen el seu interès principal en el resultat agregat, i l'interès propi depèn d'aquest. L'anàlisi de fenòmens com l'esmentat ha

---

<sup>76</sup> Una exposició es pot trobar a H. Kelley i J. Thibaut, Interpersonal Relations. A Theory of Interdependence, Wiley, Nova York, 1978.

<sup>77</sup> Vid. J. Kahan i A. Rapoport, "Social Psychology and the Theory of Games: A Mixed-Motive Relationship", Representative Research in Social Psychology, vol. 14, 1984.

<sup>78</sup> Això va ésser apuntat per D. Messick i C. McClintock ("Motivational Bases of Choice in Experimental Games", Journal of Experimental Social Psychology, vol. 4, 1968). A partir d'aquí, es varen proposar diverses formes d'analitzar aquest fenomen. Per un costat proposant jocs descomposats (per exemple, W. Liebrand, "The Effect of Social Motives, Communication and Group Size on Behavior in an N-Person Multi-Stage Mixed-Motive Game", European Journal of Social Psychology, vol. 14, 1984), per altre, desenvolupant mètodes de mesura independents de jocs concrets (per exemple, K. MacCrimmon i D. Messick, "A Framework for Social Motives", Behavioral Science, vol. 21, 1976). La possibilitat d'establir una connexió explicativa entre les orientacions socials i el comportament en jocs, iniciada per W. Hesse i U. Schulz (a "Preference Structures by Payoffs in Experimental Games", a R. W. Scholz (Ed.), Current Issues in West German Decision Research, Peter Lang, Frankfurt, 1986), apareix quan es poden interpretar les preferències dels jugadors, suposant que incorporin les seves orientacions socials. Experimentalment, la forma d'analitzar les preferències és a través dels «pagaments», pressuposant que a cada una de les distintes estructures del joc, s'han de produir diverses ordenacions de preferències sobre els «pagaments», segons les orientacions que adoptin els jugadors (i comprovant llavors si s'acompleixen).

portat a desenvolupar un conjunt de funcions d'utilitat distintes a la de la maximització de l'interès propi, per maximitzar altres orientacions d'interessos, com l'agressió, l'altruisme, la cooperació, etc., que permeten un treball experimental més exhaustiu sobre les formes de comportament. Formalment, es pot definir l'existència d'una funció lineal genèrica, que pren diversos valors segons les orientacions socials. Si tenim una funció d'utilitat composta de satisfacció pròpia (x) i satisfacció dels altres (y), es pot formular la següent funció general (amb 'u' utilitat, i 'a' i 'b' paramètres) :

$$u(x,y) = ax + by$$

A partir d'aquesta funció, les actituds individuals poden prendre valors específics en la funció. Com a punts ideals destacats, es poden definir diverses formes simples, seguint la proposta de Schulz i May<sup>79</sup>:

Andividualisme:	$u(x,y) = x$	Sacrifici:	$u(x,y) = -x$
Altruisme:	$u(x,y) = y$	Agressió:	$u(x,y) = -y$
Cooperació:	$u(x,y) = x + y$	Competició:	$u(x,y) = x - y$

També es poden definir formules més complexes, com les relacions condicionals:

Maximin:

$$u(x,y) = x + y - |x-y| \text{ (x quan } x < y, \text{ i } y \text{ quan } x > y)$$

Igualitarisme:

$$u(x,y) = - |x-y| \text{ (+x-y quan } x < y, \text{ i } -x+y \text{ quan } x > y)$$

---

<sup>79</sup> "The Recording of Social Orientations..", art. cit., 1989, pàgs. 42-43. Els autors es basen parcialment en el treball de K. MacCrimmon i D. Messick, "A Framework..", art. cit., 1976.

D'altra banda, per intentar una aproximació més realista, hi ha la possibilitat d'introduir una funció no-lineal, que prendria la forma general de:

$$u(x,y) = ax + by + c|x - y|$$

Un tema en discussió és l'estabilitat de preferències quan els agents interactuen. En aquest sentit, s'ha argumentat que la persistència d'unes determinades funcions pot estar reforçada per normes socials, genèriques o adaptades a contextos i situacions específiques. D'igual forma, la transformació d'aquestes funcions pot ser deguda a estar fonamentada solament per emocions i creences esporàdiques, que aconduïxen a unes orientacions molt volàtils, afectades substancialment per cada moviment de la interacció. En l'aplicació de les funcions a la teoria de jocs, es considera l'existència d'una «matriu donada» de pagaments objectivament definits, i d'una «matriu efectiva» que determina de fet les eleccions estratègiques dels actors: les preferències sobre pagaments que subjectivament consideren els actors, generades per una determinada orientació d'interès<sup>80</sup>. Els actors interpreten els «pagaments» (estableixen les seves preferències) a la llum de les relacions que creuen tenir envers els altres jugadors, més que en relació a unes formes prèvies donades externament (com es proposa amb la maximització de propi

<sup>80</sup> H. Kelley i J. Thibaut, *Interpersonal Relations...*, cit., 1978, pàgs. 14-17. Per exemple, en un joc del dilema del presoner, amb

		Jugador 2	
		Y1	Y2
Jugador 1	X1	a: (2, 2)	b: (4, 1)
	X2	c: (1, 4)	d: (3, 3)

L'ordenació tradicional de preferències pel jugador 1 seria  $b > d > a > c$ , arribant al resultat conjunt d'a: (2,2). Sobre aquest payoff «objectiu», basat en una orientació individualista, l'ordenació de preferències per un jugador 1 competitiu seria  $b > d = a > c$ ; i per un jugador 1 agressiu,  $b > a > d > c$ , etc.. (per una caracterització completa, vid. W. Hesse i U. Schulz, "Preference Structures...", art. cit., 1986, pàgs. 217-221). Un enfoc formal més general -i més difícil d'aplicar- sobre la qüestió de les confusions en els jocs, com a producte de la manca d'informació, es pot trobar a M. Wang, K. Hipel i N. Fraser, "Modeling Misperceptions in Games", *Behavioral Science*, vol. 33, 1988.



interès). Aquestes interpretacions subjectives poden ésser capturades per un limitat nombre de regles de transformació (per exemple, maximitzar/ minimitzar les diferències, els guanys conjunts,...)<sup>81</sup>, que permeten una precisa definició operacional de les actituds i tendències de comportament associades amb diferents entorns d'interacció social.

En poder canviar l'estructura de preferències dels jocs, que convencionalment era definida per una específica matriu de resultats, es trenca el supòsit de relació unívoca entre les preferències i els «pagaments» o resultats. Aquesta relació era mantinguda gràcies a suposar la relació de consistència i l'existència d'una sola orientació d'interès -la maximització dels propis guanys-. Així queda oberta la possibilitat que diversos jocs de difícil solució amb la regla de maximització, admetin solucions estables si es consideren altres estructures de preferències. Suposar que els jugadors realitzen interpretacions subjectives sobre les orientacions dels altres jugadors, junt a l'anàlisi de l'interacció entre jugadors amb diferents orientacions deixa oberta la possibilitat de realitzar una enriquidora anàlisi d'alguns sorprenents resultats de la interacció social, sense necessitat de desenvolupar sofisticats mecanismes de càlcul per arribar a similars conclusions. D'altra banda, es podria argumentar que disminueix la capacitat de predicció teòrica, i que s'inclou el supòsit d'utilitat transferible (que tots els individus tenen la mateixa escala per valorar satisfaccions); però tot i acceptant hipotèticament les crítiques, degut a les interessants possibilitats d'aplicació, les transformacions esmentades tenen grans possibilitats a les anàlisis interpretatives basades en l'enfocament de l'«elecció racional».

---

<sup>81</sup> Com destaca F. Scharpf, *"For empirical social scientists this implies the hope that we might finally be able to profit from the analytical power of the rational-choice apparatus without having to ignore most of what we know about the variance and contingency of human motives"*, a "Political Institutions, Decision Styles and Policy Choices", a R. Czada i A. Windhoff-Héritier (Eds.), Political Choice. Institutions, Rules and the Limits of Rationality, Campus Verlag, Frankfurt, 1991, pàg. 67).

### 2.2.3. Normes i jocs: variables exògenes o preferències endògenes.

L'amplitud del tema, així com la creixent discussió existent en el terreny de la teoria social, fa totalment inabordable en unes poques pàgines una presentació de l'estat actual de la qüestió. Tanmateix, el propòsit d'aquest apartat sols és mostrar algunes de les relacions existents entre la teoria de jocs i les normes, regles, pautes de comportament o obligacions, conceptes que, amb les seves matisacions pròpies, es refereixen a guies o indicacions col·lectives, més o menys compulsives, que limiten força estrictament la formulació de decisions<sup>82</sup>. El gran nombre d'equilibris possibles que apareix en moltes formulacions basades en la teoria de jocs, així com les dificultats per mantenir el supòsit d'informació completa en moltes situacions, ha portat a alguns investigadors a considerar que els jugadors tinguin en compte el seguiment de normes en les seves decisions. Així, estableixen el supòsit de que empàticament els jugadors creuen que les normes seran compartides pels altres jugadors, davant a l'alternativa de suposar la realització de sofisticats càlculs interactius per arribar a uns resultats similars.

Es reconeixen generalment tres grans mecanismes de formació de normes. Primer, l'establiment i difusió de normes a través d'institucions, o inclús individus, que tenen l'autoritat i la capacitat per regular aspectes del comportament d'altres individus. Segon, la formació de normes a través d'acords voluntaris entre individus que s'obliguen recíprocament al seu compliment, siguin aquestes dues persones (per exemple, un acord entre veïns) o tots els membres d'una nació (per exemple, a

---

<sup>82</sup> Seguint la formulació general de F.A. Hayek, "The term 'rule' is used for a statement by which a regularity of the conduct of individuals can be described, irrespective of whether such a rule is 'known' to the individuals in any other sense than they normally act in accordance with it" (cf. "Notes on the Evolution of Systems of Rules of Conduct", a Studies in Philosophy, Politics and Economics, Routledge & Kegan, Londres, 1967, pàg. 67). Donada la generalitat de l'enfocament adoptat, no es distingeix entre distints tipus de normes, per exemple normes socials versus normes morals, o legals, o d'altres, ja que s'entén que l'anàlisi efectuat es pot aplicar a totes. La distinció entre norma i regla, sotmesa o moltes possibles matisacions, pot ésser entesa, de forma general, per la major importància que en la norma tenen les sancions associades al seu incompliment. Com defineix J. Elster una norma, "is the propensity to feel shame and to anticipate sanctions by others at the thought of behaving in a certain, forbidden way" (a The Cement of..., op. cit., 1989, pàg. 105).

través d'un referèndum). Finalment, en tercer lloc, hi ha la formació evolutiva de normes. En aquest cas les obligacions per acomplir no han estat explícitament planejades o formulades, sinó que apareixen com a producte de la interacció social, i poden estar sotmeses a transformacions derivades del propi canvi social<sup>83</sup>. La tipologia exposada suggereix l'aparició de dos dilemes, que es poden vincular amb les distintes perspectives amb què s'interpreta la teoria de jocs (apt. 2.2.1.). El primer dilema fa referència a l'aparició d'aquestes pautes de conducta -normes-. Apareixen dues possibilitats per fonamentar l'explicació. La primera, des del propi enfocament de l'«elecció racional», amb un model d'home i una lògica explicativa basada en la teoria de la racionalitat i l'acció individual optimitzadora<sup>84</sup>. La segona, seguint les tradicions dels enfocaments sociològics o antropològics, i utilitzant teories emergents no basades en la racionalitat<sup>85</sup>. Com a derivació del primer, el segon dilema tracta de la forma de vinculació de les normes amb els models de la teoria de jocs. Aquesta vinculació pot ésser endògena o exògena a la pròpia teoria. En el

---

<sup>83</sup> Per a una caracterització d'aquesta classificació, veure K.-D. Opp, "The Evolutionary Emergence of Norms", *British Journal of Social Psychology*, vol. 21, 1989. Seguint la proposta de Ph. Mirowski, una distinció alternativa, més general, seria l'existent entre regles «naturals» (definides per les condicions tècniques, ambientals, o fins i tot cognitives), les regles «socials» (construïdes intencionalment -els dos primers casos de la distinció anterior-. Un tercer tipus de regles serien les «evolutives» -com el tercer cas anterior-, enteses com a producte de la relació entre les regles «naturals» i les «socials» (Cf. Ph. Mirowski, "Institutions as Solution Concepts in a Game Theory Context", a *The Reconstruction of Economic Theory*, Kluwer, Boston, 1986).

<sup>84</sup> Per a cada mecanisme de formació de normes es poden trobar distintes aproximacions explicatives partint de l'enfocament de l'elecció racional. La teoria de la decisió i la teoria de les votacions, pel cas de les regles institucionals (per exemple, J. Brennan i J. Buchanan, *La razón de las normas*, Norma Editorial, Madrid, 1986; també P. Ordershook, *Game Theory and...*, *op. cit.*, 1986, cap. 6). La formació de normes voluntàries es tractada per la teoria de la negociació (J. Elster, *The Cement of...*, *op. cit.*, cap. 2, 1986; també A. Roth, (ed.), *Game-Theoretic Models of Bargaining*, Cambridge University Press, Cambridge, 1985); i finalment, per a l'aparició de normes emergents, hi ha les teories sobre la cooperació i l'aparició de convencions en jocs repetits, vid. E. Ullmann-Marglit, *The Emergence of Norms*, Clarendon Press, Oxford, 1977.

<sup>85</sup> J. Elster, per exemple, després de criticar l'explicació generalitzada de les normes basada en el supòsit de racionalitat i de maximització de l'interès propi (vid. "Norms of Revenge", *Ethics*, vol. 100, 1990), tot i defensant la necessitat de formular una teoria de les normes socials en termes individualístics, assenyala que la pregunta de perquè unes determinades normes existeixen en llocs i temps específics, continua essent una incògnita per a les ciències socials, indicant que "one should adopt an inductive strategy. Careful examination of the diversity of norms in different societies may provide tentative generalizations, which in turn may suggest possible causal mechanisms" ("Rationality and Social Norms", Paper prepared for the XII Congress of the ISA, Madrid, 1990, pàgs. 20-21).

primer cas, cal que els jugadors estableixin preferències a través de la interacció sobre l'autodependència de pautes d'acció limitades o prèviament acordades<sup>86</sup>. La vinculació exògena entén que les normes són derivacions d'influències -o restriccions- de l'entorn social preexistent sobre els jocs i jugadors, que afecten directament la funció d'elecció dels individus immersos en aquest entorn, sigui l'anàlisi de la generació de normes formulat en termes d'elecció racional o no<sup>87</sup>. De les diverses combinacions que es poden produir a partir dels dilemes presentats, es discuteixen dues qüestions: i) les dificultats que presenta la teoria de jocs convencional per integrar les normes exògenes; i ii) els intents d'analitzar l'aparició de les normes emergents, des de la teoria de l'elecció racional.

En aquestes pàgines es defensa que la introducció de normes socials de forma exògena a la teoria de jocs no és quelcom que devaluï l'anàlisi de la interacció. Al contrari, es considera que pot aportar més precisió a aquesta anàlisi, en reduir l'àmbit d'explicació que es pretén per la teoria de jocs. D'altre costat, es pot pensar que és factible en casos concrets l'explicació de regles socials des d'un punt de partida individual, com a propietat emergent (igual que les ideologies o els símbols com a sistemes de creences). Tantmateix, aquesta tasca encara està mancada de

---

<sup>86</sup> Si són normes explícites i acordades voluntàriament, les teories de la negociació aplicades tradicionalment en la teoria de jocs no està clar que siguin realment endògenes per la teoria de jocs, sinó que s'aproximen més a constructes per trobar equilibris no-cooperatius. Aquests parteixen de la construcció d'un pre-joc de negociació, en què els agents exposen les seves verdaderes intencions, i es revelen els pagaments dels contraris, imposant unes condicions al pre-joc per assegurar la credibilitat de les comunicacions i acords. El mecanisme, deixant a banda el seu escàs realisme, implica també suposar uns elevats costos informacionals als agents (té el seu origen en la proposta de J. Nash, "Two Person Cooperative Games", art. cit., 1953; vid. apt. 2.1.2.). Malgrat això, en els darrers anys s'ha intentat afrontar aquest problema introduint el supòsit d'informació incompleta i la formulació de models amb equilibris seqüencials (A. Roth i M. Malouf, "Game-Theoretic Models and the Role of Information in Bargaining", *Psychological Review*, vol. 86, 1979; R. Myerson, "Negotiation in Games: A Theoretical Overview", a L. Hurwicz, D. Schmeidler i H. Sonnenschein, (Eds.) *Social Goals and Social Organization*, Cambridge University Press, Cambridge, 1985; i J. Sutton, "Non-Cooperative Bargaining Theory: An Introduction", *Review of Economic Studies*, vol. 53, 1986). Així, apareixen importants problemes en aquestes noves formulacions, com per exemple els derivats de la utilització de jugadors bayesians, encara que sí aconsegueixen el propòsit d'oferir una explicació endògena.

<sup>87</sup> Una distinció similar sobre les normes socials, però aplicada en una discussió sobre el conjunt de la macroeconomia, es pot trobar a A. Field, "Microeconomics, Norms and Rationality", a *Economic Development and Cultural Change*, vol. 32, 1984.

propostes suficientment convincents, essent possible -a més- que sorgeixin límits conceptuals<sup>88</sup>. En canvi, pot ésser més interessant buscar explicacions més complexes del funcionament de les normes en termes causals, que permetin establir finalment connexions més acurades en els contextos analítics en què les normes s'apliquen com a variables exògenes.

I. Cal esmentar que la perspectiva de vinculació exògena està obligada, per les característiques actuals de la teoria de jocs, a substituir directament la funció d'elecció per una norma<sup>89</sup>. Això és degut a què els models habituals de la teoria de jocs només integren una orientació d'interès (la maximització del propi interès), es produeix una situació en que l'estructura de preferències es pot inferir de la matriu de resultats, suposant la consistència de les eleccions. Les intensitats relatives d'utilitat per cada jugador són una funció de les diferències entre resultats dels pagaments; però aquesta relació podria conduir a diverses ordenacions de preferències. L'acceptació de diverses orientacions d'interès en una població permet una altra forma -menys simplista- de que les regles s'incorporin en els models de la teoria de jocs mitjançant la seva influència sobre les ordenacions de preferències. Aquesta perspectiva implica el supòsit de que les regles influeixen directament les orientacions d'interès, la qual cosa no ha de ser efectiva en totes les ocasions. Quan només hi ha una orientació d'interès, desapareix la possibilitat de pensar en regles exògenes que influeixin les preferències, ja que preferències i regles es converteixen en dues cares de la mateixa moneda; i si les preferències assumeixen regles, els

---

<sup>88</sup> Referint-se a la capacitat de les tècniques de la teoria de jocs per explicar les estructures de les regles «socials», A. Field opina en el seu article que *"Insofar as the origins of language or the more fundamental constitutive rules of a group or society are concerned, the answer is negative"* (Cf. "Microeconomics, Norms and Rationality", art. cit., pàg. 691)

<sup>89</sup> Al contrari d'allò que es podria anomenar regles «externes», que són completament obertes, i que tracten de qui defineix i de qui abona els pagaments que es disputen els jugadors. Però aquestes regles no estan vinculades a la teoria de jocs, sinó que formen part d'un context més general en què es pot incorporar una explicació en termes de teoria de jocs, en el cas que es formulin explícitament. Com ja assenyalava K. Arrow l'any 1951, una possibilitat per introduir regles en el joc és a través del disseny de la distribució inicial de recursos (a "Mathematical Models in the Social Sciences", a D. Lerner i H. Lasswell, (Eds.), *The Policy Sciences*, Stanford University Press, Stanford, 1959<sup>o</sup>).

individus ja no segueixen cap ordenació de preferències, sinó que apliquen directament les regles<sup>90</sup>.

Aquesta situació aconduïx a la conclusió que la teoria de jocs que la incorporació del supòsit d'interès pròpi deixa poc espai a la seva combinació amb l'adopció de regles de comportament (també es podria assenyalar en aquest context que l'única regla admissible és la de maximitzar el propi interès). Les preferències que els agents conformen, abans i mitjançant la interacció, són tota la informació que s'aporta al joc, i defineixen les possibilitats d'acció que els agents consideren. Com s'ha esmentat, una altra possibilitat d'aplicar regles exògenes en els models de la teoria de jocs és suposar que aquestes afecten directament els mecanismes individuals de decisió. Així, en comptes d'emprar procediments racionals d'escollir estratègies a partir d'ordenacions de preferències, els agents aplicarien directament una regla per actuar, independentment de les seves preferències sobre els resultats de la seva acció<sup>91</sup>. No sembla possible altra cosa, excepte les limitacions prèvies que determinen els contextos físics, tant naturals, creats per l'home, o com a resultat de la pròpia interacció (per exemple, la dimensió en l'acció col·lectiva)<sup>92</sup>. Tot això

---

<sup>90</sup> Així ho destaca C. Mironescu: *"The solution is dictated by the very structure of the game; the game alone determines -through the rationality of individual decision-makers- and legitimises the balance of individual advantages"* (a "Individual and Collective Interest in Game Theory", a P. Birnbaum (ed.), *Democracy, Consensus and Social Contract*, Sage, Londres, 1978, pàg. 323).

<sup>91</sup> Un exemple interessant és el que proposa J. Hey sobre regles de recerca. Rebutjant el concepte de maximització-optimització, descriu unes regles de comportament observades experimentalment sobre la recerca d'informació en mercats. Com a tal, aquestes regles no són integrables en teoria de jocs, però externament es poden vincular amb un model d'interacció estratègica en què els agents maximitzen un resultat, mentre que la seva recerca d'informació depèn d'aquestes regles ( Cf. J. Hey, "Search for Rules for Search", *Journal of Economic Behavior and Organization*, vol. 3, 1982).

<sup>92</sup> Aquesta és també la tesi de Ph. Mirowski: *"the mathematical formalism of game theory is best suited for the discussion of natural rules"*, destacant que només serveix com descriptora de les regles intencionals, i que no serveix per les regles evolutives, ja que aquestes no són teleològiques (cf. "Institutions as Solutions Concepts...", art. cit., 1986, pàg. 260). Per la seva banda, C. Bicchieri (a "Strategic Behavior and Counterfactuals", *Synthese*, vol. 76, 1988) aplica una concepció similar per discutir la qüestió de la selecció entre múltiples equilibris. Bicchieri defensa que per aïllar un únic equilibri entre molts possibles, cal introduir en la concepció de la racionalitat de les creences una teoria de com els jugadors revisen les seves creences en les situacions que es desvien de l'equilibri: *"if the players have common knowledge of an initial model of the game and of the rules for belief revision, they can attain common knowledge of their mutual expectations"* (pàg. 165). Bicchieri proposa atribuir als jugadors una

no impedeix, que el coneixement previ sobre la situació d'interacció, explicitat per altres mitjans, es pugui integrar en els models de teoria de jocs a través del concepte de «regles del joc», si són usades per construir les matrius de pagaments (resultats), de forma independent de les preferències i decisions dels agents<sup>93</sup>.

II. En aquest punt, es planteja si a través de la interacció entre individus racionals pot ésser explicada l'emergència de normes i institucions, sense invalidar la possibilitat d'interpretar les normes com a construccions socials<sup>94</sup>, ni la possibilitat de vincular-les de forma endògena amb els models de la teoria de jocs. El que es tracta d'explicar és com els agents desenvolupen les seves expectatives (sobre les accions dels altres) de forma convergent per estructurar les seves decisions. En l'explicació de les normes emergents a través de l'enfocament de l'«elecció racional», per solucionar la manca de confiança entre els agents a l'hora d'establir normes, i degut a la impossibilitat dels jugadors per crear condicions de mútua credibilitat, les aproximacions més destacades es centren en els jocs repetits, on la continuïtat en la interacció pot fer aparèixer la confiança en la cooperació. El «Folk Theorem» demostra que en jocs infinits o indeterminats el resultat es

---

teoria de contrafactuals (o 'norma lògica'), perquè convergeixin en les seves creences; en aquest sentit, el seu enfoc forma part de la teoria de la racionalitat, però força més sofisticat. Una alternativa dins aquest esquema conceptual, però sortint de les suposicions sobre la lògica dels jugadors, és el «Personal Principle», o «Group Principle» que proposa M. Gilbert (a "Rationality, Coordination, and Convention", *Synthese*, vol. 84, 1990), entenent per aquest l'adopció deliberada d'un principi d'acció per part dels jugadors, front a la indeterminació de la interacció racional. Comportant, a més, una acceptació conjunta en la interacció -implícita o explícita- d'aquest principi d'acció, com a resolució dels problemes de coordinació d'una col·lectivitat: "a social convention is a jointy accepted fiat, a rule of the fiat form that is properly seen as 'ours' by the members of some population: a 'group principle'" (pàg. 18).

<sup>93</sup> Per exemple, en una situació en la qual hi han dos lladres presos per un robatori comú, en els interrogatoris per separat pot ser molt diferent la seva actitud segons si es deixa en llibertat al que traïx, o bé si només se li redueix lleument la pena. Caldria determinar amb més atenció quin tipus de regles són possibles d'incorporar a través d'aquest procediment. Segurament les regles socials bàsiques d'organització i distribució s'hi podrien introduir; però les «metaregles» (regles sobre regles, com pot ser una Constitució), o les regles sobre procediments cognitius i de representació de la realitat, no està tan clar.

<sup>94</sup> Com diu A. Field, "The task must be to show how phenomena previously explained as the result of nonrational behavior or the operation of such ad hoc devices as social norms in fact result from the rational interaction of freely choosing self-interested individuals unconstrained by such norms" (Cf. "Microeconomics, Norms and rationality", art. cit., 1984, pàg. 688).

transforma en cooperatiu en fer-se dependent l'elecció d'un jugador del seu resultat en el proper moviment<sup>95</sup>.

Sense sortir de l'enfocament que assumeix la hipòtesi d'un comportament racional per explicar l'aparició de normes, cal destacar el treball de R. Axelrod sobre un dilema del presoner iterat, que mostra l'aparició d'un equilibri cooperatiu estable quan els jugadors utilitzen una estratègia de cooperació condicional<sup>96</sup> (en aquest sentit, es podria entendre el resultat de la cooperació com l'acceptació col·lectiva d'una norma). K.-D. Opp defensa la idea més general de què tot comportament recurrent es converteix en una norma. Aquest tipus de comportament es produeix quan està associat a alts beneficis i baixos costos, sigui directament, sigui per processos d'imitació; i quan l'estructura social (per exemple, comunicació-cohesió) pot afavorir la formació de preferències sobre aquest comportament, que acabin constituint-se en normes socials<sup>97</sup>. Una aproximació en una línia similar és el treball de A. Schotter. Insistent en la no intencionalitat de l'aparició de regles socials -o regularitats de comportament-, assenyala que aquestes emergeixen del comportament maximitzador individual a través d'un joc repetitiu que porta a què

---

<sup>95</sup> Un desenvolupament recent del «Folk Theorem», en condicions més realistes -però encara poc verosímils-, es pot trobar a D. Fudenberg i E. Maskin, "Folk Theorems...", art. cit., 1986. Una altra línia d'investigació per explicar l'aparició de regles endògenes en un model de teoria de jocs amb els supòsits de racionalitat, és suposar l'existència de payoffs discontinus. Es defineix un espai d'estratègies sobre el qual es determina una solució, en existir indeterminació en els pagaments, els autors demostren que existeix una possible regla que aconduïx a l'equilibri (cf. L. Simon i W. Zame, "Discontinuous Games and Endogenous Sharing Rules", *Econometrica*, vol. 58, 1990).

<sup>96</sup> L'estratègia «ull per ull» va mostrar molt més èxit que altres estratègies molt més sofisticades en la competició organitzada per R. Axelrod (*La evolución... op. cit.*, 1986). R. Axelrod explica l'aparició de normes dominants en termes de teoria de jocs a partir de les estratègies evolucionàries estables, a "An Evolutionary Approach to Norms", *American Political Science Review*, vol. 80, 1986. El model d'Axelrod també ha estat aplicat a jocs amb informació incompleta, amb resultats diferents però també satisfactoris per l'estratègia «ull per ull» (cf. J. Bendor, "In Good Times and Bad: Reciprocity in an Uncertain World", *American Journal of Political Science*, vol. 31, 1987. (per a una ampliació, vid. apt. 4.3.2.)

<sup>97</sup> K.-D. Opp, "The Evolutionary...", art. cit., 1982.



cada jugador restringeixi voluntàriament el seu espai d'estratègies<sup>98</sup>.

El problema dels models que intenten explicar l'aparició de normes emergents a través de jocs iterats és que el seu àmbit d'aplicació és molt limitat, perquè moltes interaccions socials no són infinites o indeterminades, o bé no prenen la forma d'un dilema del presoner. A més, la seva connexió amb altres jocs es manté de forma exògena, sense formar part de les interaccions que disputen altres resultats substantius<sup>99</sup>. D'altra banda, l'argument d'Opp, sense utilitzar l'aparell analític de la teoria de jocs, té un dels seus supòsits fonamentals en la maximització del propi interès, cosa que el fa dependre de les crítiques a què ha estat sotmesa l'aplicació d'aquest supòsit com a principi invariant. El problema és que les regles, com les normes, o les institucions, són constructes socials: poden conduir a actuar més racionalment, però no necessàriament, ja que també poden portar al contrari, o bé modificar les pròpies preferències anteriors<sup>100</sup>.

---

<sup>98</sup> Schotter elabora el concepte d'aparició endògena d'un conjunt d'equilibris de regles, escollides inintencionalment pels actors dins un ampli conjunt de regles possibles. El procediment de selecció -en termes estàtics- és la recerca d'un equilibri a través de la maximització d'utilitat al llarg d'un comportament repetit. A. Schotter, "The Evolution of Rules", a R. Langlois, (Ed.), Economics as a Process, Cambridge University Press, Cambridge, 1986. Del mateix autor, vegeu també The Economic Theory of Social Institutions, Cambridge University Press, Cambridge, 1981.

<sup>99</sup> Per situar el problema d'aquesta línia d'investigació, és interessant la caracterització de Mirowski: *"Game theorists have opened the Pandora's Box marked 'rationality' and do not know how to close it again (...) game theorists proposed to discuss bargaining, which lead to cooperative games, which begat noncooperative games, which begat discussions of process, which allowed the transactors the freedom to differ in their interpretations of the roles of others and the constancy of the world, all of which is now undermining the older construct of mechanistic rationality. (...) It is happening because game theory exposes the weakness of the physical metaphor that all the excessive mathematical formalism served to obscure. Game theory does not, however, suggest what to put in its place. It cannot conceptualize the reduction of a language or of an institution to a game"* (Cf. P. Mirowski, "Institutions as Solution Concepts...", art. cit., 1986, pàg.257).

<sup>100</sup> La dificultat no és tant per explicar l'aparició d'algunes normes evolutives a través de la racionalitat i l'interès propi -per evitar represàlies, per les seves bones conseqüències, etc.-, cosa que és possible en alguns casos, sinó per proposar aquesta via com una explicació general de l'aparició de normes. Hi ha multitud de contraexemples, accions no racionals que restringeixen la recerca de l'interès propi- (J. Elster, "Rationality and Social Norms", cit., 1990). La polèmica per ampliar el nombre de casos explicables a través de l'«elecció racional» és creixent. Per exemple, Ph. Petit (a "Virtus Normativa: Rational Choice Perspectives", Ethics, vol. 100, 1990) proposa l'estratègia basada en l'actitud (davant la basada en el comportament), que tracta primer de l'aprovació d'un tipus donat de comportament, per explicar després el seguiment d'un comportament per part de l'agent (de forma similar a l'enfocament proposat per M.

S'han comentat diverses possibilitats d'integrar regles en els models de la teoria de jocs, però per finalitzar, cal insistir en els problemes de l'integració de les regles que provenen d'una formulació externa a l'enfocament de l'«elecció racional». En general, el que proveeixen els sistemes de regles (igual que els sistemes de coneixement) són informacions sobre les característiques dels jocs, que permeten als jugadors tenir un coneixement suficient dels requisits informacionals d'aquests sense els elevats costos d'obtenir informació en cada interacció. Així, s'entén que els sistemes de regles poden definir externament tan els valors dels pagaments com les ordenacions de preferències o els mecanismes de decisió que els jugadors adopten. Aquesta perspectiva no discrimina el tipus de regles que pot integrar (emergents o no) per vincular amb la teoria de jocs. Accepta tan sols que hi ha uns elements del procés d'elecció que poden ésser influenciats per les regles, ja que han d'ésser formalitzables a través de matrius de pagaments (que es suposen exògenament conegudes per tots els jugadors), les estratègies adoptades o l'ordenació de preferències. En aquest sentit, es podria considerar addicionalment l'existència de regles externes (generalment institucionals o voluntàries), conegudes per tots els actors, que constreñeixin el conjunt d'estratègies que aquests poden portar a terme - els resultats accessibles- o que defineixen quins jugadors poden portar a terme el joc<sup>101</sup>, servint el conjunt com un marc per formular l'anàlisi en termes de teoria de jocs<sup>102</sup>.

---

Gilbert).

<sup>101</sup> E. Ostrom, "An Agenda for the Study of Institutions", Public Choice, vol. 48, 1986.

<sup>102</sup> Arribant a una conclusió similar, A. Field destaca que "Given shared language and other fundamental rules, the techniques of game theory do provide a framework for understanding how, in the case of a limited set of regulative rules, a political process might select from a group of possible rules or norms those which are most efficient" (Cf. "Microeconomics, Norms and Rationality", art. cit., 1984, pàg. 704). En el mateix sentit apunta el propòsit de M. Shubik, "My basic approach to economics is through the construction of mathematical models in which the 'rules of the game' derive not only from the economics and technology of the situation, but from the sociological, political and legal structure as well" (a Game Theory in the Social Sciences, MIT Press, Cambridge, 1982, pàg. 10).

#### 2.2.4. Problemes en la modelització dels actors a la teoria de jocs.

En aquest apartat s'analitza si l'estructura formal de la teoria de joc comporta limitacions a l'hora de proposar modificacions en els axiomes que componen la hipòtesi de la racionalitat. En general, el tema ha estat poc treballat a la literatura especialitzada, existint únicament -al nostre coneixement- algunes propostes aïllades i amb poca continuïtat, sense que entre els dos camps d'investigació existeixin massa vinculacions fluïdes encara que, sorprenentment, les possibilitats són més àmplies d'allò que fora previsible. S'apunta l'argument que tot i existir interessants possibilitats de transformació de la hipòtesi de racionalitat aplicada a la teoria de jocs, el reduït coneixement existent actualment en ciència social fa difícil poder integrar molts dels desenvolupaments instrumentals que es deriven d'aquestes transformacions. El problema és que s'han d'introduir més afirmacions teòriques substantives sobre la formació de preferències, o bé desenvolupar les teories existents en àmbits socials ben definits. El motiu és que l'aplicació de variacions de la hipòtesi de racionalitat fora d'una d'aquestes dues línies, només s'aconsegueix ampliar el camp de possibilitats, augmentant la indeterminació interpretativa. Com ja s'ha esmentat anteriorment (apt. 2.1.2.), la independència de la teoria de jocs de la teoria de la utilitat comporta considerar -per definir solucions estables-, com elements mínims de la teoria, l'existència dels principis de decisió i de pautes per determinar l'equilibri<sup>103</sup>. Seguidament es discuteixen els dos casos «extrem» possibles, deixant per l'apartat següent alguns dels casos intermitjos. En primer lloc, es dedica especial atenció a les conseqüències que diverses alteracions dels axiomes bàsics de l'acció racional pressuposen per la formulació d'equilibris en la teoria de jocs. En segon lloc es discuteix l'aplicació de la teoria de jocs sense els conceptes d'utilitat i racionalitat, quan solament s'utilitzen distribucions de probabilitat per a

---

<sup>103</sup> Veure B. Peleg, "The Independence of Game Theory...", art. cit., 1966. Peleg demostra l'existència de solució d'equilibri en un model de joc sense transitivitat, continuïtat i funció d'utilitat. L'únic que pressuposa és l'existència d'una funció que assigna a cada estratègia una distribució de probabilitat sobre els resultats. Això permet a aquesta autor considerar la teoria de jocs clàssica com un cas especial contingut en la generalització.

justificar les decisions.

La teoria de jocs utilitza normalment la Hipòtesi de la Utilitat Esperada, que postula l'existència de funcions de preferències convexes, com a base per a la construcció de les ordenacions de preferències entre estratègies mixtes (vid. apts. 1.3.1. i 2.1.2.). La seva exclusió implicaria limitar molt el nombre d'equilibris possible. En aquest sentit, s'analitzen seguidament els efectes que algunes modificacions de la Hipòtesi poden tenir per la teoria de jocs, especialment en relació a la determinació d'equilibris. Sense formar part de la Hipòtesi, l'aplicació de la teoria de Bayes com a predictor del comportament de les persones en situacions d'incertesa és un important problema, constantment destacat en l'aplicació de la teoria de jocs a l'anàlisi social<sup>104</sup>. Aquesta teoria s'aplica també a l'anàlisi dels jocs amb informació incompleta, per transformar-los en jocs d'informació imperfecte. Desconegudes les probabilitats objectives, es suposa que el subjecte no té la capacitat d'establir probabilitats subjectives (entre altres possibles causes, per manca d'experiència en la interacció) i es considera que el jugador assumeix la teoria de Bayes en el seu tractament de la probabilitat (vid. apt. 2.2.2.). Però a més d'aquest problema, l'aplicació més general de la utilitat subjectiva porta discutir com és possible modificar els supòsits sobre l'actitud dels jugadors front el risc.

M. Kostreva mostra que en introduir el supòsit de no-convexitat<sup>105</sup> a la teoria de jocs no-cooperatius, apareix un conjunt de punts d'equilibri (que nomena 'pre-equilibrium set') el qual inclou tots els equilibris de Nash, i és més gran que

---

<sup>104</sup> Sobretot per ésser aplicada d'una forma generalitzada, universal (vegeu per exemple, D. Grether, "Bayes Rules as a Descriptive Model: The Representativeness Heuristic", *Quarterly Journal of Economics*, vol. 95, 1980).

<sup>105</sup> La convexitat en les funcions d'utilitat és un supòsit que implica l'aversion estable davant el risc en totes les situacions de decisió possibles, tant de pèrdues com de guanys. L'axioma d'independència, en imposar que tercers factors no afecten una decisió entre altres dos elements, garanteix aquest supòsit de convexitat continuat en evitar l'aparició de funcions no-lineals que implicarien una aversió al risc variable.

el conjunt que formen aquests<sup>106</sup>. Aquesta possibilitat permet considerar la integració en la modelització de la teoria de jocs de propostes com les de M. Machina o D. Kahneman i A. Tversky, que elaboren modificacions de l'axioma d'independència, suposant l'existència de funcions d'estimació individual de probabilitats no lineals (vid. apt. 1.3.3.). En aquest sentit, ja han estat proposats conceptes d'equilibri en teoria de jocs sense l'axioma d'independència<sup>107</sup>. Malgrat tot, el problema que apareix és l'augment de la complexitat per determinar punts d'equilibri únics, ja que augmenta el seu nombre, i per tant, les dificultats per seleccionar-ne un, tot i seguir els recents mètodes per a seleccionar equilibris únics (o quasi-únics)<sup>108</sup>. En l'aspecte normatiu, i pel que fa a la teoria dels jocs cooperatius, també s'han formulat algunes propostes que no inclouen l'axioma d'independència, essent la més destacada la d'E. Kalai i M. Smorodinsky<sup>109</sup>. La seva proposta també porta a un punt únic d'equilibri, substituint l'axioma d'independència per un axioma de monocitat, que imposa una funció monòtona lineal als jugadors, però que no determina una orientació concreta davant el risc<sup>110</sup>. No obstant això, es pot deduir que si hom surt de la linealitat torna a aparèixer el

---

<sup>106</sup> M.M. Kostreva, "Nonconvexity in Noncooperative Game Theory", a International Journal of Game Theory, vol. 18, 1989. Una aproximació similar es pot trobar a J.F. McClendon, "Existence of Solutions of Games with Some Non-Convexity", a International Journal of Game Theory, vol. 15, 1986. La introducció de la no-convexitat és formulada a través de funcions d'utilitat polinòmiques, mostrant que és possible calcular tots els punts d'equilibri amb estratègies pures. La possibilitat d'aplicar funcions polinòmiques en la teoria de jocs ja es va contemplar des dels inicis de la disciplina (vid. M. Dresher, S. Karlin i L.S. Shapley, "Polynomial Games", Annals of Mathematics Study (Contributions to Theory of Games, I), vol. 24, 1950).

<sup>107</sup> V. Crawford, "Equilibrium without Independence", Journal of Economic Theory, vol. 50, 1990. Crawford parteix de la proposta de M. Machina i elabora un concepte d'equilibri que anomena «Equilibrium in Beliefs», que és una generalització de l'equilibri de Nash, en el sentit que les creences d'un jugador no depenen de les creences de l'altre.

<sup>108</sup> Per exemple, de l'article seminal de M. Kreps i R. Wilson, "Sequential Equilibria", art. cit., 1982. Per les crítiques a aquest intent, vid. apt. 2.2.5.

<sup>109</sup> E. Kalai i M. Smorodinsky, "Other Solutions to Nash's Bargaining Problem", Econometrica, vol. 43, 1975.

<sup>110</sup> Per una aplicació dinàmica d'aquest model (amb preferències sobre el temps), conduint també a un únic equilibri, veure N. Anbarci, "The Kalai/Smorodinsky Solution with Time Preferences", Economic Letters, vol. 31, 1989. Una aplicació en la mateixa línia es pot trobar a M.J. Herrero, "The Nash Program: Non-Convex Bargaining Problems", Journal of Economic Theory, vol. 49, 1989.

problema d'un augment del nombre d'equilibris possibles.

Apart de l'actitud davant el risc, una altra possibilitat és transformar els propis axiomes de racionalitat, cosa que implica modificar alguns dels supòsits bàsics de les funcions d'utilitat. Com s'ha vist al capítol anterior, en el marc de la teoria de la decisió ja existeix una àmplia diversitat de propostes. Cal afegir que tot i no ser suficientment realista, es manté sempre la hipòtesi que els jugadors, per trobar equilibris, formulen estratègies mixtes i finalment escullen a l'atzar. Una possibilitat addicional per simplificar les aplicacions a les ciències socials seria no acceptar els equilibris derivats de les estratègies mixtes. Així no caldria discutir les actituds front el risc, tot i que seguiria essent fonamental la discussió sobre els axiomes de decisió racional i els seus efectes sobre la teoria de jocs.

Abans de discutir aquestes propostes, és interessant d'esmentar el recent teorema d'impossibilitat formulat per K. Basu, que implica precisament la necessitat d'investigar com es pot incorporar una formulació més dèbil dels axiomes de la racionalitat en la teoria de jocs:

*"The Theorem shows that a definition of rational behaviour which is applicable to all extensive games and which does not suffer from the problem of unreached nodes (...) does not exist"<sup>111</sup>.*

Basu considera els jugadors amb accions imprevisibles com irracionals, i creu que no cal buscar «racionalitzacions-explicacions» d'aquest comportament, amb la finalitat de mantenir el supòsit de racionalitat. Això porta a aquest autor a la conclusió que no tots els jocs han de tenir solucions, o que la racionalitat sigui

---

<sup>111</sup> K. Basu, "On the Non-Existence of a Rationality Definition for Extensive Games", a International Journal of Game Theory, vol. 19, 1990, pàgs. 33-34. Per «unreached nodes» Basu es refereix al conegut problema de la cooperació en iteracions, en què, sense capacitat de preveure el nombre de jocs a realitzar, si cada jugador actua «racionalment» (seguint els teoremes d'equilibri), realitzarà la defecció en totes les jugades; encara que intuïtivament, es pot afirmar que en molts casos, i donades les condicions esmentades, es produiria la cooperació.

sempre definible en entorns estratègics. En aquestes situacions, Basu defensa que l'alternativa és introduir els «errors» dins el supòsit de comportament racional dels jugadors: o sigui, modificar, i per tant introduir major complexitat en la conceptualització de la lògica de l'acció dels subjectes<sup>112</sup>. En aquest marc es poden plantejar preguntes sobre quins tipus d'equilibris es mantenen quan es suposa, per exemple, que els jugadors no realitzen eleccions transitives, o quan apareix la incompletesa, o quins nous supòsits cal introduir, perquè no desapareixi la possibilitat de trobar equilibris.

Deixant a banda la no-linealitat de les funcions d'utilitat, resten dos camins per a intentar fer més complexos els supòsits de la racionalitat. Aquests parteixen de la discussió sobre els axiomes d'ordre-completesa i de continuïtat. Ambdós, tot i ésser molt criticats empíricament (vid. apt. 1.3.3.), constitueixen principis fonamentals de la teoria de la decisió. Així, cal veure si el seu rebuig o modificació implicaria renunciar a l'elaboració de models de la teoria de jocs amb una certa capacitat predictiva, restant només la possibilitat de construir models ad-hoc amb una alta variabilitat. La primera possibilitat sorgeix a través del concepte de transitivitat. L'estratègia seria treballar amb funcions d'utilitat sense aquest requisit<sup>113</sup>, com ja s'ha formulat per la teoria de l'equilibri del consumidor<sup>114</sup>. Aquest camp està relativament poc treballat, especialment pel que fa a la seva aplicació en teoria de

---

<sup>112</sup> K. Basu, "On the Non-Existence..." art.cit., 1990, pàgs. 41-43.

<sup>113</sup> Com proposa R. Aumann, "Utility Theory without the Completeness Axiom", art. cit., 1962. Aumann demostra que és possible mantenir una teoria de jocs amb resultats parcialment ordenats (no totalment comparables), calculant de forma anàloga als jocs convencionals els diversos equilibris possibles (pàgs. 447-8). Per altra banda, Aumann destaca que "el preu a pagar" és que desapareix la unicitat de la funció d'utilitat ( quan  $U(x) > U(y)$ , no vol dir que  $x P y$ , encara que sí a l'inrevés).

<sup>114</sup> Vid. H. Sonnenschein, "Demand Theory without Transitive Preferences, with Applications to the Theory of Competitive Equilibrium", a J. Chipman, L. Hurwicz, M. Richter i H. Sonnenschein, (Eds.), Preferences, Utility and Demand, Harcourt Brace Jovanovich, Nova York, 1971; i A. Mas Colell, "An Equilibrium Existence Theorem for a General Model without Complete or Transitive Preferences", Journal of Mathematical Economics, vol. 1, 1974.

jocs, existint algunes propostes aïllades<sup>115</sup>. Hom podria suggerir que rebutjant o no la completesa, la línia d'investigació seria caracteritzar estructures més complexes, multifuncionals, que a més de tractar amb les preferències no transitives, fossin menys dependents d'altres supòsits, com la convexitat<sup>116</sup>. D'altra banda, introduir funcions no-lineals a la teoria de jocs, pot servir per evitar una repetició del problema de la recursió de les preferències en cicles (entre estratègies i funcions de preferències), però apareix la dificultat d'incorporar proposicions teòriques sobre el funcionament de les relacions entre les diverses funcions de preferències, així com analitzar posteriorment les conseqüències per a la formulació d'equilibris interactius. Com a dificultat addicional, cal mencionar que generalment els equilibris han de dependre de les funcions de preferències que es determinin; i que com més discontinües siguin aquestes, major serà la multiplicitat d'equilibris possibles a assolir racionalment.

L'altra possibilitat de modificar la definició axiomàtica de la racionalitat té com a punt de partida discutir el supòsit de continuïtat. La qüestió es pot plantejar a través del tema de les funcions d'utilitat no-arquímedes, que tenen una relació de preferència o indiferència al mateix temps, la qual cosa trenca la propietat de la continuïtat. Aquestes propostes tenen un gran interès per la possibilitat de la seva aplicació a la teoria de jocs en situacions de no aplicabilitat de la racionalitat estàndard, com podria ésser l'existència d'estructures de preferències múltiples, no

---

<sup>115</sup> Apart de l'article de R. Aumann, cal destacar el treball de J.C. Cesco i A. Neme, "A Nash Type Equilibrium in Games without Transitive Preferences", Mathematical Social Sciences, vol. 17, 1989. Els autors destaquen que "the price paid for weaker conditions on the preferences of the players is the necessity of assuming stronger convexity conditions both on the preferences and on payoff functions" (pàg. 48). No obstant això, en aquest article es suposa l'existència de preferències completes i reflexives, en forma de correspondències que mantenen la propietat de la continuïtat.

<sup>116</sup> Per exemple, seguint les propostes de P. Fishburn (P. Fishburn, "Nontransitive Measurable Utility", Journal of Mathematical Psychology, vol. 26, 1982; i P. Fishburn i I. LaValle, "The Structure of SSB Utilities for Decision under Uncertainty", Mathematical Social Sciences, vol. 15, 1988), es proveeix una discriminació bilineal de preferències que no assumeix la transitivitat (però sí ordre, continuïtat i convexitat). A partir d'aquesta teoria, Y. Nakamura (a "Bilinear Utility and a Threshold Structure for Nontransitive Preferences", Mathematical Social Sciences, vol. 19, 1990) realitza una proposta que relaxa l'axioma de convexitat, definint una estructura de lliendar a la qual apareix una discriminabilitat imprecisa de les preferències i una relació asimètrica en la bilinealitat.



reduïbles a una sola funció, ja que en la utilització d'escala univalors apareix la inconsistència. Von Neumann i Morgenstern ja van reconèixer en la seva obra la possibilitat d'utilitzar utilitats no-arquímedes<sup>117</sup>, i amb posterioritat altres autors també van insistir en la mateixa possibilitat<sup>118</sup>. A començaments dels anys setantes, un article de P. Fishburn rebutjava la possibilitat d'emprar aquestes utilitats no-estàndards perquè els jocs resultants no tenien punt d'equilibri<sup>119</sup>; però un article posterior de H. Skala replantejava en altres termes el problema, demostrant la possibilitat de trobar punts d'equilibri amb un altre model d'utilitat ordinal<sup>120</sup>. La discussió és interessant, però ha tingut poca continuïtat, a més que presenta una elevada complexitat matemàtica, en treballar amb conceptes d'ordre no convencionals. Tanmateix, el seu desenvolupament per aplicacions empíriques potser podria ésser bastant més simple, si es desenvolupessin models concrets d'interacció amb multidimensionalitat de les funcions de preferències.

En tot cas, les diverses propostes de transformació de la funció d'utilitat i la seva vinculació amb els fonaments de la teoria de jocs estan molt poc desenvolupades. A partir dels resultats disponibles actualment, tan sols es pot apuntar que existeix la possibilitat de fer algunes modificacions axiomàtiques, mantenint-se la capacitat de trobar punts d'equilibri a la teoria de jocs. En certs casos, com quan es formulen funcions no lineals, augmenta força el nombre

---

<sup>117</sup> Theory of Games and..., op. cit., 1953<sup>3</sup>, pàgs. 29 i 630-631.

<sup>118</sup> Per exemple, Luce i Raiffa, Games and..., op. cit., 1957, pàg. 27. Un dels primers desenvolupaments va ser la proposta de M. Hausner per a debilitar els axiomes de Von Neumann-Morgenstern sobre la utilitat esperada ("Multidimensional Utilities", a R. Thrall et al., Decision Processes, Wiley, Nova York, 1954).

<sup>119</sup> P. Fishburn, "On the Foundations of Game Theory: The Case of Non-Archimedean Utilities", International Journal of Game Theory, vol. 1, 1972.

<sup>120</sup> "The state of affairs with non-Archimedean utilities is not so different from the results under von Neumann-Morgenstern utilities as Fishburn asserts, if we represent the utilities in an appropriate non-Archimedean ordered field (nonstandard model of the real numbers) and admit that the components of the optimal strategies also may assume values in this ordered field", H.J. Skala, "Nonstandard Utilities and the Foundation of Game Theory", International Journal of Game Theory, vol. 3, 1974, pàg. 67). Del mateix autor, sobre la teoria de la utilitat i els ordres de preferències no-estàndards, vegeu Non-Archimedean Utility Theory, Reidel, Dordrecht, 1975.

d'equilibris possible. Malgrat això, en els resultats existents apareix una certa interdependència entre la linealitat de les funcions i la possibilitat d'alterar els supòsits de transitivitat o continuïtat, presentant-se força difícil la possibilitat d'alterar simultàniament ambdós supòsits, si l'objectiu final és mantenir una teoria de la decisió basada solament en un concepte de racionalitat interactiva. Resten dos aspectes no discutits, que caldria afegir a les complexitats d'aquestes propostes en el cas de voler desenvolupar aplicacions teòriques a temes substantius. Primer, el problema dels requisits informacionals i els seus costos pels jocs d'informació completa, que no sembla que des d'aquest punt de vista es pugui resoldre. Segon, l'assumpció d'una constant no discutida, associada a la concepció de com s'assoleix l'equilibri, que és el principi de comportament optimitzador (vid. apt. 2.2.5.).

L'aplicació de la teoria de jocs a problemes de la biologia planteja sorprenentment un repte a l'aplicació de la teoria de jocs en les ciències socials, perquè en la biologia no s'utilitza el supòsit del comportament racional, i de retruc, es podria inferir l'aplicació de la teoria de jocs a les ciències socials sense el requisit de suposar el comportament racional dels agents<sup>121</sup>. A la biologia, els pagaments s'entenen com l'èxit relatiu de reproducció, i l'elecció d'estratègies és representada per l'adopció de diferents genotips en fraccions d'una mateixa població. La interacció dels elements d'una població en estructures mixtes de conflicte-cooperació, porta generalment al predomini d'una de les fraccions, que tendeix a adoptar una estratègia evolutiva estable (en el sentit que la població no pot ésser modificada per mutants), reproduint el procés de selecció natural<sup>122</sup>. Com destaca J. Maynard Smith,

---

<sup>121</sup> A. Ryll, "Die Spieltheorie...", art. cit., 1989. L'article de B. Peleg ("The Independence of Game Theory...", art. cit., 1966) defineix una aproximació formal a aquesta possibilitat.

<sup>122</sup> Vid. A. Rapoport, "Applications of Game-Theoretic Concepts in Biology", Bulletin of Mathematical Biology, vol. 47, 1985; així com J. Maynard Smith, "Evolutionary Game Theory", a D. Farner et al., Evolution, Games and Learning, North-Holland, Amsterdam, 1986; i I. Bomze i B. Pötscher, Game Theoretical Foundations of Ecological Stability, Springer Verlag, Berlin-Heidelberg, 1989.

*"Paradoxically, it has turned out that game theory is more readily applied to biology than to the field of economic behaviour for which it was originally designed. There are two reasons for this. First, the theory requires that the values of different outcomes (..) be measured on a single scale. In human applications, this measure is provided by 'utility' -a somewhat artificial and uncomfortable concept: in biology, Darwinian fitness provides a natural and genuinely one-dimensional scale. Secondly, and more important, in seeking the solution of a game, the concept of human rationality is replaced by that of evolutionary stability. The advantage here is that there are good theoretical reasons to expect populations to evolve to stable sets, whereas there are grounds for doubting whether human beings always behave rationally"*<sup>123</sup>

No cal recalcar, que els problemes de conceptualitzar la interacció humana són més complexes, però aquesta aplicació de la teoria de jocs a la biologia pot ajudar precisament a clarificar diversos d'aquests aspectes problemàtics. L'interès d'aquesta aplicació és que mostra l'aparició de cooperació sense pressuposar la racionalitat individual, ni maximització de l'interès individual, ni l'existència d'acords previs. L'únic que és pressuposa es l'adopció d'una estratègia -un genotipus- davant altres possibles i, com a resultat de la interacció, una determinada taxa de reproducció de l'espècie. Alguns autors infereixen una analogia «quasi-racional» del procés de selecció natural com a generador de decisions<sup>124</sup>. Aquesta analogia es basa en l'observació que la interacció segueix una pauta similar a la d'un equilibri de Nash, i la suposició que aquest equilibri implica la hipòtesi de

---

<sup>123</sup> J. Maynard Smith, Evolution and the Theory of Games, Cambridge University Press, Cambridge, 1982, pàg. vii.

<sup>124</sup> Per exemple, P. Hammerstein, "The Biological Counterpart to Non-Cooperative Game Theory", Working Paper 131, Institute of Mathematical Economics, Universitat de Bielefeld, 1984.

racionalitat individual<sup>125</sup>. En general però, el principal problema d'aquest enfocament és el desconeixement dels genotips sobre els que pressuposa comportaments<sup>126</sup>. Per tant, la dificultat que conseqüentment apareix en l'anàlisi és com discriminar entre les diverses estratègies evolutives estables possibles la que és més probable que estigui actuant a la natura<sup>127</sup>.

Traslladat en termes d'interacció social, una estructura similar de generar decisions podria suportar un principi (o regla) de comportament individual (per exemple, l'imperatiu categòric Kantià: *actuar com es vol que els altres actuïn, perquè s'universalitzi aquesta actitud*), conduint a un resultat conjunt distint d'aquell que implicaria actuar per maximitzar el benefici individual<sup>128</sup>. En cert sentit, aquest podria ésser un comportament intransitiu per un agent que en altres circumstàncies actuï amb uns principis de discussió diferents, que caldria integrar conjuntament en els models de la teoria de jocs. Però el problema bàsic de les estratègies evolutives estables és que considera l'existència de «jugadors miops», amb una complexitat interna molt baixa; quan hom vol que s'introdueixin mecanismes de decisió reflexius, no pot estalviar els problemes de la definició de la hipòtesi de comportament racional dins la teoria de jocs. Tanmateix, la idea de les estratègies

---

<sup>125</sup> L'argument, en relació a la teoria econòmica, és desenvolupat per L. Johansen, a "On the Status of the Nash Type of Noncooperative Equilibrium in Economic Theory", Scandinavian Journal of Economics, vol. 84, 1982. Johansen resumeix el seu argument en la següent forma, "Some conditions are stipulated for rational decisions, and it is demonstrated that, provided that the equilibrium is unique, decisions in accordance with the nash noncooperative equilibrium concept satisfy the requirements, and they are the only decisions which do so" (pàg. 439).

<sup>126</sup> El problema és similar a quan en ciències socials s'intenten deduir les preferències a través de les accions realitzades. Tot i que pot aparèixer la possibilitat d'existir més d'un ordre de preferència darrera una acció.

<sup>127</sup> Vid. G.A. Parker i P. Hammerstein, "Game Theory and Animal Behaviour", a P.J. Greenwood, P.H. Harvey i M. Slatkin, Evolution. Essays in Honour of John Maynard Smith, Cambridge University Press, Cambridge, 1985.

<sup>128</sup> Com suggereix A. Rapoport, "Prisoner's Dilemma", The New Palgrave. A Dictionary of Economics, vol. 3, 1987. En paraules de Kant, el principi s'expressa com "yo no debo obrar nunca más que de modo que pueda querer que mi máxima deba convertirse en ley universal" (M. Kant, Fundamentación de la metafísica de las costumbres, Espasa-Calpe, Madrid, 1981, traducció de M. García Morente).

evolutives estables pot ésser força aprofitable, especialment si s'aconsegueixen integrar conjuntament amb un concepte de decisió racional<sup>129</sup>.

### 2.2.5. Racionalitat limitada, aprenentatge i recerca de l'equilibri.

Entre la concepció de bastir una teoria de jocs independent de consideracions externes a la racionalitat dels agents, i la concepció d'una teoria de jocs que exclou dels seus postulats la racionalitat dels agents, apareix un ampli i divers conjunt de possibilitats que s'agrupa entorn de conceptualitzacions com la racionalitat limitada o els processos d'aprenentatge. Precisament per ésser casos intermedis, les propostes existents segueixen línies d'investigació força diverses i, en certs casos, inclús contradictòries. En general, totes tenen en comú l'intent d'articular en un mateix model regles i racionalitat per a la interpretació dels processos d'interacció estratègica.

La introducció d'aquestes línies d'investigació en el debat recent sobre la teoria de jocs es pot entendre a partir d'una sorprenent paradoxa sobre els supòsits de la racionalitat en la interacció estratègica. La paradoxa consisteix en què experimentalment -també intuïtivament- en un joc repetit de suma variable, que tingui un nombre determinat de jugades, del tipus dilema del presoner, es pot observar freqüentment que al llarg de la interacció apareix la cooperació entre els jugadors. Això no obstant, seguint la hipòtesi del comportament racional, no seria possible trobar cap equilibri cooperatiu, perquè hauria de dominar sempre la defecció. Aquest raonament considera que els jugadors descompten des de l'última interacció prevista fins la primera l'avantatge particular de la defecció, resultant

---

<sup>129</sup> El tema és tractat al següent apartat. Vegeu també K. Binmore, Essays on the Foundations of Game Theory, Basil Blackwell, Cambridge, 1990.

finalment com un joc no repetiu<sup>130</sup>.

Els nombrosos intents de cercar nocions d'equilibri més refinades per a explicar les pautes cooperatives que apareixen en aquests casos han tingut un paper important en les investigacions dels darrers anys<sup>131</sup>. Permetent que un jugador dubti de la racionalitat de l'altre -formalment considerant l'existència d'informació incompleta sobre el pagament de l'altre jugador-, s'intenta demostrar l'aparició de cooperació en un dilema del presoner finit<sup>132</sup>. Malgrat això, el problema d'aquesta proposta és que cal considerar la hipòtesi -poc plausible- que els jugadors coneixen ja, des de l'inici del joc, les seves futures pautes de comportament.

Una altra línia d'investigació recent ha estat l'aplicació d'autòmats finits. Aquests es defineixen mitjançant models de jugador amb limitacions a la seva capacitat d'accedir a factors que influeixen en el procés de decisió. Es tracta d'analitzar com s'afavoreix la cooperació en una estructura del dilema de presoner

---

<sup>130</sup> R. Luce i H. Raiffa (*Games and...*, op. cit., 1957, pàgs. 97-102), varen ésser els primers en destacar aquesta paradoxa. Per a una revisió actualitzada del problema, vid. W. Raub, "An Analysis of the Finitely Repeated Prisoner's Dilemma", *European Journal of Political Economy*, vol. 4, 1988. Raub destaca que "the incentive structure of the finitely repeated PD is similar to the one-shot PD and different from the infinitely repeated PD, where cooperation in the constituent games can be generated by equilibria of the supergame" (pàg. 377). Una possibilitat d'evitar aquesta situació seria suposar que els jugadors saben que el joc és repetit, però desconeixen el nombre de repeticions del joc (com ho proposa K. Basu, "Modeling Finitely-Repeated Games with Uncertain Termination", *Economic Letters*, vol. 23, 1987). Però aquest només és un cas particular que, pels seus efectes subjectius sobre els jugadors, s'aproxima a un joc de repetició infinita (en el qual teòricament és possible la cooperació, com ho afirma el «Folk Theorem» -vid. apt. 2.1.2.-).

<sup>131</sup> Cal destacar especialment l'article de D. Kreps i R. Wilson, "Sequential Equilibria", art. cit., 1982. Els autors assenyalen la importància de les creences, integrant-les en la descripció del model. Considerant que aquestes són «sequencialment racionals» i consistents, introdueixen un mecanisme per restringir les creences que no tenen cap probabilitat de succeir. Posteriorment, tant els propis autors com altres investigadors, han elaborat refinaments d'aquesta proposta d'equilibri. Una noció oposada d'equilibri, que inclou l'equilibri de Nash com un cas particular és la proposada per B.D. Bernheim, a "Rationalizable Strategic Behavior", *Econometrica*, vol. 52, 1984. Bernheim considera que els axiomes de l'elecció racional normalment acceptats porten a un concepte més ampli d'equilibri i, per tant, que no es pot imposar l'equilibri de Nash basant-se solament en la racionalitat. Això implica la consideració que l'equilibri de Nash -per no parlar de versions més refinades d'aquest- incorpori restriccions sobre les expectatives dels agents entorn d'allò que pot -o no- ésser plausible i, per tant, que també és possible incorporar altres supòsits alternatius de comportament basats en distintes hipòtesis de plausibilitat.

<sup>132</sup> D. Kreps, P. Milgrom, J. Roberts i R. Wilson, "Rational Cooperation...", art. cit., 1982.

iterat quan es tenen en compte els factors de complexitat<sup>133</sup>. Entre les primeres propostes desenvolupades es poden destacar els treballs de Neyman i Rubinstein. El primer, limita la complexitat de les estratègies dels jugadors, perquè sigui menor que el nombre total de jugades de la iteració<sup>134</sup>; mentre que el segon introdueix a més una variable sobre el nivell de complexitat de l'autòmat<sup>135</sup>, per aplicar als jocs infinits i mostrar paradoxalment que en aquests casos també apareix la dificultat de la cooperació.

Posteriorment, altres propostes han desenvolupat l'aplicació d'autòmats finits a la teoria de jocs, introduint la incertesa sobre la configuració de l'autòmat del jugador contrari, i discutint la selecció de la complexitat estratègica de l'autòmat, tant per investigar les conseqüències en l'equilibri dels jocs infinits<sup>136</sup>, com per analitzar els problemes informacionals de la cooperació en els jocs finits<sup>137</sup>. La complexitat apareix degut a què el jugador no coneix el seu adversari i, per tant, ha de calcular una multiplicitat d'equilibris possibles, perquè no té una regla clara de selecció. Això implica que ha d'incorporar una teoria sobre l'aprenentatge de les

---

<sup>133</sup> Per a una visió general, vid. R. Marimon, "Introducción a los Juegos...", art. cit., 1988. La primera proposta en aquest sentit va ser realitzada per R. Aumann, a "Survey of Repeated Games", Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Bibliographisches Institut, Manheim-Viena-Zurich, 1981.

<sup>134</sup> A. Neyman, "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma", Economic Letters, vol. 19, 1985.

<sup>135</sup> A. Rubinstein, "Finite Automata Play the Repeated Prisoner's Dilemma", Journal of Economic Theory, vol. 39, 1986. Per a un desenvolupament de la noció de complexitat estratègica, vid. B. Lipman i S. Srivastava, "Informational Requirements and Strategic Complexity in Repeated Games", Games and Economic Behavior, vol. 2, 1990.

<sup>136</sup> Vid. D. Abreu i A. Rubinstein, "The Structure of Nash Equilibrium in Repeated Games with Finite Automata", Econometrica, vol. 56, 1988; i J. Banks i R. Sundaram, "Repeated Games, Finite Automata and Complexity", Games and Economic Behavior, vol. 2, 1990. També, per a una aplicació a un altre concepte d'equilibri, veure E. Kalai i W. Stanford, "Finite Rationality and Interpersonal Complexity in Finite Games", Econometrica, vol. 56, 1988.

<sup>137</sup> Per exemple, E. Zemel, "Small Talk and Cooperation: A Note on Bounded Rationality", Journal of Economic Theory, vol. 49, 1989. Zemel introdueix un mecanisme de comunicació, que enviant missatges sense sentit, impedeix que els jugadors s'impliquin en estratègies més complexes que potencialment podrien impedir la cooperació entre ells.

regles de selecció i la capacitat de comprovar experimental els resultats anteriors (és a dir, amb memòria)<sup>138</sup>. Un cas especial bastant estudiat és el de l'asimetria entre els jugadors, quan un jugador es considerat plenament racional i l'altre té algunes característiques d'autòmat<sup>139</sup>; en aquests casos, els estudis tendeixen a demostrar una major probabilitat d'equilibris cooperatius. No obstant això, els problemes de l'aplicació de la teoria dels autòmats a la teoria de jocs són molt grans, degut a que apareix el conegut problema de la recursió infinita sobre el càlcul de les estratègies òptimes<sup>140</sup>, apareixent la necessitat, com destaquen Banks i Sundaram, de desenvolupar

*"A more general theory of bounded rationality and decision cost in repeated games which would address the issue of complexity both in terms of the computation of optimal strategies as well as the actual implemetation of such strategies"*<sup>141</sup>

---

<sup>138</sup> Vid. I. Gilboa, "The Complexity of Computing Best Response Automata in Repeated Games", Journal of Economic Theory, vol. 45, 1988; i E. Ben-Porath, "The Complexity of Computing a Best Response Automaton in Repeated Games with Mixed Strategies", Games and Economic Behavior, vol. 2, 1990. El problema bàsic és que, com diu Ben-Porath, "We do not have a theory about how players reason when deciding which strategy to play, we just have equilibrium concepts" (pàg. 2).

<sup>139</sup> Per un enfocament probabilístic, considerant que el jugador racional no sap exactament a quin tipus de jugador s'enfronta, vid. R. Aumann i S. Sorin, "Cooperation and Bounded Recall", Games and Economic Behavior, vol. 1, 1989. Un enfocament més determinista, però amb una major definició de la capacitat estratègica de l'autòmat, es pot trobar a I. Gilboa i D. Samet, "Bounded versus Unbounded Rationality: The Tyranny of the Weak", Games and Economic Behavior, vol. 1, 1989. Aquestes propostes connecten parcialment amb el discutit text de T. Schelling "Hacia el abandono de simetria en la teoria de juegos" (a La Estrategia del Conflicto, Tecnos, Madrid, 1964; ed. or. en anglès a The Review of Economics and Statistics, vol. 41, 1959), però amb la diferència fonamental que les noves propostes identifiquen simetria amb racionalitat i, per tant, per analitzar la simetria modifiquen les assumpcions sobre les capacitats d'un jugador. D'altra banda, Schelling, amb la teoria dels «punts focals», proposa que els jugadors siguin capaços de trobar punts d'equilibri fora de la simetria matemàtica, motivats per aspectes concrets del entorn en què es desenvolupa la interacció, o per creences compartides sobre la conducta a seguir (vid. apt. 2.1.2.).

<sup>140</sup> Entendre el jugador racional com una màquina de computació programada implica, com indica K. Binmore ("Modeling Rational Players I.", Economics and Philosophy, vol. 3, 1987) trobar-se amb el problema de la infinitat incompleta i, doncs, amb la inevitable aparició de situacions que el model no integra.

<sup>141</sup> J. Banks i R. Sundaram, "Repeated Games...", cit., 1990, pàg. 116.



El problema d'aquestes propostes és que els elements de racionalitat limitada són imposats externament, com restriccions informacionals als subjectes per aplicar la seva capacitat de cognició racional, definida segons els axiomes convencionals, incloent el principi de maximització. Per això, el problema bàsic es manté: donada una situació concreta en què no és possible imposar limitacions a la racionalitat dels jugadors, experimentalment es pot observar com aquests cooperen en una estructura tipus dilema del presoner, mentre que seguint els postulats de la racionalitat, aquesta cooperació seria impossible.

D'altra banda, des de l'enfocament evolutiu també s'han desenvolupat intents per a explicar l'aparició de la cooperació. Sintèticament, s'ha adoptat la idea d'aplicar regles de comportament en els models d'interacció, per trobar estratègies evolutives estables que siguin cooperatives<sup>142</sup>. També s'ha intentat combinar la racionalitat limitada i les estratègies evolutives. Friedman i Rosenthal<sup>143</sup> presenten un model en el qual, segons el tipus de joc, l'individu aplica una determinada regla de comportament. Llavors consideren que excepcionalment, un individu pot estar insatisfet amb una determinada regla de selecció d'estratègia, i buscar una de millor, que no ha de ser necessàriament l'òptima<sup>144</sup>.

---

<sup>142</sup> Paral·lelament, els autors que adopten aquest enfocament discuteixen teories de l'aprenentatge per donar un suport teòric -però exogen- a la formació de les regles de comportament. Un exemple d'aquests models es pot trobar a U. Witt, "Evolution and Stability of Cooperation without Enforceable Contracts", a KYKLOS, vol. 39, 1986. Una altra aproximació, centrada en l'efecte final dels jocs finits, es pot trobar a R. Selten i R. Stoecker, "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames. A Learning Theory Approach", Journal of Economic Behavior and Organization, vol. 7, 1986. Una altra línia és la incorporació de l'atzar per explicar l'aparició d'estratègies evolutives estables i eficients, com a forma d'incorporar la possibilitat que els jugadors realitzin errors en el seu comportament intencional (cf. D. Fudenberg i E. Maskin, "Evolution and Cooperation in Noisy Repeated Games", American Economic Review. Papers and Proceedings, vol. 80, 1990).

<sup>143</sup> J. Friedman i R. Rosenthal, "A Positive Approach to Non-Cooperative Games", Journal of Economic Behavior and Organization, vol. 7, 1986.

<sup>144</sup> Per això estableixen unes regles d'immigració i d'emigració, basades en comparacions amb el pagament més elevat de la matriu de pagaments del jugador (suposant que les diferències estan afectades per les característiques de l'entorn). Si la diferència és molt gran, augmenta la tendència a canviar d'estratègia. Quan ja no hi ha canvis, s'entén que la interacció s'aproxima a un equilibri de Nash.

A partir de la creixent discussió dels darrers anys sobre aquestes qüestions, es pot indicar que s'està desenvolupant una intensa competició per tal de proposar cada cop models més sofisticats, però també amb definicions de restriccions informacionals més específiques, cosa que té com a conseqüència la limitació dels seus àmbits d'aplicació possibles. En conjunt, es pot afirmar que les propostes han arribat a una situació particularment confusa<sup>145</sup>, perquè impliquen el supòsit d'uns actors molt sofisticats i d'uns requeriments informacionals molt específics per explicar en aquest context, mitjançant l'exclusivitat de la hipòtesi de racionalitat, l'aparició de cooperació. En aquestes condicions, el teorema d'impossibilitat de K. Basu (apt. 2.2.4.), il·lustra clarament el problema, quan destaca que sempre apareix la mateixa dependència en relació a la racionalitat i els seus axiomes, tot i la imposició de limitacions a les capacitats informacionals dels actors.

La línia que resta explorar dins d'aquests processos «intermedis» de decisió és el desenvolupament de models amb limitacions cognitives a la racionalitat dels jugadors, seguint les indicacions de H. Simon<sup>146</sup>. Des d'aquesta perspectiva, generalment es postula un procés de generació d'alternatives, amb el seu cost associat (en comptes d'un conjunt fixat), i no es coneix la distribució de probabilitat dels resultats (es considera incerta o estimada, però no es suposa subjectivament coneguda), ja que depèn dels límits cognitius de l'individu<sup>147</sup>. Cal afegir que la

---

<sup>145</sup> Però poden tenir altres aplicacions interessants, com per exemple a la programació informàtica. Una recent contribució, proposant un joc amb autòmats probabilístics, amb possibilitats d'ésser limitats per l'espai, el temps i la informació, es pot trobar a A. Condon, *Computational Models of Games*, MIT Press, Cambridge, 1989.

<sup>146</sup> "The term "bounded rationality" is used to designate rational choice that takes into account the cognitive limitations of the decision-maker - limitations of both knowledge and computational capacity", (H. Simon, "Bounded Rationality", art. cit., pàgs. 266-7). Vegeu també l'apt. 1.1.4..

<sup>147</sup> Analitzant el concepte de racionalitat limitada, Ph. Mongin destaca (a "Simon, Stigler et...", art. cit., 1986) l'existència de dos tipus generals de limitacions a la racionalitat en la concepció de H. Simon. La incertesa externa o limitacions informacionals (el càlcul sobre les conseqüències de cada acció, o racionalitat substantiva); i la incertesa interna o limitacions cognitives (el raonament sobre les accions possibles, o racionalitat de procediment). En ambdues es considera necessari el supòsit d'una combinació de mecanismes cognitius i regles de comportament en els procediments de decisió (que generalment tendeixen a assolir la satisfacció). D'aquesta manera, els models convencionals, apart de només considerar l'existència de incertesa externa, apliquen el mecanisme cognitiu derivat de la hipòtesi de racionalitat, més

racionalitat costosa (informacionalment i cognitivament) no pot ésser determinada de forma íntegra per l'esquema formal d'una teoria de la decisió<sup>148</sup>. Les propostes són molt diverses<sup>149</sup> i, en termes generals, es poden dividir en grans grups. El primer manté els postulats de la hipòtesi de racionalitat, incorporant a més de les limitacions sobre la capacitat d'obtenir informació, limitacions a la capacitat dels agents per discriminar les accions possibles. Per a aquest fi s'incorporen teories sobre l'aprenentatge, que integra la informació sobre les accions anteriors dels agents amb una certa capacitat de generar experimentació<sup>150</sup>.

El segon grup intenta apropar la racionalitat limitada (o de procediment) a la modelització de la teoria de jocs, en rebutjar l'aplicació constant del principi de maximització en els models sobre l'aprenentatge. Es postula una estratègia de satisfacció<sup>151</sup>, amb procediments computacionals simples i capaços d'incorporar les

---

el principi de maximització.

<sup>148</sup> Com indica H. Simon, "to make predictions about behavior on the basis of a satisficing theory requires much more empirical data about, for example, aspiration levels and their adaptivity, than does prediction on the basis of the optimizing theory. The magnitude of the difference becomes less, however, when we recognize that the optimizing theory says nothing about the shape or content of the utility function. It simply postulates a consistency of behavior over time that may not be found if the decision maker is satisficing instead of optimizing" (a "Satisficing", art. cit., 1987, pàg. 244).

<sup>149</sup> Com a cas extrem, inclús es podria postular un hipotètic «Garbage can game theory». Només es tractaria d'adoptar estratègies no basades en la racionalitat, sinó mitjançant l'atzar, de funcions amb distribucions de probabilitat i s'introduirien funcions aleatòries pel repartiment dels pagaments (cf. M. Cohen, J. Narch i J. Olsen, "A Garbage Can Model of Organizational Decision-Making", Administrative Science Quarterly, vol. 17, 1972). En aquest cas, caldria investigar quan i com conjunts de estratègies, resultats i actors es combinen per produir equilibris (o situacions inestables i canviants): la definició d'equilibri canviaria totalment, transformant-se en una noció d'èxit. Quan un actor aconsegueix el seu propòsit. Tot i que la possibilitat d'aplicació existeix, la descripció dels trets institucionals és més important per aquest enfocament que l'anàlisi de les decisions dels agents.

<sup>150</sup> Un enfocament normatiu es pot trobar a V. Crawford i H. Haller, "Learning How to Cooperate: Optimal Play in Repeated Coordination Games", Econometrica, vol. 58, 1990. Un enfocament més descriptiu, destacant el paper de l'experimentació, es troba a M. Moore i M. Chapman Moore, "Adaptive Learning, Adaptive Utility, and Rational Behavior in a Repeated Prisoner's Dilemma", Journal of Risk and Uncertainty, vol. 2, 1989 (per una ampliació, vegeu la secció 4.3.).

<sup>151</sup> El terme «satisfacció», definit pel propi H. Simon, s'entén com "a decision-maker (...) who chooses an alternative that meets or exceeds specified criteria, but that is not guaranteed to be either unique or in any sense the best, is said to satisfy" (H. Simon, "Satisficing", The New Palgrave, vol. 3, 1987).

inconsistències de les eleccions humanes. Així, davant la incorporació tradicional de la irracionalitat en forma de «fallides» lògiques, es pot proposar entendre com la primera font d'errors les imperfeccions dins el mateix procés de raonar: el procediment de procesar les dades, afectat tant per la capacitat decisional de l'actor com pels factors de l'entorn en què es juga el joc. Per desenvolupar aquest procediment en la teoria de jocs, incloent les limitacions cognitives dels jugadors, l'estratègia d'investigació en analitzar el procés d'equilibri ha d'ésser buscar un punt intermedi entre els enfocaments evolutius, basats sols en regles de comportament (amb una complexitat interna dels jugadors baixa), i els enfocaments cognitius, que entenen els jugadors com a màquines molt complexes internament.

L'objectiu que és proclama és arribar a modelar explícitament alguns aspectes del procés de pensar que porta a un equilibri en la interacció entre els jugadors, ja que aquest procés és el que determina en bona part la natura de l'equilibri assolit. En aquest sentit, cal destacar la proposta de Binmore per construir una màquina (autòmat) computacional capaç en certa mesura d'introspecció i d'imposar un grau d'arbitrarietat, amb la qual cosa introdueix una visió algorítmica de la racionalitat<sup>152</sup>. El model es formula en termes de màquines de Turing: garantir amb un algoritme d'aproximació, una resposta per tots els inputs rellevants, acceptant cert nivell d'error mitjançant la incorporació de regles d'aturada (del càlcul) per a evitar el problema de la regressió infinita. Per tant, es tractaria de màquines d'alta complexitat, que poguessin tornar a programar-se segons l'evolució del joc. En aquest model, caldria pressuposar una visió de l'oponent com un subjecte

---

L'estratègia de satisfacció es pot convertir en una de maximització si es contemplen els costos de recerca (informació i computació). Però maximitzar aquests costos imposa un problema de segon ordre que encara és més complicat que el primer, i així successivament.

<sup>152</sup> K. Binmore, "Modeling Rational Players II.", *Economics and Philosophy*, vols. 4, 1988. Binmore destaca també que els factors de l'entorn poden provocar desviacions de l'estratègia perfecta, sense que això tingui res d'irracional. En aquesta concepció es pot detectar una continuïtat amb la teoria dels «punts focals» de T. Schelling. Un altre aspecte que també destaca Binmore és la necessitat de sortir de l'ortodòxia bayesiana i adoptar alguna arbitrarietat en l'aplicació de probabilitats.

racional, però considerant possible que l'individu segueixi alternativament diverses variacions de la hipòtesi de racionalitat. Segons la forma i els costos del càlcul, es podria considerar l'existència de regles heurístiques que determinessin quina variació aplicar en cada cas. Aquestes regles podrien ésser enteses com una metàfora d'un procés evolutiu, amb la possibilitat que existeixen diverses regles alternatives en una població donada, en competició entre elles per la seva supervivència<sup>153</sup>. En aquest context, es suposa que en alguns casos els jugadors poden revisar les seves prediccions sobre les estratègies dels contraris, per millorar la seva utilitat esperada, donada la distribució de probabilitats observada de les estratègies del contrari<sup>154</sup>. Si es limiten els càlculs a un nivell determinat d'aproximació, les màquines actuen com a satisfactòries i els equilibris aproximats poden ésser molt diferents entre ells. A més incertesa sobre el contrari, el domini de les regles de comportament és superior, deixant menys lloc per als càlculs amb models que es basen en els postulats de la racionalitat<sup>155</sup>. Donada aquesta línia de modelització, en general, i aplicant la interpretació de Binmore<sup>156</sup>, quan dos diferents tipus de jugadors fan dues anàlisis diferents del joc, i ambdós pensen que la seva anàlisi és aproximadament correcta, depèn de si el nombre de dades objectives és molt elevat, que la

---

<sup>153</sup> Com diu K. Binmore, "evolutionary processes work on a master-program which has the capacity to choose strategies in a wide variety of disparate games, some of which may never have been played before" (a "Modeling...II.", art. cit., 1988, pàg. 16). Es pot arribar a una situació de domini d'un program amb unes determinades característiques, essent ja de coneixement comú; però també poden aparèixer nous models mutants, i tornar a iniciar-se de nou la competició.

<sup>154</sup> Així ho assenyala K. Binmore, "as the game proceeds, players will learn about the structure of the machines they are playing against by studying the moves they choose" (a "Modeling... II.", art. cit., 1988, pàg. 22).

<sup>155</sup> Aquesta seria la proposta de R. Heiner, "Origin of Predictable Behavior: Further Modeling and Applications", *American Economic Review. Papers and Proceedings*, vol. 75, 1985. També K. Wilde et al., "Knowledge, Uncertainty and Behavior", *American Economic Review. Papers and Proceedings*, vol. 75, 1985. De forma més complexa K. Binmore proposa modelar una teoria de la introspecció, suposant que faria la màquina oposada en la mateixa situació, però també accepta la necessitat d'incorporar un algoritme conjectural, forcosament arbitrari o fruit d'un accident històric. Per simular l'altra màquina, una màquina treballa sobre la hipòtesi que ambdues tenen el mateix programa. Els jugadors poden aprendre el procés de pensar dels seus oponents observant el seu comportament, abandonant el principi de que les desviacions del comportament d'equilibri han d'ésser enteses com incorrelacionades.

<sup>156</sup> K. Binmore, "Modeling...II.", art. cit., 1988.

probabilitat de què existeixen grans diferències en les regles d'aproximació i els sistemes de creences sigui baixa. Si la probabilitat segueix essent baixa i persisteixen grans diferències, hom ha de pensar en la possibilitat que, des d'una perspectiva evolutiva, existeixen regles de comportament molt diferents entre els jugadors.

Una altra proposta a destacar, degut a les seves possibilitats de combinació dins aquesta perspectiva, és l'intent de Rosenthal d'aïllar l'acció maximitzadora dels jugadors com un cas extrem, considerant una àmplia gamma d'altres possibilitats<sup>157</sup>. El model requereix que la probabilitat d'utilitzar una estratègia per part d'un agent sigui una funció monòtona no decreixent del seu pagament esperat. Rosenthal defineix aquesta funció a partir d'una variable,  $t$ , com una mesura de la limitació de la racionalitat de cada actor. El seu sentit és que com més gran és  $t$ , més sensible és el jugador a les desviacions pròpies en relació a les seves millor respostes (quan  $t$  és infinit per a tots els jugadors, apareix l'equilibri de Nash)<sup>158</sup>. Així, en ésser una funció de  $t$  el conjunt de solucions del joc, es pot observar com es redueixen els comportaments possibles a mesura que es limita la desviació respecte a la racionalitat. El model de Rosenthal també suposa l'existència d'una «racionalitat aproximada», que pot incloure de forma més simple diferents tipus de jugadors amb models diferents de racionalitat, segons els supòsits d'aquesta que incorporin (a la vegada, dins cada jugador, poden haver diferents tipus disponibles).

Sense proposar una modelització concreta, objectiu que queda al marge

---

<sup>157</sup> R.W. Rosenthal, "A Bounded-Rationality Approach to the Study of Noncooperative Games", *International Journal of Game Theory*, vol. 18, 1989. Per una aplicació més simple d'un principi similar, vegeu R. Goodin, "Some New Sources of Social Conflict: Transformations of Mixed-Motive Games", *British Journal of Sociology*, vol. 39, 1988.

<sup>158</sup> També es pot entendre com un indicador de la intensitat de les seves preferències o de la precisió de les seves estimacions: d'aplicar una probabilitat igual, a establir un càlcul absolutament racional. Com el propi Rosenthal afirma, "*I interpret the parameter  $t$  as a measure of the importance of the game (with its particular utility scaling) to the player relative to his interests outside the game. Thus, if  $t = 0$ , the player has no interest in payoff differences in the game, and the equal-probability solution to the conditions above reflects the player's indifference and disinclination to devote effort to the chore of making decisions about how to play. As  $t$  grows, the player becomes more sensitive to the payoff differences induced by his strategy choices*", ("A Bounded-Rationality...", art. cit., 1989, pàg. 274).

d'aquest treball, es pot destacar que els eixos per a aplicar un concepció de la racionalitat no normativa a l'estructura de la teoria de jocs queden força perfilats. *Suprimir el supòsit de maximització, alterar o reduir els axiomes que conformen la hipòtesi de la racionalitat, introduir regles de cognició, o realitzar combinacions entre les diverses possibilitats, són les estratègies d'investigació possibles. Quina és la més afortunada és quelcom que resta encara depenent d'un major desenvolupament teòric, tot i que cal recordar que no es tracta de cap competició.* Probablement existirà un estret vincle entre la teoria substantiva que hom vulgui elaborar, l'àmbit d'aplicació i les característiques del model de decisió que es proposi.

### 2.3. Teoria de jocs, ciència social i acció col·lectiva.

Al llarg d'aquest capítol s'han discutit les possibilitats actuals d'ampliar la capacitat descriptiva de la teoria de jocs, a través de diverses línies de recerca. D'aquestes possibilitats, es pot deduir que crítiques contundents a les aplicacions de la teoria de jocs, com les que realitza M. Bunge en el cas de la ciència política, no estan ben fonamentades<sup>159</sup>. Sembla estrany que alguns excel·lents filòsofs de la ciència, com M. Bunge, tinguin quelcom essencial contra qualsevol forma d'aplicació de la teoria de jocs a les ciències socials. En tot cas, dels apartats precedents hom pot recordar que aquestes acusacions no afecten de forma destacada el conjunt de possibilitats de la teoria de jocs, sinó específicament alguns dels seus desenvolupaments, els més vinculats amb l'enfocament metodològic criticat en aquestes pàgines, com els que mantenen en exclusiva el supòsit de racionalitat per a resoldre tots els seus problemes teòrics.

D'aquests possibilitats, algunes es poden interconnectar fàcilment, altres són mútuament incompatibles. En termes generals, les propostes comentades mostren l'estat d'efervescència actual de la teoria de jocs i ofereixen un ampli panorama de nous problemes que qüestionen les teories -econòmiques i polítiques- basades en els postulats tradicionals de la teoria de la decisió racional i la teoria de jocs. Malgrat tot, hom ha de reconèixer que molts dels problemes de la teoria de jocs ja han estat prèviament esmentats en el primer capítol, en tractar dels problemes de l'anàlisi de les decisions. Per aquest motiu, és fàcil de suposar que tant l'elaboració de millors teories sobre els processos de decisió, com l'aclariment conceptual de les distintes formulacions existents, poden tenir un important impacte en els nous models de la teoria de jocs que es formulin.

---

<sup>159</sup> Per exemple, la següent afirmació: "*(Game theory) boils down to two postulates, neither of which is true: that there is a utility function for everybody and everything, and that to be rational is to always act so as to maximize one's expected utility*" (cf. M. Bunge, "Game Theory is Not...", art. cit., 1989, pàg. 210).



Tanmateix, també s'han analitzat altres problemes de la teoria de jocs no directament relacionats amb les dificultats de conceptualitzar els processos de decisió. Entre aquests, alguns fan referència a la manca d'elements i proposicions fàctiques en la pròpia teoria. De les propostes analitzades en aquest capítol, s'han ressenyat diverses dificultats, així com suggeridores possibilitats de desenvolupament. Precisament per aquest motiu també s'ha fet més necessària la disposició d'un camp teòric amb referències fàctiques, per a poder aprofundir en l'anàlisi que es desenvolupa. Aquest camp, l'anàlisi del qual constitueix la segona part de la tesi, es refereix als processos d'acció col·lectiva i als problemes que existeixen per la seva formació.

Avançant el que serà un dels temes centrals de la segona part de la tesi, es pot destacar que la introducció de la perspectiva de les limitacions a la racionalitat en l'anàlisi de l'acció col·lectiva trenca amb la concepció tradicional que presentava el problema de l'acció col·lectiva de forma unitària. *Seguint l'enfocament presentat, que contempla una àmplia varietat de possibilitats de modelitzar l'acció individual, el problema es pot fragmentar. Aquesta fragmentació permet afrontar l'acció col·lectiva com un problema teòric múltiple, amb diversos nivells i perspectives d'anàlisi, que alhora possibiliten la introducció de proposicions teòriques emergents.* En el sentit esmentat, cal indicar també que igual d'interessant que la fonamentació de l'acció col·lectiva en termes d'accions individuals, és la discussió del problema sobre com relacionar aquesta fonamentació amb proposicions emergents. Hauria de permetre explicar l'aparició i el manteniment de la interacció i d'accions comunes, així com de les normes i els trets culturals -interns i externs a la pròpia acció-compartits pels protagonistes d'una acció col·lectiva. Aquest és un camp d'aplicació diferent a l'anàlisi de les xarxes, els conjunts de relacions i dependències entre agents, institucions, col·lectius, etc.<sup>160</sup>. Per explicar l'acció col·lectiva interessa

---

<sup>160</sup> Per exemple, sobre aquestes qüestions, vegeu l'aproximació de F. Scharpf, "Games Real Actors Could Play: The Problem of Connectedness", Max-Planck-Institut für Gesellschaftsforschung, Discussion Paper 90/8, Colonia, 1990.

especialment les preferències i les normes, els equilibris, els espais de decisió individuals similars, la valoració subjectiva de les recompenses; i també la discussió sobre el paper del principi d'optimització. En els següents capítols, per analitzar els problemes de l'acció col·lectiva, s'aplica aquest enfocament, amb el doble propòsit, com assenyala T. Schelling, de mostrar les limitacions de la teoria, així com proposar línies per un desenvolupament dels models teòrics basats fonamentalment en pressupòsits sobre l'acció individual:

*"A game may be useful in the articulation of a theoretical model if it is designed for that purpose, just as games richer in detail may help to fit theory into its institutional context (...) one kind of game may be used to elucidate a theoretical model, the other to show its limitations"*<sup>161</sup>

La creixent complexitat en la formulació de models d'home presentada en aquest capítol aproxima els problemes de conceptualització existents a la dinàmica de la ciència cognitiva. Encara que el tema no es tracta a la segona part de la tesi, és interessant de destacar algunes coincidències sorprenents, mostrant alhora les seves similituds i diferències. L'aprofundiment en la conceptualització dels models d'home hauria de conduir a la discussió de temes propers als de la ciència cognitiva, i aquest seria un desenvolupament coherent dels temes plantejats fins al moment en aquest treball. Les creixents possibilitats de realitzar simulacions informatitzades augmenten el seu interès, encara que la seva base teòrica en relació als processos agregats és més reduïda. És bàsicament per aquest darrer motiu que no s'ha desenvolupat aquesta alternativa, preferint afrontar directament els problemes de construcció teòrica de la teoria de jocs en un àmbit aplicat.

Diversos enfocaments conflueixen en les ciències cognitives (filosofia, psicologia, intel·ligència artificial, neurociència, antropologia i lingüística), amb

---

<sup>161</sup> T. Schelling, "Experimental Games and Bargaining Theory", *World Politics*, vol. 14, 1961, pàg. 47.

vincles de diversa intensitat entre ells. El propòsit que els connecta és descobrir les capacitats de representació i computació de la ment, i la seva representació funcional i estructural en el cervell<sup>162</sup>. El problema bàsic d'aquest propòsit és que els sistemes que s'estudien són massa complexos per a ser tractats de forma purament descriptiva i, per tant, no es pot evitar de fer proposicions sobre estructures i relacions causals. L'enfocament, tot i possibilitar suggeridores interpretacions parcials, ha comportat algunes decepcions degut a les insospitades dificultats aparegudes en tractar la gran complexitat d'aquests sistemes. En una primera fase es desenvoluparen procediments generalitzats d'inferència -o màquines lògiques-, de forma similar a les propostes basades essencialment en la racionalitat esmentades als apartats anteriors, però aquests models -programes per a jugar als escacs, o per a provar teoremes matemàtics- van presentar uns resultats dolents en comparar-se amb l'activitat humana en aquests camps. Una segona fase es va centrar en els procediments de representació del coneixement. La hipòtesi subjacent a aquest enfocament era que per tractar problemes complexos, les persones acostumen a utilitzar regles, procediments pautats, experiències anteriors pròpies o acumulades, incorporant generalment molt poca creació pròpia per a implementar una decisió. Però aquesta línia d'investigació també va acabar en un cul de sac, ja que l'acumulació de coneixement i regles provocava incertesa i confusió, i calia introduir moltes condicions específiques als problemes de la decisió<sup>163</sup>.

Comparant les dues perspectives, es pot destacar que en les teories sobre la

---

<sup>162</sup> Citat per D. Heckathorn ("Cognitive Science, Sociology, and the Theoretic Analysis of Complex Systems", Journal of Mathematical Sociology, vol. 14, 1989), referint-se a Sloan Foundation, Cognitive Science, 1978: Report of the State of Art Committee, Nova York, 1978.

<sup>163</sup> Per a aquesta descripció, D. Heckathorn, "Cognitive Science...", art. cit., 1989. Un exemple d'aquests models mixtes aplicats a la ciència social és ofert per K. Carley, a "The Value of Cognitive Foundations for Dynamic Social Theory" Journal of Mathematical Sociology, vol. 14, 1989. Carley formula una estructura cognitiva del comportament en tres parts: i) Un esquema estructurat de representació del coneixement, ii) un conjunt de principis cognitius (definició relativa, comprensió immediata, consistència local, canvis monotònics, etc.), i iii) un conjunt d'operadors cognitius (adquisició de coneixement per descobriment, per generalització, per explicació; comunicació del coneixement; elecció del company d'interacció).

racionalitat, apareixen més dificultats per modelitzar el comportament humà (com ja s'ha discutit àmpliament), però en canvi, ofereix més possibilitats per a explicar la interacció des d'un punt de vista col·lectiu, mitjançant la formulació d'equilibris. En aquest sentit, els models cognitius han de desenvolupar complicades simulacions per formular les seves possibles prediccions, amb una escassa capacitat d'integració teòrica i mantenint-se a un nivell força descriptiu dels processos socials. No obstant això, aquesta perspectiva, en centrar-se en les formes de cognició humana, especificant i modelant diversos mecanismes psicològics, aporta una visió més plausible -tot i existir molts problemes pendents- sobre el procés d'interacció des del punt de vista individual<sup>164</sup>. D'altra banda, els models que treballen amb la noció d'equilibri ofereixen la possibilitat de postular els trets determinants del comportament col·lectiu amb menors pressupòsits i una major abstracció (entre d'altres, amb la hipòtesi de racionalitat). Potser una integració teòrica d'elements del enfocament cognitiu, amb l'ajuda de simulacions per ordinador, junt a models de la teoria de jocs, podria augmentar força la capacitat interpretativa d'aquestes perspectives. Però aquesta ja és una qüestió que s'allunya dels plantejaments de la tesi, encara que podria contribuir-hi indirectament l'anàlisi de les possibilitats d'integració teòrica que permet la teoria de jocs.

---

<sup>164</sup> Tot i així, també presenta les seves fronteres per definir l'ésser humà, com destaca H. Putnam en la seva crítica de la identificació entre els estats mentals i els estats que es poden caracteritzar computacionalment, ja que els primers tenen una capacitat emergent que no pot ésser descrita en termes de funcionalitat física (Representation and Reality, MIT Press, Cambridge, 1988).

**SEGONA PART:**

**ACCIÓ COLLECTIVA I RACIONALITAT**

**CAPÍTOL TERCER:**

**L'acció col.lectiva interpretada des del  
paradigma de l'elecció racional.**

### 3.1. L'anàlisi de l'acció col·lectiva.

Per aprofundir en l'anàlisi dels problemes metodològics sobre la racionalitat plantejats a la primera part de la tesi, és necessari disposar de referents teòrics provinents de les ciències socials. Per a desenvolupar aquest propòsit, aquesta segona part està dedicada al terreny de les teories de l'acció col·lectiva, proper a la ciència política. El capítol comença amb una discussió sobre la definició d'aquest camp, a la qual segueix una anàlisi de la formació d'aquestes teories i dels problemes interns que ha generat i genera el seu procés de construcció teòrica. En aquest sentit, la tasca de reflexió metodològica apareix a través de la pròpia discussió sobre els problemes més rellevants de la lògica de l'acció col·lectiva i de les propostes analítiques que es consideren.

#### 3.1.1. Definició de l'acció col·lectiva.

Una primera aproximació al concepte d'acció col·lectiva pot ser assenyalar que aquesta es refereix a les interaccions entre individus que comparteixen un objectiu comú i tenen en compte els efectes de les decisions dels altres en establir les seves pròpies decisions. Per aquest motiu, es considera que una acció col·lectiva implica l'existència d'una estructura d'interacció estratègica. També es pot definir en contraposició a la interacció en una estructura de mercat, que implica la recerca d'objectius diversos, de vegades enfrontats, en un context en què els individus no calculen els efectes de les decisions dels altres sobre les seves accions i, per tant, decideixen com si es tractés d'una decisió front la natura.

La capacitat d'un conjunt d'individus -conscients de la seva interdependència- per a portar a terme un propòsit comú pot explicar-se des de dues vessants. La que

cerca una solució «interna» per a resoldre el problema de com es produeix l'acció col·lectiva, «interna» perquè únicament introdueix en els seus models característiques dels individus, tot analitzant les condicions per l'aparició de coordinació individual amb o sense relacions comunicatives. La segona vessant proposa la necessitat de buscar solucions «externes», en el sentit d'afegir a l'anàlisi el paper dels mecanismes institucionals que impulsen l'acció col·lectiva<sup>1</sup>. Entre els defensors d'ambdues orientacions analítiques hi ha un debat teòric sobre la potencialitat dels seus respectius enfocaments. S'argumenta, per una banda, que la solució «interna» és la clau bàsica de l'acció col·lectiva i que, per tant, és prioritari analitzar la seva lògica; mentre que d'altra banda es defensa que sense mecanismes institucionals no es possible explicar gairebé cap tipus d'acció col·lectiva<sup>2</sup>. La teoria de jocs, com tot instrument formal, pot ser aplicada per desenvolupar els dos tipus de solucions, adoptant diferents status metodològics (vid. apt. 2.2.1.). No obstant això, cal senyalar que tot i que l'acció col·lectiva pot ser interpretada a través de la teoria de jocs, l'acció col·lectiva no és pròpiament la teoria de jocs, ja que hom ha d'afegir a la teoria de jocs altres teories, amb contingut fàctic, per analitzar una situació d'acció col·lectiva.

Aquesta aproximació a la noció d'acció col·lectiva pot incloure fenòmens tractats per diverses ciències socials, però hom pot argumentar que els aspectes referents a la seva lògica interna estan força vinculats amb un dels problemes centrals que analitza la ciència política: tot problema d'acció col·lectiva implica una dimensió política, degut a la necessitat de resoldre les dificultats per a aconseguir els objectius desitjats i

---

<sup>1</sup> La divisió està inspirada en la caracterització de M. Taylor, *The Possibility of Cooperation*, Cambridge University Press, Cambridge, 1987. Aquesta divisió recull implícitament les dues grans posicions metodològiques sobre l'autonomia de les hipòtesis sobre el comportament individual esmentades a la primera part de la tesi (especialment apt. 1.1.3.).

<sup>2</sup> Com una mostra d'aquesta discussió, comentada més àmpliament al llarg del capítol, vegeu l'article de M. Hechter "On the Inadequacy of Game Theory for the Solution of Real-World Collective Action Problems", a K.S. Cook i M. Levy (Eds.), *The Limits of Rationality*, University of Chicago Press, Chicago, 1990.



reconèixer com s'han obtingut els resultats<sup>3</sup>. Però la ciència política és una disciplina que incorpora diversos paradigmes, i tan sols alguns d'aquests fan ús de la noció d'acció col·lectiva en els termes plantejats per abordar la qüestió esmentada<sup>4</sup>. Donada aquesta situació, sense defensar que la ciència política s'hagi de dedicar exclusivament a l'estudi dels problemes d'acció col·lectiva, es pren només l'enfocament que té el seu origen en models econòmics, per a construir seguidament una aproximació més detallada a aquest concepte.

### 3.1.2. La gènesi del problema de l'acció col·lectiva.

Des del punt de vista de la seva gènesi històrica, cal destacar que la relació entre grups d'interessos i interessos individuals fou esbossada durant els anys cinquantes per W. Baumol de forma força clara. Posteriorment, als anys seixantes, els problemes sobre el conflicte d'interessos en grups varen ser formulats en termes teòrics per M. Olson. Una línia d'investigació propera, centrada en el paper de l'Estat, fou iniciada formalment a primers de segle per Frank Knight. Aquesta línia ha influït l'obra d'autors com A. Downs, J. Buchanan, G. Stigler i M. Friedman, els quals constitueixen

---

<sup>3</sup> Inclús hi ha autors, com M. Taylor, que inclou dins la definició de 'política' l'expressió acció col·lectiva: "*Politics is the study of ways of solving collective action problems*" (a "Cooperation and Rationality: Notes on the Collective Action Problem and Its Solutions", a K.S. Cook i M. Levi (Eds.), *The Limits of...op. cit.*, 1990, pàg. 224). També Elster menciona que la política és "*The study of ways of transcending the Prisoners' Dilemma*" (a "Some Conceptual Problems in Political Theory", B. Barry (ed.), *Power and Political Theory*, Wiley, Londres, 1976, pàgs. 248-249).

<sup>4</sup> Per una revisió panoràmica dels paradigmes que s'aproximen a aquest problema, vegeu P. Dunleavy, *Democracy, Bureaucracy and Public Choice*, Harvester-Wheatsheaf, Nova York-Londres, 1991, pàgs. 13-44. Dunleavy destaca quatre grans corrents sobre interessos i acció col·lectiva: l'aproximació pluralista, la corporatista, els models de la 'nova dreta' i els models basats en la 'lògica de l'acció col·lectiva'. Aquests darrers són els que es discuteixen en aquestes pàgines.

l'enfocament conegut com l'escola de l'«elecció pública»<sup>5</sup>. Ambdós enfocaments comparteixen força pressupòsits, però el primer afronta més centralment els problemes bàsics de la lògica de l'acció col·lectiva que l'altre, dedicat a l'estudi dels efectes dels mecanismes d'intervenció estatal sobre la societat i el mercat. Els autors de l'escola de l'«elecció pública» comparteixen l'heurística d'aplicar l'aparell conceptual de l'economia al comportament *polític* dels individus, en aquest cas els que actuen en el sector públic (electors, buròcrates, polítics...): afirmen que l'acció col·lectiva que impulsa l'Estat (per a la provisió de béns públics) és una forma de resoldre les fallides en el funcionament de mercats perfectament competitius (problemes de competència imperfecta)<sup>6</sup>. Amb aquest argument, entenen que l'acció col·lectiva es pot produir tant mitjançant la imposició de mesures legals per part d'una estructura de poder, com per la realització d'acords polítics entre individus o organitzacions representatives. Quan el mercat és incapaç de produir diversos béns públics necessaris per a una comunitat (com la defensa del territori) o d'evitar «externalitats» negatives (com la contaminació ambiental), cal arribar a trobar formules col·lectives de proveir aquestes necessitats. Degut a les dificultats per l'aparició de l'acció col·lectiva espontània, aquesta ha de ser molt sovint impulsada per un poder centralitzat, amb capacitat d'imposar mesures coercitives pel compliment de les seves polítiques<sup>7</sup>. Conseqüentment, els autors que treballen des

---

<sup>5</sup> Vid. pel primer enfocament, W.J. Baumol, Welfare Economics and the Theory of the State, Longmans, Londres, 1952, i M. Olson, The Logic of Collective Action, Harvard University Press, Cambridge, 1965; i per el segon, J. Coleman, "Competition and Cooperation", Ethics, vol. 98, 1987; també E. Fuentes Quintana, "James Buchanan. Premio Nobel de Economía 1986", Papeles de Economía Española, vol. 31-32, 1986. Cal destacar que, tot i que Baumol aplica la seva anàlisi per a explicar el paper de l'Estat i entén la intervenció d'aquest com la solució al problema de l'acció col·lectiva, també discuteix àmpliament la lògica de la formació voluntària de grups (sobre la seva contribució, vid. R. Hardin, Collective Action, Johns Hopkins University Press, Baltimore, 1982, pàg. 52).

<sup>6</sup> D'altra banda, la noció de cooperació perfecta de Knight també inclou una noció de cooperació -esquemes de cooperació-: tant el benefici propi que produeix l'intercanvi en el mercat, sense importar la persona amb qui es realitza, com pel que fa a la definició de les normes de la competició, evitant l'aparició de frau o imposició violenta en els intercanvis. (J. Coleman, "Competition..", art. cit., 1987, pàgs. 84-85).

<sup>7</sup> Per exemple, vid. J. Buchanan, "Las bases de la acción colectiva" (a J. Casahuga, (Ed.), Teoría económica y democracia, Instituto de Estudios Fiscales, Madrid, 1971).

d'aquesta perspectiva, es centren en l'anàlisi de les polítiques i regles de decisió de les institucions estatals, així com en els problemes que generen aquestes en intervenir sobre allò que és percebut com fallides del lliure mercat<sup>8</sup>. Aquesta perspectiva sobre l'acció col·lectiva no serà tractada al llarg de tot el capítol, el qual es concentra en examinar l'altra línia teòrica esmentada. El motiu és doble. Per una banda, l'enfocament de l'«elecció pública» contempla els problemes de l'acció col·lectiva de forma associada amb la intervenció estatal sobre el mercat; i d'altra banda, no discuteix centralment el problema de l'aparició de l'acció col·lectiva -amb o sense capacitat coercitiva-<sup>9</sup>.

L'aportació teòrica d'Olson parteix de la detecció d'una «anomalia» a la interpretació convencional de l'acció col·lectiva<sup>10</sup>. En absència de capacitat coercitiva, es suposava generalment que la provisió d'un bé col·lectiu es produïa quan existia una coincidència d'interessos entre un grup d'individus racionals. Però Olson, fent servir la teoria dels béns públics, va mostrar que en la major part de les ocasions la solució «interna» al problema de l'acció col·lectiva no era tan senzilla, i que era necessari incorporar solucions «externes». Aquesta «anomalia» es presenta en termes de la teoria de jocs a través del joc del dilema del presoner. En aquest joc, si l'individu es guia pel seu propi interès, i és consistent amb les seves preferències, aprofita les contribucions dels altres, sense participar (aquesta actitud s'anomena 'free-rider'). El problema detectat per Olson és que si tots els individus adopten aquesta actitud, la cooperació no es produeix, tot i que hi pugui haver un interès comú sobre els beneficis del bé col·lectiu.

---

<sup>8</sup>Per exemple, una perspectiva global -relativament crítica- sobre aquest corrent es pot trobar a V. Ostrom i E. Ostrom, "Public Choice: A Different Approach to the Study of Public Administration", Public Administration Review, vol. 31, 1971.

<sup>9</sup> Un dels aspectes més desenvolupats per aquesta corrent és el que fa referència a l'anàlisi de la votació (per exemple, D. Mueller, Elección pública, Alianza, Madrid, 1984).

<sup>10</sup> M. Olson, The Logic of..., op. cit., 1965.

Mancur Olson va connectar la interpretació de l'acció col·lectiva basada en la racionalitat individual amb alguns supòsits institucionalistes (com solucions «externes»). Proposant la influència d'incentius i coercions sobre un únic interès individual, la seva teoria assolí major capacitat predictiva i convertia la motivació en una variable independent, predeterminada. Junt a la provisió d'incentius individuals per tal de contrarestar l'efecte 'free-rider', també va introduir altres nocions, com les restriccions a l'acció, per a explicar el propi interès i la maximització dels guanys en la cooperació. Al mateix temps, Olson assenyalava que l'interès comú no proveeix cap incentiu per l'acció individual i que, per tant, una solució «interna» basada solament en la racionalitat individual tenia moltes dificultats per a formular explicacions consistents. Tot i incorporar alguns elements de caire institucionalista (el paper dels incentius i la concepció de l'economia política com a camp d'interacció entre actors col·lectius), la formulació d'Olson es mantenia bàsicament dins la perspectiva dominant de l'economia, integrant alguns elements de la teoria dels béns públics desenvolupada uns anys abans per P. Samuelson<sup>11</sup>:

*"The fundamental reality that unifies the theory of public goods with the more general logic of collective action is that ordinary market or voluntary action fails to obtain the objective in question"*<sup>12</sup>

La influència del model d'Olson va traspasar els àmbits de la teoria econòmica de la qual partia, per provocar, des de finals dels anys seixantes, una de les transformacions probablement més importants en la investigació social i política

---

<sup>11</sup> Especialment en dos articles: P. Samuelson, "The Pure Theory of Public Expenditure", Review of Economics and Statistics, vol. 36, 1954; i "Diagrammatic Exposition of a Theory of Public Expenditure", Review of Economics and Statistics, vol. 37, 1955. Una aproximació a l'estat actual de la teoria dels béns públics es pot trobar a W. Blumel et al., "The Theory of Public Goods. A Survey", Journal of Institutional and Theoretical Economics, vol. 142, 1986.

<sup>12</sup> M. Olson, "Collective Action", The New Palgrave. A Dictionary of Economics, vol. 1, 1987, pàg. 474.

contemporània sobre el comportament col·lectiu. Degut a la seva atracció explicativa, la lògica 'olsoniana' sobre la racionalitat de la participació va ser integrada amb els enfocaments tradicionals sobre organitzacions i moviments socials. Aquesta integració interparadigmàtica ha estat un punt continuat de conflictes interpretatius, on poques vegades es fan explícits tots els supòsits que intervenen<sup>13</sup>. Malgrat l'augment de la capacitat interpretativa assolida per Olson amb la introducció de supòsits de caire institucionalista, sovint les explicacions s'han mostrat ineficients en diverses investigacions experimentals i històriques, ja que de vegades funciona la cooperació tot i no existir cap mena d'incentius, i d'altres vegades no funciona tot i existir diversos incentius.

Al llarg d'aquest capítol s'assumeix generalment que els problemes més importants per explicar l'acció col·lectiva apareixen en les situacions no-centralitzades, a les quals no hi intervenen organitzacions per assegurar o facilitar la contribució. La resolució d'aquests problemes pot ser fàcilment aplicable posteriorment -en ésser aclarits- als fenòmens d'acció col·lectiva centralitzada<sup>14</sup>. Per tant, és sobre aquest pla que es desenvolupa la discussió. Primer s'analitza detalladament l'estructura teòrica de M. Olson i la justificació dels seus pressupòsits. Els següents apartats són dedicats a una discussió crítica dels models actuals sobre l'acció col·lectiva, inspirats en la perspectiva 'olsoniana', passats més de vint-i-cinc anys des de la publicació de la seva obra. Es discuteix la relació entre les dues lògiques centrals de l'acció col·lectiva: la individual, vinculada amb la natura del bé públic, i la interactiva, derivada de les necessitats de

---

<sup>13</sup> Les investigacions sociològiques de base empírica, han utilitzat o desestimat a conveniència aquesta hipòtesi, sense una major discussió teòrica. B. Keller recalca aquesta situació a "Olson's 'Logik des kollektiven Handelns'. Entwicklung, Kritik -und eine Alternative", *Politische Vierteljahresschrift*, vol. 29, 1988, pàg. 388.

<sup>14</sup> En aquest sentit argumenta J. Elster que "*solutions can be centralized or decentralized, depending on whether they require externally imposed force and inducements. Decentralized solutions are more basic than centralized ones, since to ensure compliance with a central institution is itself a collective action problem*" (Nuts and Bolts..., *op. cit.*, 1989, pàg. 131).

coordinació entre els individus, i es presenten detalladament propostes de solucions «internes» i «externes» als problemes de l'acció col·lectiva.

### 3.2. El model neoclàssic d'acció col·lectiva i les seves limitacions.

#### 3.2.1. Racionalitat econòmica i empelts institucionalistes en la teoria d'Olson sobre l'acció col·lectiva.

L'obra de Mancur Olson mostra clarament que la tensió entre les solucions «interna» i «externa» per explicar els processos d'acció col·lectiva ja era present des dels inicis de la seva formulació explícita en termes teòrics. La principal hipòtesi formulada per M. Olson sobre l'acció col·lectiva és que, si el nombre d'individus d'un col·lectiu amb un interès determinat en la producció d'un bé públic és força gran, el bé col·lectiu no es produirà a no ser que existeixin incentius selectius o coerció. L'argument destaca que l'interès individual d'aquells qui estan interessats en la producció del bé col·lectiu actua com a restricció estratègica per a la consecució d'aquest<sup>15</sup>. Olson explica perquè els grups grans tenen menys probabilitats d'èxit que els petits mitjançant dos factors. Destaca primer que en augmentar la dimensió d'un grup es redueixen els beneficis de la participació i augmenten els costos de l'organització (incloent els incentius selectius); en segon lloc, assenyala que en créixer el nombre de participants, la relació d'interdependència és cada cop més feble, i augmenta la probabilitat de les actituds no cooperatives.

A la teoria d'Olson es poden trobar influències de l'escola institucionalista americana. Tot i que la concepció de l'acció col·lectiva que aquesta escola va desenvolupar era completament distinta a la que aplicava l'explicació econòmica a la ciència política, cal destacar què, sorprenentment, la noció d'acció col·lectiva, com a

---

<sup>15</sup> R. Balme, "L'action collective rationnelle dans le paradigme d'Olson", *L'Année sociologique*, vol. 40, 1990, pàg. 263.

caracterització general d'una acció per proveir un bé públic no regulada per una estructura de mercat, sorgeix com a concepte analític des d'aquesta perspectiva.

*"Collective action proceeds, indeed, not from the intellectual logic of philosophers and economists, but from the arguments, debates, conferences, compromises, mass meetings, agreements, disagreements, negotiations, propaganda -among ordinary people themselves, like business men, labouring men, farmers, or professional classes, when forced or persuaded to consider their common interests. The psychology of this give-and-take process of conciliation and agreement may be named negotiational psychology, to distinguish it from the pleasure-pain psychology of the individualistic economists since the eighteenth century"*<sup>16</sup>.

La concepció de l'acció col·lectiva com a forma de comportament que ha d'ésser impulsat per l'Estat -en el cas de l'enfocament de l'«elecció pública»-, o la idea que és necessari introduir constriccions o incentius per facilitar l'acció col·lectiva voluntària -el cas d'Olson- tenen la seva arrel inicial en interpretacions institucionalistes<sup>17</sup>. La vinculació és interessant perquè s'introdueixen problemes teòrics, o elements explicatius en un cos interpretatiu gairebé contrari al que havia generat inicialment la discussió

---

<sup>16</sup> J. R. Commons, The Economics of Collective Action, University of Wisconsin Press, Madison, 1970 (primera ed. 1950), pàgs. 28-29.

<sup>17</sup> D'una forma força intuïtiva, anys abans Commons havia indicat ja aquesta necessitat: *"Collective action, as well as individual action, has always been there; but from Smith to the Twentieth Century it has been excluded or ignored, except as attacks on trade unions or as postscripts on ethics or public policy. The problem now is not to create a different kind of economics -'institutional' economics- divorced from preceding schools, but how to give collective action, in all its varieties, its due place throughout economic theory"* (J.R. Commons, Institutional Economics: Its place in Politics and Economy, University of Wisconsin Press, Madison, 1959 (primera ed. 1934), pàg. 5. Sobre Commons, vegeu entre d'altres: A.W. Coats, "John R. Commons as a Historian of Economics: The Quest for the Antecedents of Collective Action", a Research in the History of Economic Thought and Methodology, vol. 1, 1983. Sobre el desenvolupament posterior de les interpretacions institucionalistes en el camp de l'acció col·lectiva, vegeu E. Ostrom, Governing the Commons: The Evolution of Institutions for Collective Action, Cambridge University Press, Nova York, 1990.



sobre l'acció col·lectiva, i precisament aquesta dificultat d'integració real va provocar en bona part els problemes interpretatius que han complicat força el desenvolupament teòric posterior d'aquest enfocament. D'altra banda, el fenomen no és singular, ja que l'aplicació a nous problemes dels instruments analítics de l'economia neoclàssica és conseqüent amb la dinàmica expansiva d'aquest corrent al llarg de la segona meitat del segle vint. A la resta de l'apartat s'analitza amb més detall la síntesi teòrica realitzada per Olson en el seu llibre *The Logic of Collective Action*, que ha donat lloc a la formació d'un paradigma 'olsonià' sobre l'acció col·lectiva voluntària. El motiu és que, apart de ser una de les obres més significativa sobre els problemes de l'acció col·lectiva, és la que millor expressa les tensions existents en l'intent d'integrar aquests enfocaments contraposats.

Cal mencionar una limitació que Olson s'imposa des de l'inici del seu llibre: la seva interpretació es circumscriu a les organitzacions que suposadament busquen l'interès dels seus membres<sup>18</sup> (LCA, pàg. 6). Acceptant expressament en una nota que la lògica de la teoria també pot ésser aplicada a organitzacions comunals, religioses o filantròpiques, matisa que els resultats que pot aportar per aquests camps no són massa bons (LCA, pàg. 6, not. 6). La limitació és important perquè posteriorment aquesta interpretació s'ha aplicat freqüentment a grups en els quals aquesta premissa no era clara, i s'ha rebutjat o criticat seguidament la lògica de funcionament del grup, en no coincidir amb les prediccions 'olsonianes'.

Olson diferencia entre interessos individuals i interessos comuns. Per interès comú, entén un objectiu singular la consecució del qual es compartida per un conjunt

---

<sup>18</sup> Les referències al llibre de Mancur Olson, *The Logic of Collective Action*, molt nombroses en aquest apartat, són totes de la versió revisada i ampliada (amb un apèndix) que Olson va realitzar l'any 1971 (Harvard University Press, Cambridge). Per referir l'obra, tan sols es mencionen les inicials LCA, i el número de la pàgina.

d'individus, afegint que la recerca d'un interès comú és el que defineix principalment una organització, ja que els interessos particulars de cada individu es poden aconseguir millor actuant individualment. Quins són els tipus d'interessos que reconeix com efectius per l'acció col·lectiva és quelcom que resta bastant fosc en el seu llibre, però sí menciona que l'interès sobre l'objectiu comú pot ser de distinta intensitat segons les persones (LCA, pàg. 22). Els exemples que ofereix es refereixen constantment a interessos materials, bàsicament econòmics (o que poden derivar en beneficis econòmics), tot i que no afirma explícitament que sigui necessari referir-se sempre a aquests interessos. Olson proposa una analogia amb la noció del mercat perfectament competitiu per defensar la idea que no és racional la participació en els grans grups, malgrat l'existència d'un interès comú per tots els possibles participants:

*"In a large organization, the loss of one dues payer will not noticeable increase the burden for any other one dues payer, and so a rational person would not believe that if he were to withdraw from an organization he would drive others to do so"* (LCA, pàg. 12)

Restant importància als factors emocionals o ideològics que poden estar presents en una organització i fomentar la participació, Olson defensa la preeminència de l'interès propi. Per justificar la seva posició, destaca l'exemple de l'Estat nacional, com el cas més potent d'acció col·lectiva. L'organització estatal, segons Olson, té uns enormes recursos emocionals i ideològics, basats en el nacionalisme, però tot i així no pot evitar la necessitat d'imposar mesures de força per finançar-se i finançar la provisió de béns públics (LCA, pàgs. 12-14)<sup>19</sup>.

---

<sup>19</sup> És simptomàtic que hagi de recórrer a l'exemple, i no a arguments de principis o lògics. Aquesta tècnica és un indici de la introducció d'empelts institucionalistes en els plantejaments d'Olson, així com de la seva dificultat per integrar-los formalment.

*"And just a state cannot support itself by voluntary contributions, or by selling its basic services on the market, neither can other large organizations support themselves without providing some sanction, or some attraction distinct from the public good itself, that will lead individuals to help bear the burdens of maintaining the organization"* (LCA, pàgs. 15-16).

Seguidament exposa la tesi de que les organitzacions que no tenen la capacitat de fer obligatòria la participació d'un individu que comparteixi l'interès que defensa l'organització. Per tant, assenyala que aquestes han d'oferir alguns béns no-col·lectius perquè els membres potencials tinguin un incentiu per participar.

Olson distingeix entre béns públics exclusius i inclusius, tot considerant que els primers són propis de les situacions de mercat, i els segons de les situacions de no-mercat:

*"In non market situations the benefit from a collective good is not fixed in the supply (...) any number of people can join a lobbying organization without necessarily reducing the benefits for the others"* (LCA, pàg. 37).

La diferència és que, quan treuen profit dels béns exclusius més individus que aquells qui formen part del grup, es redueix el consum dels membres que hi contribueixen (degut a què la contribució individual afecta directament el volum del bé públic), mentre que els segons poden ésser també gaudits per membres addicionals sense que es redueixi el consum dels membres antics<sup>20</sup>. Finalment, Olson destaca que hi ha

---

<sup>20</sup> En tot cas, com el propi Olson reconeix (LCA, pàg. 38, not. 58), és més adequada la distinció de J. Head (a "Public Goods and Public Police", Public Finance, vol. 17, 1962) entre «impossibilitat d'exclusió» i «oferta conjunta» ('Jointness of supply'); que és la que ha estat generalitzada posteriorment. Restant la referència a inclusiu/exclusiu per caracteritzar el nivell de rivalitat en el seu consum (o discriminabilitat en l'oferta). Com a exemple de béns exclusius hom pot esmentar les carreteres -tant per l'ús com per la

grups, com els sindicats, que operen en les dues esferes, cosa que produeix actituds ambivalents respecte a la incorporació de nous membres al grup. Així, en relació als béns exclusius, la no participació d'un membre, si la natura del bé col·lectiu permet que s'aprofiti dels esforços dels altres, comporta de reducció dels beneficis d'aquells qui sí participen en l'acció col·lectiva<sup>21</sup>. Però si tots els membres poden efectuar aquesta estratègia, i es requereix la unanimitat, cal suposar que serà necessària molta negociació i, inclús així, la probabilitat d'èxit serà força més petita quan és necessària la unanimitat perquè els beneficis del grup es redueixen poc, tot i existir membres que no participen. Pel contrari, en els grups que proveeixen béns públics inclusivament aquests problemes tenen molta menys importància, ja que la participació unànime no és necessària normalment, i inclús la interacció estratègica entre els membres del grup pot ésser molt reduïda (LCA, pàg. 43).

Així, el problema central és analitzar com es constitueix la «massa crítica» d'individus que contribueixen a la provisió de bé col·lectiu. Paral·lelament apareix un tema important: quina és la relació existent entre la dimensió del grup i el comportament individual que respon al càlcul cost-benefici? Hom pot suposar que en grups petits les contribucions individuals sí són significatives, i això porta a un comportament diferent -més participatiu- pel que fa als membres d'aquests grups. Per respondre a aquesta qüestió, Olson assenyala que cada grup que busca un bé col·lectiu, afronta una particular funció de costos, segons les característiques del bé. La funció de costos total serà sempre creixent (si  $C$  són els costos, i  $T$  el nivell de provisió del bé, tenim  $C = f(T)$ ), i la funció de costos mitjans tindrà la típica forma d' $U$ . Així, en els

---

localització-, o el repartiment dels pressupostos de l'Estat; i com exemples de béns inclusivament, es pot destacar l'aire pur, la defensa d'un país o la llengua.

<sup>21</sup> Quan la reducció és molt gran, la participació unànime és absolutament necessària: "Any single holdout has extraordinary bargaining power; it may be able to demand for himself most of the gain that would come from any group-oriented action" (LCA, pàg. 41).

grups petits, si la relació entre costos i beneficis -en termes d'utilitat<sup>22</sup>- per obtenir el bé col·lectiu és favorable a l'individu, es produeix directament la participació degut a la seva atracció pels membres individuals. Buscant el punt òptim en la relació entre guanys i cost-benefici, Olson indica que aquest es produeix quan el benefici (pel grup) que implica l'existència del bé col·lectiu sobrepassa el cost total de provisió (pel grup) en major quantitat que el guany particular que n'obtenen un o més individus dins el grup (LCA, pàg. 34)<sup>23</sup>. Olson considera que per la natura dels béns col·lectius -no exclusivitat- en els grups petits hi ha una tendència a la sub-optimalitat en la seva provisió, que incrementa en augmentar el volum del grup, ja que els individus tenen la tendència a contribuir menys d'allò que seria necessari per la provisió òptima del bé.

Només quan apareix una important desigualtat en la intensitat de la motivació, pot ser més fàcil la provisió òptima del bé col·lectiu -perquè els membres amb més recursos insisteixen més- però en aquesta situació es pot produir una explotació d'aquests per part dels membres més petits, que contribueixen força menys d'allò que els pertoca, aprofitant-se del fort interès dels membres més potents. No obstant això, com més gran sigui el grup que es pot beneficiar d'un bé col·lectiu, menys probable serà l'existència d'una interacció oligopolística que ajudi a obtenir el bé, ja que en termes proporcionals s'haurà de mantenir una gran diferència d'interès.

---

<sup>22</sup> Com una mostra dels seus «equilibris» teòrics, es pot destacar que Olson parla de 'value', i no d'utilitat, tot i que la seva funció dins el model és idèntica a la de la utilitat subjectiva: "The total gain to the group will depend upon the rate or level at which the collective good is obtained (T), and the 'size' of the group (S), which depends not only upon the number of individuals in the group, but also on the value of a unit of the collective good to each individual in the group" (LCA, pàg. 23).

<sup>23</sup> Si és un bé col·lectiu pur (indivisibilitat d'oferta, no-rivalitat en el consum) no hi ha cap raó perquè el benefici individual disminueixi amb un augment de la grandària del grup. A més, la funció de costos no ha de ser necessàriament creixent en augmentar el nombre de membres del grup. B. Barry presenta una analogia a aquesta idea, destacant que es tracta d'una generalització del concepte de votar en termes de racionalitat econòmica. La recompensa és la diferència partidària multiplicada per la probabilitat de que un vot alteri els resultats. Si la contribució d'un individu és infinitesimal i no altera els resultats (la provisió del bé col·lectiu), no és racional fer l'esforç de contribuir (cf. B. Barry, *Los sociólogos..... op. cit.*, 1974).

Olson analitza tres tipus de grups, amb comportaments distints segons la seva grandària: a) grups petits (o grans) amb alguna contribució individual molt gran, b) grups petits amb contribucions significatives però sense cap de determinant, i c) grups grans amb contribucions individuals no perceptibles pel conjunt del grup. En el primer cas, afirma que el bé col·lectiu serà proveït, perquè el contribuent principal en trauria beneficis tot i havent de pagar tots els costos. En el segon cas, reconeix que el resultat resta indeterminat<sup>24</sup>. En el darrer cas, estableix la seva coneguda proposició segons la qual el bé col·lectiu no serà proveït a no ser que hi hagi coerció o estímuls externs que portin als membres del grup a actuar pel seu interès comú (LCA, pàg. 44)<sup>25</sup>. Sobre la relació entre bé col·lectiu i la necessitat de coordinació informal o organització formal, afirma que en els grups del tipus (a), és poc necessari el desenvolupament d'un acord o organització formal. També assenyalava que quan no es produeix la situació de gran desigualtat en la contribució en augmentar el grup, més necessària és l'organització per obtenir un bé col·lectiu. No cal que tot el grup que pot rebre el bé s'organitzi, però almenys una part d'aquest sí que ho ha de fer<sup>26</sup>. Però això implica uns costos addicionals per aquells que sí s'organitzen (comunicació, negociació, etc.) per la provisió d'un bé col·lectiu. Per tant, com més gran sigui la suma conjunta d'aquests costos per començar a proveir el bé, menys probable serà que es formi el grup sense

---

<sup>24</sup> Olson es refereix a la perspectiva històrica i als estudis experimentals dels psicòlegs socials per argumentar la «fàcil» possibilitat de cooperació: "*The fairly small (or intermediate) group has a fair chance that voluntary action will solve its collective problems*" (LCA, pàg. 58).

<sup>25</sup> Així, insisteix en l'argumentació: "*Whether a group will have the possibility of providing itself with a collective good without coercion or outside inducements therefore depends to a striking degree upon the number of individuals in the group, since the larger the group, the less the likelihood that the contribution of any one will be perceptible. It is not, however, strictly accurate to say that it depends solely on the number of individuals in the group.*" (LCA, pàg. 45). Cal destacar que aquesta afirmació d'Olson, sobre l'augment de la propensió a la no participació en augmentar el grup, no és integrada dins el seu model sobre el càlcul cost-benefici individual de la participació, constituint així dues explicacions paral·leles sobre el mateix fenomen.

<sup>26</sup> Si el grup és molt gran, pot ser més difícil aconseguir que un subgrup s'organitzi: "*The cost of organization are an increasing function of the number of individuals in the group*" (LCA, pàgs. 46-7).

coerció o incentius externs. Aquells grups que per no aconseguir les condicions esmentades no s'arriben a formar, són anomenats per Olson 'grups latents' (LCA, pàg. 50).

Precisament sobre els 'grups latents' Olson aplica la seva teoria dels incentius selectius, per destacar que aquests grups només s'activen quan poden ésser mobilitzats gràcies a l'ajuda d'incentius, positius o negatius:

*"Group action can be obtained only through an incentive that operates, not indiscriminately, like the collective good, upon the group as a whole, but rather selectively toward the individuals in the group. The incentive must be 'selective' so that those who do not join the organization working for the group's interest, or in other ways contribute to the attainment of the group's interest, can be treated differently from those who do". (LCA, pàg. 51).*

Olson planteja la seva anàlisi sobre els incentius selectius, considerant d'entrada la benèvola hipòtesi segons la qual existeix un perfecte consens en els grups que es poden formar (LCA, pàg. 60), ja que els problemes de consens -que no analitza- poden ésser un altre factor addicional que impedeixi la provisió de béns col·lectius<sup>27</sup>. Encara que per aquest autor els incentius per excel·lència són els que afecten als interessos monetaris o materials, també considera altres tipus d'incentius no econòmics (status, acceptació, etc.), que en la seva major part integra amb la denominació d'incentius socials, destacant el seu caràcter individual (no són béns col·lectius). Olson es recolza en interpretacions institucionalistes de les organitzacions, que han desenvolupat

---

<sup>27</sup> Aquí es pot apreciar una certa ingenuïtat d'Olson, en suposar implícitament que tot el que no és consens absolut pot anar en contra de les possibilitats de formació d'un grup. No seria del tot estrany considerar una hipòtesi completament contrària: la manca de consens pot afavorir el manteniment de l'acció col·lectiva, si s'accepten les veus discrepants (Hirschman, *Voz, salida y lealtad*, FCE, Mèxic, 1977). En tot cas, la qüestió és força més complexa que la linealitat que suposa Olson.

àmpliament el concepte d'incentiu<sup>28</sup>, per defensar que els incentius socials poden ésser considerats en el seu model igual que els incentius monetaris, tot i que finalment considera que la seva efectivitat resta reduïda als grups petits, als quals hi ha interacció entre els membres -o en els grups d'acció dins els grups grans, quan aquests estan dividits- (LCA, pàgs. 61-63)<sup>29</sup>. També menciona altres tipus d'incentius (eròtics, psicològics, morals,..) destacant que podrien ésser integrats igualment dins la seva proposta, ja que poden permetre discriminar entre els individus que donen suport a l'acció per proveir un bé col·lectiu, i aquells que no ho fan. Tantmateix, renuncia a considerar aquests incentius en el seu estudi per diverses raons: a) les dificultats de contrastació empírica de motivacions morals, o altres de les esmentades, b) perquè no ho necessita, ja que segons la seva opinió, amb els incentius econòmics -i addicionalment els socials- pot explicar-se tot tipus d'acció encaminada a aconseguir béns col·lectius en els grups d'interessos, c) perquè en els grans grups aquest factors són generalment residuals (LCA, pàgs. 61-2, not. 17).

Olson no manté la presumpció d'egoisme quan considera els incentius morals, que poden ésser integrats en el seu model. Però relativitza molt la seva importància, i només els aplicaria en un cas extrem als que sí participen, però no a tots aquells qui s'aprofiten del bé col·lectiu, els quals mantindrien el seu comportament egoista. Tot i la preeminència que Olson atorga al comportament egoista, Olson reconeix que aquest

---

<sup>28</sup> Començant per Ch. Barnard, *The Functions of the Executive*, Harvard University Press, Cambridge, 1938, i desenvolupat per P. Clark i J. Wilson a "Incentive Systems: A Theory of Organizations", *Administrative Science Quarterly*, vol. 6, 1961. Malgrat tot, Olson utilitza el terme com una noció global, sense analitzar la seva efectivitat per separat i conjuntament: "*Olson provides neither a theory nor a history of selective incentives. Polemically, one may say that this notion serves as a cover for a number of ad hoc explanations*" (L. Mjoset, "Review Essay. The Limits of Neoclassical Institutionalism", a *Journal of Peace Studies*, vol. 22, 1985, pàg. 80).

<sup>29</sup> Això sembla una contradicció, perquè en tots els grups acostuma a existir un cert nivell d'interacció entre els membres, a través de mitjans molt diversos, excepte en casos molt especials. No obstant això, en una nota Olson menciona (LCA, pàg. 63, not. 18) el paper dels mitjans de comunicació com una forma de proveir incentius socials (pressió/propaganda), tot i menysvalorant la seva eficàcia per proveir béns col·lectius.



no és una de les seves premisses:

*"The concept of the latent group (...) does not necessarily assume the selfish, profit-maximizing behavior that economists usually find in the marketplace. The concept of the large or latent group offered here holds true whether behavior is selfish or unselfish, so long as it is strictly speaking 'rational'" (LCA, pàg. 64).*

En relació als interessos, Olson té unes preferències analítiques molt clares: la gent cerca els interessos materials o econòmics. Però també suggereix la possibilitat d'integrar en el seu model altres motivacions, encara que, en el que hom podria inferir com una altra influència institucionalista, no presenta un model interpretatiu per integrar-les a la seva anàlisi. Per a Olson, és una qüestió de matisos, que cal discutir sobre casos concrets. Persuasivament intenta reduir al mínim la seva necessitat donant exemples concrets<sup>30</sup>. Segurament amb la pretensió que la capacitat predictiva -no tant explicativa- del seu model podria augmentar molt si la hipòtesi de racionalitat i una determinada motivació poden ésser presumiblement ja fixats d'entrada com d'aplicació generalitzada<sup>31</sup>.

---

<sup>30</sup> En discutir sobre els «lobbies» no-econòmics, Olson marca implícitament els límits de la seva teoria: *"As chapter II explained, the theory of large groups, at least, is not even limited to situations where there is self-interested behaviour, or where only monetary or material interests are at stake. Accordingly the generality of the theory is clear; on the other hand it is true that this theory, like any other theory, is less helpful in some cases than in others (...) it is evident that the theory sheds new light on some essentially social and political organizations (...), and that it is not especially useful in studying some other noneconomic lobbies"* (LCA, pàgs. 159-160) (es refereix com exemple als filantròpics o religiosos). Com resumeix D. Reisman, *"Olson's theory is not a general theory of human action (...) Olson's use of the self-interest construct is appropriate where it is appropriate but inappropriate where it is not"* (a *Theories of Collective Action*, MacMillan, Londres, 1990, pàg. 146).

<sup>31</sup> Referint-se a la limitació d'Olson sobre els incentius selectius de caire econòmic, Brian Barry destaca també: *"Esta restricción confiere a la teoría de Olson cierta capacidad positiva de predicción, pero en esta forma no parece capaz de explicar todos los fenómenos pertinentes"*, a *Los sociólogos...op.cit.*, 1974, pàg. 43. En aquest sentit, Olson és un defensor del principi *"That the theory that explains the most with the least is most likely to be true"* (cf. M. Olson, *The Rise and Decline of Nations*, Yale University Press, New Haven, 1982, pàg. 161).

Això no obstant, hi ha una altra raó: si en una teoria es consideren possibles tot tipus d'interessos, incentius, i orientacions, la seva aplicació directa pot caure molt fàcilment en una explicació tautològica, que consideri que tota acció col·lectiva observada ja té els ingredients adequats. Ara bé, com es mostra posteriorment, no cal plantejar sempre aplicacions directes de la teoria, aquesta pot limitar-se a definir impossibilitats i relacions entre les variables, deixant per la modelització dels fenòmens que s'analitzen la caracterització de les variables significatives.

Respecte a les motivacions individuals, Olson deixa oberta una possibilitat per admetre -en últim extrem- un cert marge de flexibilitat, però hi ha una condició sobre la que es mostra inflexible:

*"The only requirement is that the behavior of individuals in large groups or organizations of the kind considered should generally be rational, in the sense that their objectives, whether selfish or unselfish, should be pursued by means that are efficient and effective for achieving these objectives"* (LCA, pàgs. 64-65)<sup>32</sup>

Aquesta absència de flexibilitat és deguda a l'essencialitat del supòsit de racionalitat en la «conducció» de les motivacions per a mantenir la capacitat analítica del seu model. Si no fos així, suposa Olson, hom no deduiria que de la seva contribució en un grup no depèn el bé col·lectiu, i continuaria contribuint a la provisió d'un bé col·lectiu, convençut del valor de la seva participació<sup>33</sup>. Com destaca Reisman, el

---

<sup>32</sup> Cal mencionar que eficiència i eficàcia són també dos conceptes emprats i definits per Ch. Barnard, l'autor que inicialment va conceptualitzar la «lògica» dels incentius i la participació (vid. infra nota 27).

<sup>33</sup> Ara bé, això no implica que Olson consideri que el comportament individual sigui sempre racional. Referint-se als grups que defensen 'causes perdudes', menciona: *"There is probably less rationality, at least in the sense in which that world is used in economics, in noneconomic groups than in economic groups"* (LCA, pàg. 161, not. 93). La defensa d'Olson és simple, reconeix que el seu model no serveix per analitzar

debat sobre el principi de racionalitat en el model d'Olson s'ha de centrar en discutir en quins terrenys factics pot ésser útil la seva aplicació<sup>34</sup>. És precisament l'essencialitat d'aquest principi defensada per Olson un dels punts més importants que s'analitzen en el proper capítol<sup>35</sup>.

La tensió esmentada es pot observar en els darrers capítols de la seva obra, quan la influència de l'escola institucionalista americana apareix especialment. En aquests realitza anàlisis de casos concrets, com per exemple el seu tractament de les organitzacions sindicals (LCA, cap. 3). Amb un intens repàs de l'emergència i desenvolupament del sindicalisme nord-americà a través de la literatura institucionalista, mostra que l'aparició de grans organitzacions de treballadors amb capacitat de negociació col·lectiva ha estat sempre vinculat amb l'existència d'incentius selectius o, de forma més destacada, de mecanismes de coerció (destacant especialment el paper del 'closed-shop'). Paral·lelament, Olson argumenta precisament que el comportament dels treballadors ha estat racional en relació als elements definits pel seu model, i que sense coerció (i incentius), no hagués funcionat voluntàriament la negociació col·lectiva (LCA,

---

els grups que es caracteritzen per un baix nivell de racionalitat (LCA, pàg. 161)

<sup>34</sup> Concretament, assenyalava que "It is not about to include and exclude but about where to draw the line" (Cf. D. Reisman, *Theories of...* cit., 1990, pàg. 149. Per una banda, Reisman defensa l'auto-limitació d'Olson indicant que "the incorporation of any driving force but rationality of choice based upon personal self-interest tends to make the theory of collective action tautologous, non-testable, and useless for practical purposes" (*Ibidem.*, pàg. 159). D'altra banda, destaca que la motivació no pot ésser induïda directament de l'acció i molt difícil de determinar, i per tant el supòsit d'interès propi és el que pot cobrir un espectre més ampli de fenòmens.

<sup>35</sup> Olson defensa el concepte de racionalitat econòmica per la seva teoria assenyalant que "when all action-even charitable action- is defined or assumed to be rational, then this theory (or any other theory) becomes correct simply by virtue of its logical consistency, and is no longer capable of empirical refutation" (LCA, pàg. 160, n. 91). A més de ser una crítica a la pràctica de l'escola de Chicago, mostra que per Olson -no clarament, potser inclús contradictòriament- la racionalitat, en el sentit que la utilitza, ha d'ésser quelcom més que un conjunt de propietats lògiques.

pàgs. 86-87)<sup>36</sup>.

En altres capítols, que són de menor interès pels objectius d'aquest treball, Olson aplica la seva proposta teòrica per a criticar les teories tradicionals sobre l'Estat (destaca les perspectives institucionalista, econòmica i marxista) i les teories dels grups de pressió<sup>37</sup>. Finalment, dedica un darrer capítol a proposar una interpretació sobre els grups de pressió a partir de l'aplicació dels seus principis sobre la lògica de l'acció col·lectiva (LCA, cap. 6)<sup>38</sup>. En aquest capítol entra plenament, especialment en l'anàlisi de les relacions entre grups, en la interpretació influenciada per la visió institucionalista de la política (LCA, pàgs. 141-159), sense intentar aplicar una teoria econòmica de la política de tipus competitiu o maximitzador<sup>39</sup>. Per a Olson, la línia de la racionalitat no passa -segons es pot inferir indirectament- pel comportament estratègic de les organitzacions.

### 3.2.2. Cruïlles teòriques i ambigüitats en l'obra d'Olson.

Olson no aplica l'enfocament econòmic a les relacions entre organitzacions polítiques, però en canvi assumeix el supòsit de maximització o d'optimització en el seu

---

<sup>36</sup> En aquest sentit, Olson conclueix: "*The union member, like the individual taxpayer, has no incentive to sacrifice any more than he is forced to sacrifice*" (LCA, pàg. 91). Com és força evident, aquestes interpretacions són molt distants del funcionament dels sindicalismes europeus, els quals no presenten majoritàriament la institució del 'closed-shop'.

<sup>37</sup> En aquest capítol és central la discussió crítica de l'obra de Commons, de la qual extreu el paper institucional dels grups de pressió, criticant a aquest autor no haver descobert -sinó només insinuat- (The Economics of Collective..., op.cit., 1950, pàg. 213) la lògica de l'acció col·lectiva. Però el camp de discussió d'Olson -els grups de pressió i la seva lògica política-, és el camp tradicional del institucionalisme.

<sup>38</sup> L'argument d'Olson és que "*it will argue that the main types of large economic lobbies -the labor unions, the farm organizations, and the professional organizations- obtain their support mainly because they perform some function besides lobbying*" (LCA, pàg. 135).

<sup>39</sup> Com es desenvolupa, per exemple, a l'obra d'A. Downs, Una teoria econòmica de la democràcia, Aguilar, Madrid, 1973 (ed. or. 1957).

model d'individu i també en el seu model de grup. Destaca que els grups tenen una tendència a proveir sub-òptimament els béns col·lectius (LCA, pàg. 28), i que aquesta situació crea inclús oportunitats en els grans grups per l'aparició d'emprenedors, ja que la diferència en relació a l'òptim possible és molt gran (LCA, pàg. 177). Però hom podria suggerir que l'optimalitat en la provisió d'un bé col·lectiu es deriva de l'optimalitat de les contribucions individuals i, per tant, que els individus no es comporten òptimament. Tanmateix, per a Olson, són alhora el grup i l'individu els qui optimitzen, amb la diferència que tenen funcions d'elecció diferents. Pel grup, la seva funció és optimitzar la provisió d'un bé col·lectiu, per l'individu, la seva funció és optimitzar la seva contribució a la provisió del bé. El que no s'introdueix com a element de discussió, és si és possible transformar la funció d'optimització individual perquè sigui coincident amb la col·lectiva, o en tot cas, quines possibilitats hi ha perquè sigui més propera.

Com s'ha esmentat, Olson no considera explícitament el concepte d'utilitat en la seva obra<sup>40</sup>. Però si hom vol analitzar més formalment les hipòtesis que estableix, ha de fer ús del concepte d'utilitat, ja que la seva obra s'inserta en la tradició neoclàssica. En aquest sentit, tant el propi Olson en alguns articles dirigits a un públic més especialitzat<sup>41</sup>, com la major part dels continuadors de la seva obra, han hagut d'aplicar la utilitat com un instrument analític necessari pels seus models. A més de les

---

<sup>40</sup> Vid. nota 22. L'única menció és en una nota (LCA, pàg. 160, n. 91), per definir la noció de racionalitat aplicada a la filantropia, entenant aquí 'utilitat' com la satisfacció que proveeix actuar filantròpicament. Però seguidament indica que "*all the situations analyzed so far in this book require no such comprehensive and questionable definition of rationality*". Olson està més segur quan es refereix a interessos econòmics, i sembla inferir-se que el seu concepte de racionalitat econòmica s'associa a la persecució d'interessos que donin una satisfacció directa (però no es dedueix un rebuig del concepte d'utilitat, en aquest sentit restringit).

<sup>41</sup> Per exemple, a "An Economic Theory of Alliances" (amb R. Zeckhauser, *Review of Economic Studies*, vol. 48, 1966), s'utilitzen corbes d'indiferència sobre el valor de la defensa nacional. Vid. també M. Olson i R. Zeckhauser, "The Efficient Production of External Economies", *American Economic Review*, vol. 60, 1970; i M. Olson, "Microeconomic Incentives and Macroeconomic Decline", *Weltwirtschaftliches Archiv*, vol. 120, 1984.

dificultats per elaborar una aparell analític alternatiu amb què definir les preferències individuals, un motiu important d'aquesta dependència és que Olson no elabora cap interpretació o modelització de la lògica de la conducta individual. Tan sols pressuposa l'existència d'un interès, i d'una ordenació de preferències derivada, per establir les conseqüències col·lectives de les eleccions individuals, basant-se en la hipòtesi de la racionalitat.

El caràcter seminal de l'obra d'Olson fa comprensible l'existència d'algunes dificultats de coherència en el seu argument. Un exemple és la seva tècnica d'argumentació, informal, que li permet realitzar amb relativa facilitat alteracions en el significat de les variables que considera més importants. Al mateix temps, amb aquesta tècnica mantè un marge de maniobra per insistir en la potencialitat explicativa dels principis que inspiren la noció d'«homo oeconomicus», sense haver de definir completa i normativament les característiques d'aquest, i deixant sense resoldre, per tant, el problema de la seva plausibilitat.

Un aspecte rellevant de la interpretació d'Olson sobre l'acció col·lectiva que manca destacar, és la forma com enfoca la seva teoria. El principal objecte que Olson vol explicar no és el funcionament de les organitzacions voluntàries, o la lògica de la participació en aquestes, sinó el funcionament de la provisió de béns col·lectius, en relació a la dinàmica dels grups no estatals. Això és molt important per la crítica a Olson, que s'ha de referir centralment a aquest aspecte, o bé per prendre consciència de què si s'extrapola a altres terrenys la seva lògica, aquesta no és culpable completament dels problemes que hi apareixen<sup>42</sup>. Una possible causa de l'acumulació

---

<sup>42</sup> En un sentit similar apunta Reisman el perill d'utilitzar l'estructura d'Olson (basada en dos principis: racionalitat i interès propi) per "a whole kaleidoscope of combinations and permutations that is situated just out of reach of the two axioms and the three groups" (a *Theories of...*, 1990, pàg. 209), ja que es pot confondre erròniament la capacitat dels fonaments.

de problemes i confusions interpretatives en la lògica de l'acció col·lectiva pot ser deguda a què no sempre ha estat ben entès aquest punt de partida en el model d'Olson.

Malgrat els problemes de l'obra d'Olson, així com d'altres que s'exposaran seguidament, el principal valor del seu treball es la realització d'una síntesi teòrica. Partint d'un problema general -la formació de grups d'interessos- i d'un argument central -que la voluntat dels membres ofereix una explicació molt pobre-, Olson proposa una explicació alternativa, basada en la racionalitat individual, els interessos propis dels individus, i la natura dels seus objectius. A més, i aquest és potser el valor afegit més important de la seva obra, Olson intenta relacionar les diverses explicacions que proposa per aproximar-se a una teoria general sobre la lògica de l'acció col·lectiva<sup>43</sup>.

---

<sup>43</sup> L'article de F. Aguiar "La lógica de la cooperación" (a Zona Abierta, 54-55, 1990), apunta també en aquesta direcció (pàg. 15). Vid. també M. Groser, "Die Logik organisierter Interessen", Jahrbuch für Neue Politische Ökonomie, vol. 8, 1989.

### 3.3. Confrontacions per una teoria general de l'acció col·lectiva (1965-1990).

#### 3.3.1. L'aparició de diferents perspectives analítiques.

Els models que interpreten l'acció col·lectiva des del paradigma de l'elecció racional han tingut una evolució força intensa a partir de mitjans dels anys seixantes, quan Mancur Olson va formular la seva aproximació a una teoria general. Des de diferents enfocaments s'ha accentuat diversos dels components presents en l'obra d'Olson, com els incentius selectius, o la teoria dels béns públics, junt a la introducció d'alguns elements formals nous, com la teoria de jocs. Per l'elevada generalitat i dificultat de síntesi teòrica que implicava, així com per la diversitat de camps en què s'ha aplicat, les connexions entre aquestes diferents branques no s'han desenvolupat sistemàticament. I tal vegada són difícilment connectables entre si. És el propòsit del següent capítol discutir sobre tals connexions, en analitzar la utilitat de la noció d'interacció estratègica per explicar els fenòmens d'acció col·lectiva.

Entre els continuadors de la proposta teòrica d'Olson, hom hi pot trobar -en termes generals- dues grans perspectives. Per una banda, els autors que han tendit a integrar completament el model d'Olson dins l'estructura neoclàssica, criticant alguns aspectes però formalitzant completament la teoria, per integrar-la més detalladament en els models de decisió amb les distincions analítiques que proveeix la teoria econòmica (els seguidors de la solució «interna»). D'altra banda, apareixen les elaboracions que han intentat mantenir elements institucionals, rebutjant la formalització completa, i vinculant el concepte de racionalitat instrumental amb altres variables que fan força més complexes els models (les solucions «externes»). Dins aquestes dues orientacions generals, hi ha línies de treball força diverses, que es troben en un procés de discussió i crítica continuada: quelcom normal en un camp d'investigació teòrica en evolució.



Per a explicar el funcionament de l'acció col·lectiva cal referir-se prèviament als problemes que genera la producció i consum de béns públics, ja que s'entén que l'acció col·lectiva és generada per les necessitats individuals d'aquests béns. Segons les característiques específiques de cada bé col·lectiu, per exemple, el seu nivell d'exclusivitat, la seva funció de producció, el grau de rivalitat en el seu consum, etc., pot aparèixer una particular configuració que expliqui les possibilitats de portar a terme una acció col·lectiva. Com a bé públic, hom pot considerar de forma àmplia que es tracta de béns o altres productes -serveis, o inclús resultats polítics- que en una situació determinada no poden ésser produïts o consumits (o les dues coses a la vegada) de forma completament individual.

Apart de la noció de bé públic pur<sup>44</sup>, també es pot fer un ús més ampli del concepte de bé públic, relativitzant algunes de les proposicions que recull la seva definició original. Per exemple, potser que aquest bé públic sigui possible d'exclusió, però amb un cost elevat pels productors d'impedir que un no-contribuidor en faci ús; potser no sigui possible d'exclusió, però si sigui possible dificultar l'accés al seu consum. Segons com es produeixi el consum del bé públic, es poden introduir elements de selecció o direccionalitat en l'oferta que, sense trencar amb el seu caràcter públic, limitin el seu consum; o també pot aparèixer la rivalitat a partir d'un cert nivell de consum. Un altre aspecte important és la forma com es produeix el bé públic. Les possibilitats d'exclusió i d'oferta són molt distintes si -per destacar els dos extrems- el bé públic el pot produir una sola persona per tothom, o si hi han de contribuir necessàriament tots els possibles consumidors.

---

<sup>44</sup> Un bé públic pur implica les propietats d'oferta conjunta (el bé no es susceptible de ser gaudit individualment, per tant el seu consum no implica rivalitat) i impossibilitat d'exclusió (no es possible impedir l'accés a aquest bé). Sobre la teoria actual dels béns públics, vegeu R. Cornes i T. Sandler, The Theory of Externalities, Public Goods, and Club Goods, Cambridge University Press, Cambridge, 1986.

Junt a aquestes condicions, que impliquen en cada cas unes configuracions possibles de l'acció col·lectiva, cal suposar l'existència de determinades orientacions de comportament en els grups d'individus potencials consumidors del bé col·lectiu. Per exemple, l'interès propi i la racionalitat, són les orientacions suposades per Olson i les més convencionals; però també podrien ésser presents altres orientacions, o combinacions d'individus amb diverses orientacions. Segons quines orientacions, o combinacions d'orientacions, siguin presents, la formació de l'acció col·lectiva per la producció d'un determinat bé públic tindrà unes característiques específiques. La natura del bé públic no determina necessàriament les orientacions de comportament, però té una gran influència en la forma en què es produeix la interacció entre els individus per produir el bé públic; i en aquest sentit, cal afegir que en determinades condicions el propi procés d'interacció pot arribar a transformar les orientacions de comportament dels individus. En moltes de les configuracions possibles, els individus es troben en una situació d'elecció en que tot i què la provisió del bé públic és millor per tots, el bé no és proveït perquè els membres del grup prefereixen individualment no contribuir. Aquesta situació es coneguda com el problema bàsic de l'acció col·lectiva. Explicitada per M. Olson, diverses aportacions posteriors han matisat i formulat amb més claredat la qüestió. Segons quina sigui la funció de producció del bé col·lectiu, el problema bàsic de l'acció col·lectiva també pot ésser formulat en termes no dicotòmics, considerant la possibilitat de què tant sols es realitzi una provisió del bé col·lectiu en menor quantitat que la desitjada pels potencials consumidors o membres del grup.

En aquest capítol s'assumeixen alguns supòsits simplificadors, per a facilitar la presentació de les diferents perspectives teòriques que intenten solucionar aquest problema: es considera que cada individu té dues estratègies en relació a la provisió del bé públic: cooperar o no cooperar; tot i ésser poc plausible, es considera també que tots els individus són idèntics i, per tant, que tenen la mateixa estructura de preferències en relació al bé públic. Amb aquests supòsits, apareix dins el paradigma una opinió

compartida sobre el problema bàsic de l'acció col·lectiva. Aquest apareix en una elecció binària, quan tothom fa l'elecció no cooperativa, encara que el resultat global és inferior a una situació en què tothom hagués fet l'elecció cooperativa. El joc del dilema del presoner recull bona part d'aquestes situacions d'interacció, encara que no de forma exclusiva<sup>45</sup>. Es poden proposar també altres models en termes de teoria de jocs, segons les característiques del bé col·lectiu (vid. apt. 4.2.). D'altra banda, també poden existir situacions d'acció col·lectiva en les quals no es produeixi aquest problema, ja que d'entrada les eleccions individuals són cooperatives<sup>46</sup>. Aquest darrer cas no té menys interès, degut a què tals situacions il·lustren quines constel·lacions d'interessos, comportaments, condicions d'interacció i tipus de béns públics tenen més èxit per la formació i estabilitat d'accions col·lectives.

---

<sup>45</sup> A la literatura sobre acció col·lectiva, s'ha presentat sovint que el dilema del presoner amb un nombre  $N$  de jugadors era el model formal de la lògica de l'acció col·lectiva, en el sentit d'Olson (per exemple, R. Hardin, "Collective Action as an Agreeable N-Prisoners' Dilemma", *Behavioral Science*, vol. 16, 1971). Tot i l'elevada utilitat d'aquestes modelitzacions, cal destacar que només recullen un possible problema d'acció col·lectiva, pressuposant -sense explicitar normalment- unes determinades característiques del bé col·lectiu i del comportament dels individus.

<sup>46</sup> Estiguin o no basades en supòsits de comportament egoista. Mantenint el supòsits de comportament defensats per Olson, la cooperació es produiria quan el bé públic tingués suficient capacitat d'exclusió -o bé dèficit d'interès- cosa que permetria limitar als cooperadors el seu profit (es tractaria d'un joc d'assegurança), o quan el benefici individual fos prou elevat perquè alguns individus poguessin proveir el bé col·lectiu aïlladament (un joc del «gallina»). Aquestes dues situacions ja varen ésser contemplades per Olson com excepcions al seu model (vid. apt. 3.2.). Posteriorment han estat modelitzades a través de la teoria de jocs (cf. J. Hampton, "Free-Rider Problems in the Production of Collective Goods", *Economics and Philosophy*, vol. 3, 1987; i C. Ford Runge, "Institutions and the Free Rider: The Assurance Problem in Collective Action", *Journal of Politics*, vol. 46, 1984). Així, en formalitzar-se aquestes situacions, apareixen més clares les dependències amb característiques dels béns col·lectius. Un cop aclarides aquestes dependències, també es podrien pressuposar altres orientacions i comportaments amb els models de teoria de jocs (per un desenvolupament d'aquesta qüestió, vid. capítol quart). Un cas diferent és proposat per Elster, el qual introdueix la presència de processos de dissonància cognitiva en un fenomen d'acció col·lectiva, que condueixen a què tothom pensi que millora la situació col·lectiva tot i fent l'elecció no-cooperativa, quan en realitat no és així. En aquest sentit, com diu Elster, el dilema del presoner porta a una millora en el sentit de Pareto, en termes d'utilitat subjectiva. Depèn de la natura del bé col·lectiu i de l'estructura social que provoca la seva no provisió que aquests fenomen sigui un autèntic problema o no; en tot cas, la qüestió es vincula directament amb el paper de les creences dins la producció d'acció col·lectiva. No obstant això, la definició d'Elster sobre l'acció col·lectiva és quelcom contradictòria, perquè es basa en el supòsit d'interès propi, quan posteriorment rebutja l'aplicació d'aquest supòsit en moltes situacions específiques (*The Cement of...*, op. cit., 1989, pàgs. 24-49).

A partir del desenvolupament de la teoria de producció de béns públics, les dues gran perspectives esmentades anteriorment (que defensen respectivament les solucions «internes» i les «externes») han emfasitzat aspectes diversos. Hom pot destacar especialment tres aspectes sobre la investigació dels problemes de l'acció col·lectiva en els quals hi competeixen interpretacions alternatives. En primer lloc, l'observació i especificació de la natura d'una situació de no provisió de bé públic (o provisió parcial) quan existeix la demanda individual. Sobre aquesta qüestió es discuteixen les causes de la tensió entre objectius col·lectius i motius individuals. En el cas de la perspectiva que incorpora elements institucionalistes s'analitza la relació amb les possibles variables claus (coerció, incentius, etc.) que faciliten una acció col·lectiva<sup>47</sup>. En segon lloc, quan existeix l'acció col·lectiva, apareix un conflicte interpretatiu per explicar com es manté una situació de provisió de bé públic, ja que segons quins supòsits s'adoptin en el model interpretatiu, previsiblement aquesta provisió no es produiria<sup>48</sup>. En tercer lloc, els processos d'acció col·lectiva serveixen de focus per discutir i realitzar proposicions de

---

<sup>47</sup> En aquest sentit, es defineixen com intervencions de segon ordre les accions «polítiques» que s'implementen per estimular l'acció col·lectiva. Via organitzacions o lideratge individual, entre altres mètodes, existeix una àmplia col·lecció d'instruments útils per estimular l'acció col·lectiva, a utilitzar segons les característiques concretes d'aquesta. Olson, per augmentar de forma simple la capacitat deductiva de la seva teoria, restringia els instruments possibles gairebé a dos: els incentius selectius materials, i la coerció directa. Com es discuteix més endavant, la discussió sobre les intervencions de segon ordre pot desvincular-se en bona part de la lògica bàsica de l'acció col·lectiva. Més ben dit: segons els problemes de l'acció col·lectiva, o els objectius que hom es proposi per actuar sobre aquesta, l'elecció d'uns o altres instruments per intervenir-hi pot ésser més o menys eficaç. Però determinar perquè s'han emprat uns instruments concrets és un problema extern a l'explicació del funcionament de l'acció col·lectiva.

<sup>48</sup> Aquí es planteja un dels problemes teòrics més importants: com no existeixen -segons la perspectiva aquí adoptada- uns supòsits fixats prèviament (sobre el comportament, els béns públics, etc..) l'anàlisi de l'acció col·lectiva pot convertir-se en una completa tautologia, sense cap capacitat explicativa o predictiva. Així ho reconeix parcialment J. Elster, *"an implication of this approach which some may find depressing is that we will never have a general theory of collective action"* i defensa la necessitat de desenvolupar *"small and medium-sized mechanisms that apply across a wide spectrum of social situations"* (cf. *The Cement of...*, 1989, pàg. 205). Una investigació recent mostra que els individus no reaccionen igual en situacions d'acció col·lectiva idèntiques, i argumenten que cal buscar les relacions existents entre el supòsits de càlcul racional i les diverses identitats individuals (D. Goetze i P. Galderisi, *"Explaining Collective Action with Rational Models"*, *Public Choice*, vol. 62, 1989). Acceptant la impossibilitat d'una teoria general de l'acció col·lectiva, al següent capítol es planteja la possibilitat de desenvolupar teories parcials que relacionin comportaments i béns públics.

natura teòrica sobre les relacions entre el comportament interactiu i les constriccions materials (derivades de la natura del bé col·lectiu), així com amb les relacions socials<sup>49</sup>; aspectes plantejats de forma molt diferent des de cadascuna d'ambdues perspectives.

### 3.3.2. Acció col·lectiva i mobilització de recursos.

Abans realitzar el següent pas en l'anàlisi de la teoria de l'acció col·lectiva, cal destacar l'existència d'una tercera perspectiva d'investigació, no tractada directament en aquest treball, la qual recull els estudis que analitzen la dinàmica històrica de l'acció col·lectiva (per exemple, l'ocurrència i freqüència de successos en relació a determinades variables econòmiques o socials). Aquests estudis desenvolupen, entre altres elaboracions, contrastacions amb el model d'Olson. Coneguda com la teoria de la «mobilització de recursos», aquesta perspectiva es dedica especialment a la interpretació de la dinàmica històrica dels moviments socials, on els investigadors cerquen les relacions existents entre les característiques dels fenòmens col·lectius (durada, nombre de participants, resultats, etc.) i els seus entorns socials, per explicar

---

<sup>49</sup> Els vincles de la producció d'acció col·lectiva amb les estructures socials i econòmiques són escassament coneguts, però reconeguts com importants (vid. per exemple M. Granovetter, "Economic Action and..", art. cit., 1985). D'altra banda, el model d'home que presenta Olson està poc socialitzat, si es consideren aquests factors. Com un element constitutiu de l'estructura social, el paper de la comunicació es molt important. Per exemple, s'ha argumentat en aquest cas que la percepció de les contribucions i els efectes d'aquestes redueixen la incertesa sobre la productivitat de la participació, afavorint aquesta (cf. P. Oliver, G. Marwell i R. Teixeira, "A Theory of the Critical Mass I. Interdependence, Group Heterogeneity, and the Production of Collective Action", American Journal of Sociology, vol. 91, 1985). Un altre aspecte a considerar són les condicions físiques de la producció de l'acció col·lectiva: per exemple, la dispersió atomitzada (per exemple, els pagesos dispersos en masies i pobles petits) fa molt més difícil de combatre l'actitud de 'free-rider' que no la concentració en un espai concret (per exemple, els obrers situats en fàbriques i tallers): el mateix M. Olson ha tractat aquesta qüestió, que ja fou reconeguda per A. Smith (Cf. "Space, Organization and Agriculture", American Journal of Agricultural Economics, vol. 67, 1985); i de forma similar a les condicions físiques, la forma estructural com es relacionen els individus per l'acció col·lectiva. Les formes més centralitzades i les menys centralitzades són les que tenen menys vulnerabilitat a l'aparició del 'free-rider' (vid. Ph. Bonacich, "Communication Networks and Collective Action", Social Networks, vol. 9, 1987).

l'acció col·lectiva racional que té per objecte la participació política<sup>50</sup>.

Aquesta tercera perspectiva intenta combinar l'anàlisi del comportament estratègic dels grups i moviments socials amb la interpretació dels processos interns de formació i mobilització d'aquests conjunts d'individus. El sistema polític es el camp privilegiat en el que s'analitzen aquestes relacions. Per a tal fi, pressuposen que les organitzacions i els moviments integren, cohesionen i estimulen l'aportació de recursos individuals (temps, diners, etc.), i concebeixen la societat com un camp de recursos i conflictes polítics que evoluciona contínuament. No obstant això, la construcció d'una perspectiva teòrica unificada resta pendent dels problemes d'integració entre els supòsits instrumentals derivats de l'enfocament d'Olson i els supòsits referits al funcionament dels processos socials i polítics<sup>51</sup>. Les solucions que es proposen són heterogènies i contradictòries en força casos, basant-se -per exemple- en elements utilitaristes i simbòlics, sense integrar conceptualment les diverses fonamentacions teòriques de les quals són dependents. Des de la perspectiva de les aproximacions teòriques a l'acció col·lectiva esmentades anteriorment, s'argumenta que l'aplicació a camps específics, com pot ésser l'estudi dels moviments socials, es refereix a un nivell d'anàlisi distint, que ha de ser connectat -però no inferit directament- amb un pla teòric sobre l'acció

---

<sup>50</sup> Una visió general es pot trobar a C. Jenkins, "Resource Mobilization Theory and the Study of Social Movements", Annual Review of Sociology, vol. 9, 1983; J. Mac Carthy i M. Zald, "Resource Mobilization and Social Movements: A Partial Theory", American Journal of Sociology, vol. 82, 1977; R. Turner, "Collective Behavior and Resource Mobilization as Approaches to Social Movements: Issues and Continuities", Research in Social Movements, Conflict and Change, vol. 4, 1981; i D. Lapeyronnie, "Mouvements sociaux et action politique. Existe-t-il une théorie de la mobilisation des ressources?", Revue Française de Sociologie, vol. 29, 1988. Per a una revisió dels mètodes d'aquestes aproximacions, vegeu S. Olzak, "Analysis of Events in the Study of Collective Action", Annual Review of Sociology, vol. 15, 1989. Una de les obres clàssiques d'aquesta perspectiva és la de Ch. Tilly, From Mobilization to Revolution, Addison-Wesley, Reading, 1978.

<sup>51</sup> Com destaca D. Lapeyronnie, "L'affirmation de la nature rationnelle et instrumentale de l'action politique et de son caractère stratégique est difficilement compatible avec la prise en charge des problèmes posés par l'analyse des processus de mobilisation de groupes sociaux ou d'individus qui font intervenir des notions de revendications, d'engagements, d'idéologie ou de solidarité. (...) Les notions centrales qui servent à opérer le passage d'un niveau à l'autre sont de nature descriptive et n'ont pas de valeur analytique" (cf. "Mouvements sociaux..", art. cit, 1988, pàg. 615).

col·lectiva<sup>52</sup>.

R. Hardin proposa connectar els fenòmens d'acció col·lectiva que investiguen els historiadors amb l'anàlisi teòrica de l'acció col·lectiva mitjançant la distinció entre els fenòmens de «contribuir» i «actuar» conjuntament<sup>53</sup>. En el primer cas, en augmentar el nombre de participants es redueix l'estímul de contribuir, mentre que en el segon cas augmenta l'estímul, perquè es redueix el seu cost (o el seu risc). Així, els fenòmens històrics de mobilització<sup>54</sup> freqüentment analitzats des de la teoria de la «mobilització de recursos», podrien ésser interpretats a partir d'estructures d'acció col·lectiva on l'augment del nombre de participants provoca una reducció del cost de la participació, perquè l'únic recurs que s'aporta és la pròpia presència (com més gran es una manifestació o una revolta, majors són les possibilitats de triomfar, i en tot cas, el que és segur és l'augment de l'anonimat pels participants i, per tant, la reducció del risc de repressió) .

---

<sup>52</sup> Un intent de definir aquestes relacions es pot trobar a G. Marwell i P. Oliver, "Collective Action and Social Movements" (Research in Social Movements, Conflict and Change, vol. 7, 1984): "We think it is a mistake to try to equate collective action and social movements. We believe that social movements are most usefully understood as complex aggregates of collective actions or events, aggregates which meet certain criteria of scope and size" (pàg. 6). L'element central que proposen és el de «campanyes col·lectives»: conjunts d'activitats limitades en el temps i en l'espai i orientades al mateix objectiu. Per un major aprofundiment d'aquesta interpretació, vid. P. Oliver, "Bringing the Crowd Back In: The Non-Organizational Elements of Social Movements", Research in Social Movements, Conflict and Change, vol. 11, 1989.

<sup>53</sup> Així ho proposa R. Hardin, a "Acting Together, Contributing Together" (Rationality and Society, vol. 3, 1991).

<sup>54</sup> Hardin assenyala que el fenomen d'actuar conjuntament està històricament relacionat amb els grups socials que han estat mancats de recursos per desenvolupar acció col·lectiva en forma de contribucions materials. El creixement contemporani de les classes mitjanes ha produït que en l'actualitat s'observin amb força més freqüència mecanismes d'acció col·lectiva basats en la contribució de recursos (cf. R. Hardin, "Acting Together,..", art. cit., 1991).

### 3.4. Reconstruccions parcials i crítiques al model teòric d'Olson.

El model d'Olson sobre l'acció col·lectiva, despulat de matisos interpretatius, té com a tesi central l'afirmació que els grans grups tenen menys probabilitats de formar-se que els petits (considerant inclús els incentius positius o negatius com a factors correctors). Per aquesta tesi, Olson estableix al llarg del seu llibre dues grans línies d'argumentació. La primera, es centra en demostrar que el guany individual derivat de la participació en l'acció col·lectiva disminueix en augmentar el grup. La segona línia defensa que, amb independència de l'evolució dels guanys individuals, els efectes de l'increment del nombre d'individus del grup sobre la interacció entre els individus afavoreixen les actituds de 'free-rider', en el sentit de poder aprofitar el bé públic sense contribuir-hi. Seguidament, s'analitzen aquestes línies d'argumentació en relació amb els desenvolupaments posteriors per part de les dues perspectives teòriques esmentades anteriorment, amb el propòsit de destacar els seus respectius problemes.

#### 3.4.1. Individu i bé públic.

Per la primera línia d'argumentació Olson destaca dos factors bàsics. El primer, que un augment del grup provoca un augment més que proporcional dels costos de provisió del bé col·lectiu i, per tant, una reducció dels beneficis nets de la participació individual. Entre els costos, Olson hi inclou tant els costos directes de la provisió del bé públic, com els costos indirectes derivats de la formació i manteniment de l'acció col·lectiva, per exemple de la coordinació i de la provisió d'incentius selectius (LCA, pàgs. 22-25). Però molts crítics posteriors han coincidit en destacar que del model d'Olson no es pot deduir directament que es produeixi un descens dels beneficis nets de



l'individu en augmentar el grup<sup>55</sup>. Encara que Olson afirmi l'existència d'aquesta relació, en realitat és una qüestió a determinar empíricament, segons les característiques concretes que presenti el bé col·lectiu. En aquest sentit, s'ha criticat que Olson utilitza confusament les nocions de costos totals i costos mitjans, presentant, de forma excessivament determinista, la funció de costos mitjans amb forma d' $U$ , o sia, monòtonament creixent a partir d'un nombre determinat d'individus. No obstant això, la definició d'aquestes funcions de costos és essencialment una qüestió de natura empírica. Així, pot donar-se el cas que la corba tingui forma de  $S$ , o que la funció sigui sempre decreixent. Per tant, tal com destaca R. Balme, es pot afirmar que

*"La seule dimension des couts ne rend pas nécessairement l'action collective plus improbable dans les grands groupes"*<sup>56</sup>.

El segon factor bàsic que Olson destaca per defensar aquesta primera argumentació és que disminueix la fracció del bé col·lectiu que reben els participants en el grup quan augmenta aquest. Per tant, argumenta Olson, en créixer el nombre d'individus, també incrementa la sub-optimalitat en la provisió de béns col·lectius, estimulants la temptació individual d'aprofitar-se del consum del bé col·lectiu sense pagar, en reduir-se el benefici individual. Però aquesta relació discordant entre l'optimització individual i la provisió òptima del bé col·lectiu tampoc es pot deduir del

---

<sup>55</sup> Entre els crítics de l'estructura formal, es pot destacar especialment K. Kimura, "Large Groups and a Tendency towards Failure: A Critique of M. Olson's Model of Collective Action", Journal of Mathematical Sociology, vol. 40, 1989. També J. Chamberlin, "A Diagrammatic Exposition of the Logic of Collective Action", a Hooker, Leach i McClennen (eds.), Foundations and Applications of Decision Theory, vol. II, Reidel, Dordrecht, 1978.

<sup>56</sup> R. Balme, "L'action collective...", art. cit., 1990, pàgs. 265-6. També K. Kimura, "Large Groups..", art. cit., 1989, pàgs. 266-268. Si el cost del bé públic és estable, o incrementa menys que la contribució individual, aquesta es pot anar reduint segons augmenta el nombre de participants. En el fons és una qüestió de punts d'equilibri, amb una funció decreixent -o creixent- linial de costos mitjans, el punt d'equilibri és infinit o zero; en funcions no-linials més complexes que la de forma d' $U$ , llavors cal discriminar entre diversos equilibris possibles.

seu model. No es pot deduir perquè Olson no pot assenyalar quina forma tenen aquestes funcions d'optimització i com es relacionen amb la dimensió del grup, degut a què depenen de les característiques concretes de cada bé col·lectiu. L'única hipòtesi que hi ha al darrera és que el bé col·lectiu presenta rivalitat en el seu consum, i això implica el supòsit d'augment dels costos individuals en la contribució per la producció del bé col·lectiu quan el grup és més nombrós si es vol mantenir la mateixa fracció (però podria ésser que les diferències fossin infinitesimals, o inclús negatives)<sup>57</sup>.

Així, es pot concloure que la primera argumentació d'Olson, sobre la disminució del guany individual en augmentar el volum del grup, no té una capacitat d'aplicació general, sinó que depèn de la natura del bé públic que provoca una acció col·lectiva<sup>58</sup>. Aplicant aquest primer argument d'Olson sobre el cas dels béns col·lectius completament purs (oferta conjunta, no-rivalitat, no-exclusibilitat), es pot afirmar que amb un increment de la dimensió del grup el seu cost no augmenta, ni disminueix el consum individual i que, per tant, tampoc hi ha motius perquè disminueixi el benefici individual. En el cas dels béns col·lectius impurs (que els manca alguna de les propietats esmentades), segons les seves característiques concretes i al marge de les possibles actituds de 'free-rider', sí pot produir-se una certa dependència entre la participació i

---

<sup>57</sup> K. Kimura, "Large Groups..", art. cit., 1989, pàgs. 268-9. Aquests autor afirma que el descens o increment de l'òptim individual en augmentar el grup, depèn de: i) el comportament del rati del descens de la fracció del bé col·lectiu que aprofita un individu quan augmenta el nombre de participants, ii) del comportament del cost marginal del bé col·lectiu en augmentar el grup. Per exemple, si el cost marginal és constant i el rati és nul, l'òptim individual no canvia en créixer el grup.

<sup>58</sup> Per exemple, C. Bliss i B. Nalebuff presenten un model en el qual el bé públic pot ésser produït per l'esforç d'una sola persona (com obrir una finestra en una gran habitació plena de gent i calor), perquè els beneficis privats justifiquen el cost de l'acció. El seu argument és que "*the free rider problem will be greatest when there is only a limited pool of potential volunteers. Even in small groups, company is appreciated while waiting for a volunteer; more people in the same boat helps make everyone better off*" (a "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good", *Journal of Public Economics*, vol. 25, 1984, pàgs. 10-11).

la grandària del grup<sup>59</sup>. Addicionalment caldria matisar, des de la perspectiva més allunyada de les formulacions clàssiques, que segons les orientacions de comportament presents en la producció de l'acció col·lectiva, es poden alterar els càlculs de la relació cost-benefici en la funció d'utilitat<sup>60</sup>.

La major part dels béns públics que es produeixen mitjançant l'acció col·lectiva no són completament purs, i presenten un cert grau de separació en el seu consum i la producció, amb possibilitats d'imposar exclusions, i inclús pot aparèixer la rivalitat en el consum. Per tant, cada configuració concreta pot influir d'una determinada forma la probabilitat de participació; per exemple reduint aquesta si disminueix la fracció del bé col·lectiu que l'individu pot consumir o si augmenta el seu cost mig en créixer el nombre d'integrants del grup<sup>61</sup>. A partir d'aquestes possibilitats sobre les propietats dels béns col·lectius es pot ampliar la capacitat explicativa de la teoria de l'acció col·lectiva. Aquestes propietats no són d'individus, sinó de característiques físiques, tècniques, territorials, etc. -inclús de relacions socials-, la qual cosa permet tenir un punt d'escapada sobre la tautologia de basar totes les explicacions en supòsits sobre comportament individual. Seguint la perspectiva esmentada, és necessari afegir que cada bé col·lectiu té una determinada funció de satisfacció vinculada a la grandària del grup

---

<sup>59</sup> Sobre aquesta qüestió, vegeu J. Chamberlin, "Provision of Collective Goods as a Function of Group Size", *American Political Science Review*, vol. 68, 1974; i "A Diagrammatic Exposition...", art. cit., 1978. En tot cas, el que no es pot afirmar -a no ser que es determini prèviament la funció d'utilitat dels individus- és una relació de tipus general entre el nivell de suboptimalitat i el preu del bé col·lectiu, el canvi de riquesa dels individus o el canvi en la intensitat de les seves preferències.

<sup>60</sup> Per exemple, basant-se en un «principi de divisió», A. Leist proposa que un comportament altruista racional pot motivar una decisió individual de contribuir en un gran grup (a "Kollektive Güter und individuelle Verantwortung", *Analyse und Kritik*, vol. 11, 1989).

<sup>61</sup> Una organització pot oferir un conjunt molt ampli de béns col·lectius, no un de sol o principal, com suggereix Olson. La pròpia existència d'un mecanisme de coordinació és un bé col·lectiu. També hi ha la possibilitat d'oferir béns col·lectius diferenciats per sub-grups, junt a béns col·lectius generals, amb constel·lacions força complexes.

que el consumeix. Aquesta es pot entendre com la utilitat subjectiva que afecta a la utilitat directa del consum del bé. Per exemple, en una platja es gaudeix menys quan està plena que quan hi ha poca gent; però en un partit de futbol es a l'inrevés, ple dona més satisfacció que buit. Per aquest motiu, les percepcions subjectives, a més de l'estructura de costos i beneficis (o utilitat 'material' derivada de les propietats del bé públic) també tenen influència en la participació per la producció del bé col·lectiu.

Cal notar que en aquest apartat es suposa que l'individu reacciona racionalment front als resultats d'una funció d'optimització que té en compte les característiques concretes del bé públic. Però el supòsit de comportament racional és extern a l'anàlisi de les característiques del bé públic i a la funció que determina el benefici òptim de la participació. Per tant, també es podrien pressuposar altres formes de valorar els resultats de la funció, si hom considerés altres procediments de decisió no completament racionals. D'altra banda, diverses elaboracions, que s'allunyen de l'enfocament econòmic, mostren la possibilitat d'introduir alteracions en la percepció de les característiques del bé col·lectiu (per manca d'informació, per criteris de valoració subjectius<sup>62</sup>). El problema de no existència d'informació completa (sobre beneficis i costos, del tipus que siguin) en la decisió de participar, porta a l'aplicació de probabilitats, amb tots els seus problemes<sup>63</sup>. Com a alternativa, també es proposa que l'entrada en la acció col·lectiva sigui entesa com una estratègia de recerca d'informació,

---

<sup>62</sup> Dunleavy critica el supòsit d'informació perfecta dels actors utilitzat freqüentment en els models d'acció col·lectiva. Proposa un model de decisió sobre participació molt més complexa -amb elements objectius i subjectius-, en el qual la noció d'identitat col·lectiva és un element dinàmic clau (P. Dunleavy, "Group Identities and Individual Influence: Reconstructing the Theory of Interest Groups", British Journal of Political Science, vol. 18, 1988). Vid. també les seccions 3.2. i 3.6.

<sup>63</sup> Per exemple, la manca d'informació pot portar a que els individus calculin erròniament l'intercanvi entre contribució i incentius, oferint més del compte (T. Moe, The Organization of Interests, University of Chicago Press, Chicago, 1980, pàgs. 30-33). Aquest pot ésser un indicatiu de l'excés realisme de pressuposar que la gent utilitza correctament la teoria de la probabilitat (Cf. R. Boudon, L'Ideologie, op.cit., 1986). Sobre els problemes de l'aplicació de la teoria de Bayes en el context de la teoria de jocs, vegeu els apts. 2.2.1. i 2.2.4.

de forma similar a les teories de l'aprenentatge. No s'optimitza d'entrada, però aquesta és la tendència a llarg termini. Així, la revisió de les creences anteriors pot ser un pas cap a la sortida del grup, si està en la lògica de l'optimització<sup>64</sup>.

### 3.4.2. Interacció, grandària i les temptacions del 'free-rider'.

El procés d'interacció és el terreny de la segona línia argumental de Mancur Olson per defensar la tesi de què en augmentar el nombre de membres d'un grup es redueix la probabilitat que aquest tingui èxit en la provisió de béns públics. L'argument d'Olson és que en créixer el volum d'un grup, la freqüència de l'actitud de no col·laborar i aprofitar-se del bé públic, augmenta d'una forma més que proporcional. Olson presenta el seu argument en forma d'interpretació general, partint del supòsit previ que, sigui quina sigui la dimensió d'un grup, sempre hi ha latent una possible actitud de 'free-rider'.

El supòsit d'Olson s'ha criticat perquè només és defensable sempre i quan l'estructura de la interacció en un problema d'acció col·lectiva sigui la d'un dilema del presoner, que implica el supòsit de rivalitat en el consum<sup>65</sup>. Amb aquesta premissa, Olson afirma que en els grups petits és menys probable que el comportament dels individus segueixi els principis de la racionalitat individual, ja que la comunicació i el reconeixement directe entre els membres pot produir més fàcilment la cooperació. La dificultat de mantenir estructures fluides de comunicació que afavoreixin la confiança

---

<sup>64</sup> Aquest és l'argument de L. Rothenberg, "Organizational Maintenance and the Retention Decision in Groups", *American Political Science Review*, vol. 82, 1988, pàg. 1132. Això no obstant, una altra explicació podria ésser la voluntat dels individus per reduir la incertesa (vid. apt. 2.2.3.).

<sup>65</sup> El mecanisme només funciona pels béns «exclusius», en els quals amb un augment de la dimensió del grup es provoca un augment la rivalitat, si la producció del bé resta estancada (J. Chamberlin, "Provision of Collective Goods...", art. cit., 1974).

seria el motiu d'aquesta major probabilitat de no cooperació, quan el bé públic és ofert de forma conjunta, no exclusiva, i provoca una certa rivalitat en el consum<sup>66</sup>.

Només en relació als béns públics purs es podria aplicar la hipòtesi del 'free-rider' sense efectes indirectes, ja que pel que fa als béns col·lectius impurs, les eleccions entre cooperar o no cooperar són alterades per variacions en les relacions entre el volum del grup i costos del bé públic amb els beneficis que aporta la participació. Però degut a que en la major part dels casos reals cal tractar amb béns impurs, és molt important poder establir hipòtesis sobre les relacions entre el grup, el bé públic i els beneficis de la participació. D'altra banda, també hi ha força casos en que el consum de béns col·lectius purs requereix una producció conjunta pura (tot els consumidors han de ser alhora productors) no és susceptible de sofrir el 'free-riding' per individus racionals, ja que si no participa tothom, el bé no es pot produir<sup>67</sup> (per exemple, mantenir un secret: el dilema clàssic del silenci col·lectiu presentat en l'obra *El alcalde de Zalamea*).

Formalitzada posteriorment la noció d'Olson sobre la interacció mitjançant un dilema del presoner amb 'N' jugadors, s'ha presentat amb major claredat el procés de càlcul interactiu dels individus en relació a les característiques del bé públic. Però la teoria de jocs ofereix també la possibilitat de formular altres models d'interacció, si la

---

<sup>66</sup> Sobre el paper destacat de la comunicació, vid. A. Van de Kragt, J. Orbell i R. Dawes, "The Minimal Contribution Set as a Solution to Public Good Problems", *American Political Science Review*, vol. 77, 1983.

<sup>67</sup> Cf. M. Laver, "Political Solutions to the Collective Action Problem", *Political Studies*, vol. 28, 1980, pàg. 200. Una tipologia sobre la producció de béns públics que recull aquestes situacions la proposa J. Hirschleifer (a "From Weakest-Link to Best Shot: The Voluntary Provision of Public Goods", a *Public Choice*, vol. 41, 1983). Hirschleifer destaca tres formes de producció: i) la convencional, agregant les aportacions individuals; ii) el lligam dèbil, on la quantitat socialment disponible és el mínim de les quantitats individuals disponibles (per exemple, en una pressa d'aigua d'una comunitat, feta individualment per parts, la part més dèbil assenyala el nivell de bé públic que assoleix tota la comunitat); iii) la millor acció, la quantitat socialment disponible és el màxim de les quantitats individuals (per exemple, qui comunica més ràpid un missatge d'interès col·lectiu).

natura del bé públic o altres factors ho requereixen. Malgrat això, la teoria de jocs té problemes per incloure en el seu tractament formal el concepte d'Olson sobre la tendència decreixent a la cooperació en augmentar el nombre d'individus en un grup, perquè aquest concepte implica la introducció de nocions alienes al càlcul racional, o en tot cas, la consideració de què tercers factors afecten una elecció donada. Es tracta de la violació de l'axioma d'independència, assumpte ja discutit al primer capítol. Aquí es pot observar una de les incoherències d'Olson, ja que per un costat defensa la preeminència de la hipòtesi de racionalitat (apt. 3.2.1.), i d'altre, en la seva tesi més important introdueix de forma central el concepte de confiança, no basat en la racionalitat individual.

El problema del 'free-rider' ha estat intensament estudiat des de moltes perspectives. Olson argumenta que l'aparició d'aquesta actitud en la provisió voluntària de béns públics és més freqüent en els grups més nombrosos, perquè les possibilitats de comunicació directa hi són molt més reduïdes. L'argument d'Olson escapa al supòsit de racionalitat econòmica, per incorporar elements de psicologia social sobre la dinàmica dels grups petits. Entre d'altres, es pot esmentar que -contràriament als grans grups- existeix el coneixement comú de les contribucions (LCA, pàgs. 53-57)<sup>68</sup>. Aquesta formulació teòrica ha produït dificultats en els intents posteriors de modelitzar la lògica de l'acció col·lectiva a través de la teoria de jocs. A la vegada, l'argumentació d'Olson ha comportat que el fenomen del 'free-rider' fos analitzat atentament -de forma experimental i teòrica- des de perspectives no sempre fonamentades en els supòsits de

---

<sup>68</sup> Olson adopta l'enfoc de Homans sobre la dinàmica dels grups petits, destacant però, que aquesta no és adequada per als grans grups (G. Homans, *Social Behavior: Its Elementary Forms*, Harcourt, Brace and World, Nova York, 1961; i del mateix autor, "Fifty Years of Sociology", *Annual Review of Sociology*, vol. 12, 1986). Un intent de reincorporar en el model neoclàssic l'argument d'Olson es pot trobar a G. Stigler, "Free Riders and Collective Action: An Appendix to Theories of Economic Regulation", *Bell Journal of Economics*, vol. 5, 1974. Stigler argumenta que en un grup petit el possible 'free-rider' depèn més de la seva contribució per la provisió del bé col·lectiu i, per tant, més probable la seva participació; però si es suposa la racionalitat del comportament -com ho fa Stigler- la lògica del dilema del presoner invalida el seu argument.

racionalitat i interès propi.

La definició més general d'aquest problema es refereix al comportament no contributiu de béns públics produïts conjuntament, no susceptibles d'exclusibilitat i amb un cert grau de rivalitat en el consum. Però contràriament a l'afirmació d'Olson, moltes de les investigacions experimentals mostren que en circumstàncies força diverses, sense existir una oferta de béns privats (incentius selectius), també es poden produir contribucions per la producció de béns col·lectius en els grans grups que no tenen la capacitat d'excloure altres individus del consum del bé que produeixen<sup>69</sup>.

És precisament l'intent de formular una teorització d'aquests fenòmens observats empíricament el que ha portat a refinar la teoria de l'acció col·lectiva, tant pel que fa a la lògica de la provisió voluntària de béns públics, com a l'anàlisi del comportament individual i dels processos de decisió. En aquest sentit, es formulen anàlisis dels elements de decisió individual davant el bé públic i les actituds derivades dels processos d'interacció, per poder arribar a establir una interpretació més acurada sobre l'aparició del fenomen del 'free-rider'. Entre d'altres, un dels punts més destacats és la discussió

---

<sup>69</sup> Com destaca en la seva revisió B. Keller, "*individuen sehr wohl auch in großen Gruppen freiwillige Beiträge in nennenswertem Umfang leisten*" ("Olson's logik..", art. cit., 1988, pàg. 395). Hom pot destacar les investigacions de G. Marwell i R. Ames ("Experiments on the Provision of Public Goods I. Resources, Interest, Group Size, and the Free-Rider Problem", *American Journal of Sociology*, vol. 84, 1979; i "Economists Free Ride, does Anyone Else?: Experiments on the Provision of Public Goods IV.", *Journal of Public Economics*, vol. 15, 1981) i de W. Pommerehne i F. Scheider ("Wie steht's mit dem Trittbrettfahren? - Eine experimentelle Untersuchung", *Zeitschrift für die gesamte Staatswissenschaft*, vol. 136, 1980). Com a molt, s'observa una reducció del desig de cooperar quan augmenta el volum del grup, però sense indicis que les contribucions individuals tendeixin a zero (Cf. W. Stroebe i B. Frey, "Self-Interest and Collective Action: The Economics and Psychology of Public Goods", *British Journal of Social Psychology*, vol. 21, 1982, pàg. 131). Aquests autors, tot i defensar els predominis del principi de l'interès propi, reconeixen aquesta dificultat. L'argument de què és precisament sobre els grups que no existeixen -i que podrien existir (latents)- que es demostra l'efectivitat del 'free-rider' té el problema de tota argumentació contrafàctica. Si l'element de contrast en el món real no és inferior, la variable contrafàctica perd el seu sentit. Així, als defensors de la tesi del 'free-rider' no els queda altre argument que la crítica dels experiments realitzats, manifestant que són poc convincents (per exemple, T. McCaleb i R. Wagner, "The Experimental Search for Free Riders: Some Reflections and Observations", *Public Choice*, vol. 47, 1985).



sobre si mantenir una defensa estricta del supòsit d'interès propi, o bé acceptar la introducció moderada de supòsits sobre comportament no egoista i d'interès públic, en la qual es reflecteixen les orientacions de les distintes perspectives existents dins el paradigma de l'elecció racional<sup>70</sup>.

Un tema força discutit és el comportament que apareix quan la connexió entre la contribució i el consum d'un bé públic és més evident, reduint-se l'actitud de 'free-rider'. La percepció de l'efectivitat es comentada marginalment per Olson (LCA, pàg. 43-45), el qual també considera que aquesta actitud es redueix en augmentar la grandària del grup<sup>71</sup>. Tot i que és força acceptat que l'efecte de desmotivació provoca la pèrdua de control sobre el resultat de les accions pròpies, no està tan clar que el nombre d'individus del grup tingui una influència determinant en reduir la percepció de l'efectivitat, ja que depèn de les característiques del bé col·lectiu i les necessitats de

---

<sup>70</sup> Un sector defensa que l'interès propi també pot conduir a relacions cooperatives estables en determinades circumstàncies, que cal aclarir a través de models més complexos (per exemple, M. Crenson, "The Private Stake in Public Goods: Overcoming the Illogic of Collective Action", *Policy Sciences*, vol. 20, 1987). Altres autors destaquen que "*altruistic motives may reduce the likelihood and degree of free-riding, but they do not eliminate it*" (per exemple, R. Albanese i D. Van Fleet, "Rational Behavior in Groups: The Free-Riding Tendency", *Academy of Management Review*, vol. 10, 1986, pàg. 252); també afirmen que "*increases in group size lead to a decrease in the expected benefits of a contribution to the public good while, at the same time, decreasing the expected cost of free-riding*", augmentant finalment els beneficis nets del 'free-rider' en augmentar la grandària del grup (W. Stroebe i B. Frey, "Self-Interest and...", art. cit., 1982, pàg. 135). Però aquestes relacions depenen en gran part de les característiques del bé col·lectiu i, a més, tampoc està clara la relació entre beneficis i acció, ja que depèn de com es calculin els beneficis, i de com s'interpreti el model d'interacció. En aquest sentit, existeixen diverses propostes d'incorporar altres actituds en els models d'acció col·lectiva, modificant les funcions d'utilitat per incorporar satisfaccions no pròpies (vid. M. Taylor, *The Possibility of...*, op. cit., 1987, cap. 5; J. Elster, *The Cement of...*, op. cit., 1989, pàgs. 46-49). Per un tractament d'aquesta qüestió, en connexió amb la discussió sobre l'interès propi realitzada al primer capítol, vegeu la secció 4.3..

<sup>71</sup> Olson no distingia aquest efecte d'aquell altre que era produït per la publicitat de la contribució. La diferència, destacada per J. Sweeney, és que l'efecte de publicitat afecta la interacció entre els membres del grup, mentre que l'efecte de percepció només afecta al procés de càlcul i decisió individual (Cf. "Altruism, the Free Rider Problem and Group Size", *Theory and Decision*, vol. 4, 1974, pàgs. 259-61).

coordinació que imposa<sup>72</sup>, així com de l'entorn social que realitza la producció i el consum d'aquest. Així, és més plausible l'afirmació de Silver:

*"Appropriate contexts for participation are those in which strong causal connections tie the solitary acts of individuals with the prospect that a group will succeed in providing a desired good"*<sup>73</sup>

Una distinció important per a analitzar el fenomen del 'free-rider' és present entre el càlcul dels rendiments de la participació (la diferència del cost-benefici entre participar i no participar, en el sentit d'Olson) i el càlcul del cost d'oportunitat de la no participació<sup>74</sup>. En determinades circumstàncies, aquest darrer càlcul pot influir de forma determinant en una decisió de participar. Coneguda com la lògica de la «veu» i la «sortida»<sup>75</sup>, la seva aplicació a la producció de béns públics en forma col·lectiva no es contempla en el model d'Olson, però això no impedeix discutir les possibilitats de connexió. Si aquestes formes del càlcul de participar són contradictòries o

---

<sup>72</sup> L'argument de Sweeney és destacar que en augmentar la grandària del grup, la consecució d'un bé col·lectiu requereix una contribució col·lectiva tan gran en relació a la capacitat de contribució d'un subjecte individual, que aquest creu que la seva contribució no provoca cap diferència en la provisió del bé col·lectiu ("Altruism, the Free Rider..", art. cit., 1974, pàgs. 271-72. Això no obstant, aquesta situació pressuposa un bé col·lectiu que tendeix a ser perfectament divisible en la seva producció.

<sup>73</sup> Ch. Silver, "Utilitarian Participation" (Social Science Information, vol. 23, 1984, pàg. 725). En el mateix sentit argumenta Balme que *"en absence d'un mécanisme de coordination, la difficulté de perception des contributions change la nature du problème de l'action collective: alors que l'égoïste rationnel préfère le ticket gratuit s'il échappe aux sanctions, l'utilitariste est incapable d'évaluer l'efficacité de sa participation, et dans le doute peut être tenté de s'abstenir"* (Cf. "L'action collective..", art. cit., 1990, pàg. 275)

<sup>74</sup> Vid. B. Keller, "Olson 'Logik..", art. cit., 1988, pàg. 392.

<sup>75</sup> En el sentit que proposa A. Hirschman (a Voz, Salida..op. cit., 1977). Hirschman assenyala que quan hi ha altres organitzacions que proveeixen similar oferta, i no es difícil passar d'una a altra, el cost de la «sortida» és molt més reduït que quan no hi ha altres alternatives. A la vegada, aquesta darrera situació, o les dificultats per vincular-se amb altres organitzacions, estimula i facilita l'exercici de la «veu», o protesta i control a l'interior de les organitzacions. Per una recopilació del treball recent sobre aquesta teoria, vegeu del mateix A. Hirschman, "Salida y voz: una esfera de influencia en expansión", a Enfoques alternativos sobre la sociedad de mercado, FCE, Mèxic, 1989.

complementàries en una determinada acció, depèn en cada cas de les condicions estructurals del bé públic i de les formes organitzatives desenvolupades per la seva provisió. Tant la «veu» com la «sortida» tenen els seus propis costos, i aquests són diferents en cada cas<sup>76</sup>. En aquest context, es poden proposar models amb moltes variables i refinats càlculs per a determinar les dues funcions<sup>77</sup>. A més, no es pot afirmar que l'actitud de 'free-rider' sigui exclusiva del comportament racional, o que només es produeixi en estructures de dilema del presoner, encara que sigui en aquestes situacions quan és més freqüent la seva aparició (vid. la secció 4.2.).

Un darrer aspecte a destacar és la motivació orientada al procés de participació, i no als seus resultats, com s'ha entès fins ara. Per exemple, assistir a una manifestació política perquè té també els atractius d'una festa i ofereix diversió als participants, o implicar-se en una discussió perquè reforça la pròpia realització. Hirschman argumenta que aquestes motivacions poden transformar el càlcul de la participació, afegint-se a les motivacions pels resultats, tot convertint els costos també en beneficis. Per tant, des d'aquesta perspectiva seria molt menys probable l'actitud de 'free-rider' basada en una maximització de la funció de cost-benefici individual<sup>78</sup>. Però és difícil considerar que les motivacions pels processos transformin completament els costos en beneficis, excepte en casos patològics. Generalment poden contribuir a rebaixar aquests, encara que serien

---

<sup>76</sup> El cas dels sindicats és un exemple de què el cost de «sortida» és molt més gran si no hi ha cap organització alternativa en la qual integrar-se (i per tant menys probable) que si hi ha organitzacions en competència. Aquest procés reforça les formes de veu com formes de control individual (vid. B. Keller, "Olson's 'Logik..", art. cit., 1988, pàg. 393)

<sup>77</sup> Com per exemple fa J.-H. Hansen, "The Political Economy of Group Membership", American Political Science Review, vol. 79, 1985.

<sup>78</sup> Vid. A. Hirschman, Interés privado..., op. cit., 1986, pàgs. 88-102. Un tema connectat però no idèntic és el del canvi endogen de preferències. Amb el canvi de preferències pot variar la valoració de costos i beneficis, però cal recordar que el canvi és posterior a la decisió, i no anterior, com sembla deduir-se de Hirschman. Per una crítica d'Elster a l'argument de Hirschman, no excessivament encertada, vegeu The Cement of..., op. cit., 1989, pàgs. 44-46.

necessàries constel·lacions força concretes d'acció col·lectiva perquè aquestes transformacions es convertissin en especialment efectives, i no sols en elements secundaris. Hom pot considerar com un exemple dels efectes combinats de la motivació pel procés i per la finalitat el cas de les organitzacions polítiques que organitzen festivals musicals, les quals entren en competència amb els professionals del sector, generalment més experimentats. Malgrat tot, per entendre el possible èxit de l'organització política, cal contemplar que aquesta incorpora generalment altres finalitats (com uns objectius polítics compartits) que ajuden a explicar l'elecció participativa de bona part del públic.

### 3.4.3. Bé públic, interacció i càlcul individual.

Malgrat estar definida i elaborada la complexa relació entre càlcul individual i característiques del bé públic, el problema de la mútua connexió entre els dos àmbits d'anàlisi (càlcul individual i càlcul interactiu, en relació amb el bé col·lectiu) presenta encara algunes qüestions pendents<sup>79</sup>.

Olson desenvolupa el seu argument de forma no axiomàtica-deductiva, sense elaborar un model que relacioni el problema del 'free-rider' (la interdependència en la provisió de béns col·lectius) amb el guany o satisfacció individual (funció d'utilitat individual dependent del volum del grup). Tampoc estableix simultàniament una relació entre ambdues qüestions amb l'anàlisi de les propietats del bé públic (per exemple, el

---

<sup>79</sup> Quan la funció individual és negativa, ja no cal plantejar un model d'interacció, perquè la decisió és individual (un altre tema és la definició de la funció de decisió). El problema apareix quan la contribució és sub-òptima o inexistent, i la funció de decisió individual és positiva, en el cas hipotètic de contribuir òptimament. En el cas normal, hom pot aplicar com a segon nivell d'anàlisi la lògica interactiva en el procés de decisió per explicar el resultat de l'acció col·lectiva. Ara bé, hipotèticament, també es podria produir una situació en què el càlcul individual sobre la participació fos negatiu, i la interacció provoqués la participació.

grau de rivalitat pel seu consum, o el grau d'heterogeneïtat en el consum). S'han produït molts models que presenten l'argument d'Olson en termes de teoria de jocs, bàsicament utilitzant el joc del dilema del presoner amb 'N' jugadors, però generalment no integren de forma completa l'anàlisi de les propietats del bé col·lectiu<sup>80</sup>, i només suposen -igual que Olson- que existeix rivalitat en el seu consum. Un problema formal addicional és com influeix la funció d'utilitat individual o l'ordenació de preferències -variables segons la grandària del grup- a l'estructura d'interacció definida a partir del bé col·lectiu (considerant que el resultat de la interacció pot ésser també afectat durant el creixement del grup per altres variables). Aquesta és una qüestió important, especialment perquè també és un problema de caràcter computacional quan es volen aplicar aquests models<sup>81</sup>.

Les característiques dels béns col·lectius que implícitament es suposen en el model d'Olson, poden ésser transformades (per la qual cosa apareixen altres béns col·lectius) si hom utilitza altres interaccions en termes de teoria de jocs que no siguin el dilema del presoner<sup>82</sup>. Per exemple, en un cas que l'estructura de producció d'un bé col·lectiu sigui del tipus d'esglaó (per exemple, la construcció d'un pont), es

---

<sup>80</sup> Per exemple, R. Hardin, "Collective Action as...", art. cit., 1971. Aquestes modelitzacions, per necessitat de simplificació, suposen que els individus són homogenis i que només disposen de dues alternatives, contribuir o no contribuir; mentre que Olson pressuposa que les alternatives són infinites, ja que estan representades per la contribució -quantificada- que l'individu realitza (cf. K. Kimura, "Large Groups..", art. cit., 1989, pàg. 270).

<sup>81</sup> La modelització és excessivament complexa per a ésser tractada de forma plausible. Cal desenvolupar estratègies parcials de simplificació, que recullin les principals interdependències existents. També cal considerar que aquest càlcul pot estar basat en altres formes de decisió, o pot referir-se a diversos models de teoria de jocs. Els vincles entre la formulació d'un determinat model d'interacció i la relació individu-bé públic són complexes de definir. A més, en un problema de provisió de bé públic hi poden haver diversos tipus d'individus i, per tant, de forma associada, també diversos models d'interacció (per aquesta qüestió, vegeu el següent capítol).

<sup>82</sup> L'anàlisi de J. Hampton destaca que el grau de la rivalitat o no rivalitat en el consum depèn de l'estructura (o funció de producció) del bé col·lectiu, i si no existeix rivalitat, cal aplicar altres models d'interacció (cf. "Free-Rider Problems...", art. cit., 1987).

necessiten uns recursos exactes per produir el bé (amb menys recursos no es produeix, i més no són necessaris). Quan aquest bé públic no implica rivalitat en el consum, és possible que l'estructura d'interacció associada no sigui la del dilema del presoner, sinó la d'un joc de pura coordinació, com pot ésser el conegut com «la batalla dels sexes» (la coordinació pot ésser tant sobre qui paga, com sobre com es reparteixen els pagaments)<sup>83</sup>.

El supòsit d'igualtat de tots els actors en un procés d'acció col·lectiva és clarament implausible, i si s'ha mantingut prèviament ha estat degut a les necessitats de simplificació per distingir altres aspectes de la teoria. Amb tota probabilitat es pot afirmar que les motivacions mixtes, la coincidència d'individus que participen per diverses raons, estan al darrera de la major part dels casos d'acció col·lectiva que es produeixen realment<sup>84</sup>. Per aquest motiu, un dels principals problemes pel desenvolupament teòric sobre l'acció col·lectiva és com analitzar les situacions en què, per exemple, unes motivacions actuen com a detonadores d'altres, que a la vegada funcionen com a mecanisme per tornar a amplificar el nombre de les primeres i dels seus efectes secundaris.

---

<sup>83</sup> En el cas dels béns públics incrementals (l'oferta del bé augmenta amb la contribució, més o menys contínuament), l'estructura de producció del bé implica un dilema del presoner, però es tracta d'un conjunt de possibles dilemes, segons la forma concreta de la funció de producció: *"It is also possible that at certain levels of an incremental good's production (where a certain number of others in the group cooperate to produce this increments) the increment of benefit produced for that individual really does exceed the cost of that increment to her (Indeed, this may be true at all levels)"* (Cf. J. Hampton, "Free-Rider Problems..", art. cit., 1987, pàg. 267). Quan es produeix aquesta situació, la interacció ja no és la d'un dilema del presoner, ja que contribuir és preferit a no contribuir, encara que els altres no contribueixin (pot aparèixer un problema de coordinació per escollir els que han de contribuir).

<sup>84</sup> Vid. J. Elster, *The Cement of... op. cit.*, 1989, cap. 5. El tema dels incentius selectius, que es poden combinar de forma específica i discriminatòria amb les diferents motivacions existents entre els individus interessats en un bé col·lectiu, per reforçar-les (o crear-les); implica un problema analític posterior a la interacció entre diverses motivacions: la formació i manteniment de l'estructura de provisió dels incentius. En aquest sentit argumenta J. Elster: *"The provision of selective incentives cannot be the general solution to the collective action problem. To assume that there is a central authority offering incentives often requires another collective action problem to have been solved already. (...) To assume that the incentives are offered in a decentralized way, by mutual monitoring, gives rise to a second-order free-rider problem"* (*The Cement of... op.cit.*, 1989, pàgs. 40-41).

Sense entrar plenament a presentar les relacions entre preferències, motivacions, interessos i formes de comportament que es produeixen en el marc de les interaccions existents en l'acció col·lectiva, s'introdueix una formulació recent en aquest sentit, per il·lustrar algunes de les possibilitats analítiques existents. La teoria de la «massa crítica», desenvolupada per principalment G. Marwell i P. Oliver<sup>85</sup>, es centra en investigar els conjunts necessaris de participants perquè funcioni una acció col·lectiva, la probabilitat de la seva aparició, així com l'extensió i efectivitat de les seves accions. Pressuposen d'entrada l'existència d'una heterogeneïtat d'interessos i recursos en la població, amb una distribució que depèn de la funció de producció de bé col·lectiu, variable segons les característiques d'aquest. Junt a l'anàlisi de les decisions aïllades sobre les contribucions individuals, assumeixen la interdependència com la influència en aquestes decisions de les decisions prèvies dels altres sobre la contribució. Inspirant-se en simulacions per ordinador, els autors defensen que en força funcions de producció pot aparèixer un subgrup -que concentra els individus amb major intensitat de recursos i interès- que escull fer grans contribucions a l'acció col·lectiva, mentre la majoria fa poca contribució (encara que tampoc adopta l'actitud de 'free-rider', ja que en molts casos la possibilitat de benefici és molt reduïda). L'activitat d'aquest subgrup, o «massa crítica», pot desenvolupar una dinàmica que arrastri els membres menys interessats fins a la provisió òptima del bé col·lectiu. Sobre la relació amb la grandària del grup, segueixen el criteri -ja senyalat- que depèn de la funció de producció del bé col·lectiu. Això no obstant, introdueixen el paper de l'organització social vinculat amb la qüestió de la grandària, per a senyalar que un grup amb un elevat nombre de participants pot tenir menys problemes per reunir la «massa crítica», perquè representa un menor percentatge sobre

---

<sup>85</sup> P. Oliver, G. Marwell i R. Teixeira, "A Theory of the Critical Mass I...", art. cit., 1985; P. Oliver i G. Marwell, "The Paradox of Group Size in Collective Action: A Theory of the Critical Mass. II", American Sociological Review, vol. 53, 1988; i G. Marwell, P. Oliver i P. Prah, "Social Networks and Collective Action. A Theory of the Critical Mass. III", American Journal of Sociology, vol. 94, 1988. Aquesta teoria comparteix els seus supòsits addicionals amb els supòsits habituals de l'economia: es suposa que els membres disposen d'informació completa, que les seves decisions segueixen el criteri de maximització de la utilitat esperada.

el total potencial. Tanmateix, també destaquen que la relació social entre els membres de la «massa crítica» és allò que importa veritablement per a la realització d'acció col·lectiva: com més dispersa estigui -problema freqüent en un gran grup-, és més difícil la formació d'aquesta «massa».



### 3.5. Les intervencions de segon ordre en l'acció col·lectiva.

La possibilitat d'aplicar sancions és una altra via per resoldre el problema de l'acció col·lectiva. Però com els incentius i l'activitat de lideratge, la seva explicació constitueix un problema de segon ordre, en el sentit que impliquen l'existència d'algú que realitzi la provisió del bé col·lectiu 'sanció'<sup>86</sup>. Des de la perspectiva d'una solució «externa», no hi ha problema per suposar l'aparició d'elements institucionals, com si fossin variables exògenes al model, per definir la gènesi de les intervencions de segon ordre. Per a la perspectiva més vinculada a les solucions «internes», l'objectiu és poder explicar de forma endògena, tot basant-se gairebé exclusivament en el supòsit de racionalitat, l'aparició de les intervencions de segon ordre en l'acció col·lectiva.

Per ambdós casos, encara que formulada de distinta forma i amb recursos instrumentals diferents, la qüestió més important és com aconseguir formular explicacions sobre aquestes intervencions de segon ordre a partir de les eleccions individuals, cosa que es manté implícita en el paradigma compartit. L'alternativa de deixar aquesta connexió en forma de caixa negra, possibilitant la incorporació de postulats de caire institucionalista, com ho fa la primera perspectiva, deixa molts interrogants oberts. D'altra banda, elaborar models que necessiten uns elevats requisits informacionals és una alternativa que tampoc és força més prometedora, degut la manca de plausibilitat d'aquests requisits. En tot cas, aquestes intervencions són fàcilment observables en situacions quotidianes d'acció col·lectiva, i les dificultats de la seva integració teòrica, especialment pel que es refereix a la seva aparició, es mostren com un dels grans problemes pendents en l'anàlisi de l'acció col·lectiva.

---

<sup>86</sup> Tot i no aprofundir en la qüestió, cal previndre contra l'efecte 'Robinson Crusoe'. Quan algú genera una intervenció de segon ordre, sigui del tipus que sigui, apart d'afectar intencionalment als receptors de la seva intervenció, també resta afectat personalment. En determinades circumstàncies, aquest efecte de «tercer ordre» pot tenir una repercusió destacada en el procés de producció de l'acció col·lectiva (vid. G. Tsebelis, "Are Sanctions Effective? A Game-Theoretic Analysis", *Journal of Conflict Resolution*, vol. 34, 1990).

### 3.5.1. Sancions com a garantia.

Les formes d'aparició de les sancions són molt diverses. Poden ésser imposades pels propis individus que interaccionen de forma conjunta; per una de les parts front a la resta de membres del grup, o bé externament, per una autoritat amb suficient capacitat coercitiva (sancions socials, sancions personals o sancions pels altres, sancions legals). L'aparició de les normes socials, com a mecanismes endògens per a sancionar és un dels aspectes més complexos de la dinàmica de l'ordre social<sup>87</sup>, i afecta directament la lògica de l'acció col·lectiva. Un cas a destacar seria la imposició de sancions prèvies, o l'autolimitació de l'elecció, que es pot entendre com una variant especial del fenomen general d'articulació de sancions, en el qual opera de forma determinant l'asimetria temporal. Així, la coerció en el sentit d'Olson seria un cas concret dins aquest conjunt de possibilitats<sup>88</sup>. Els costos de les sancions poden variar en cada cas; i en algunes circumstàncies poden ésser superiors als beneficis que produiria la cooperació.

El paper de les sancions, especialment quan està vinculat amb l'aparició i persistència de normes, té una gran importància per a explicar una bona part dels fenòmens d'acció col·lectiva descentralitzada<sup>89</sup>. Però aquest és un tema que queda al

---

<sup>87</sup> J. Elster, *The Cement of...*, op. cit., 1989, i J. Coleman, *Foundations of Social Theory*, Harvard University Press, Cambridge, 1990.

<sup>88</sup> D. Reisman, *Theories of...* op. cit., 1990, pàg. 198.

<sup>89</sup> Un exemple interessant és el de la moralitat en una societat. Evidentment, el manteniment d'aquesta està vinculat amb un sistema endogen de sancions molt refinat. Així, l'existència d'una certa moralitat personal en una societat comporta grans beneficis per aquesta (vid. per exemple D. Reisman, *Theories of...* op. cit., 1990, pàg. 177; J. Elster, "Rationality, Morality and Collective Action", *Ethics*, vol. 96, 1985; i T. Schelling, *Choice and...* op. cit., 1984, pàg. 55). Un aspecte que es podria anomenar com a «subproducte de tercer ordre», però de rellevància especial en determinats contextos, com poden ésser els que estan dominats per fortes «cultures organitzacionals», és quan la pròpia norma es converteix en un incentiu especial, donant una satisfacció pròpia el seu compliment, degut a la intensa identificació de l'individu amb l'entorn (vid. W. Ouchi, *Teoria Z*, Editorial Orbis, Barcelona, 1985).

marge de la present discussió, i solament s'incorpora molt parcialment en la discussió dels propers apartats sobre els processos d'interacció en l'acció col·lectiva que segueixen un comportament guiat per normes.

### 3.5.2. Incentius com a foment.

La noció d'incentius selectius ha estat central en l'argumentació de que en els grups grans l'oferta de béns privats era l'única possibilitat de motivar la contribució voluntària per assolir un bé públic. Es suposa que aquests motiven la participació, ja que fan augmentar els beneficis d'aquesta, fins que la relació cost-benefici de participar sigui superior a la de no participar. Per a tal fi, han de poder ésser dirigits individualment per a alterar la funció de decisió de cada individu que pot consumir el bé públic. La principal crítica rebuda per aquest concepte és que els incentius selectius tenen un caràcter exògen al model d'acció col·lectiva que relaciona béns públics, racionalitat individual i interacció. En aquest sentit, existeix un problema teòric per a definir els incentius selectius, degut a que hom ha d'explicar l'origen d'aquests béns. La seva existència implica que prèviament una organització o un empresari controla i dirigeix la seva provisió, la qual cosa representa un problema addicional en l'explicació de l'acció col·lectiva, perquè l'associació o iniciativa d'aplicar els incentius selectius es pot considerar també com un bé col·lectiu. Però aquesta és una solució exògena al problema bàsic de l'acció col·lectiva, més útil per explicar la seva permanència, a través d'instruments de 'segon ordre' (que controlen les possibles estratègies 'free-rider' dels consumidors del bé públic) que per explicar els seus orígens<sup>90</sup>. Tanmateix, resta també explicar l'aparició d'acció col·lectiva descentralitzada, no organitzada, ja que aquesta pot no tenir capacitat d'aplicar incentius selectius. Una darrera possibilitat analítica seria

---

<sup>90</sup> Elster, "Rationality, Morality...", art. cit., 1985.

entendre l'aparició d'incentius selectius com a subproducte de la pròpia acció col·lectiva; però en tot cas, aquestes situacions serien força específiques dins l'ampli espectre de possibilitats que ofereix la lògica de l'acció col·lectiva<sup>91</sup>.

En el seu llibre, Olson discuteix sobre la natura dels incentius selectius, per a defensar que són bàsicament els incentius de natura econòmica els que poden atraure l'interès propi dels individus en un gran nombre de situacions. Olson defensa que amb aquest supòsit més restringit, la seva teoria assoleix major capacitat predictiva (vid. apt. 3.2.1.). D'altra banda, la perspectiva sociològica, sense excessives preocupacions formals, ha desenvolupat un extensa fenomenologia experimental del paper dels incentius selectius en els processos d'acció col·lectiva. Distingint entre incentius positius i negatius (els primers recompensen la contribució, i els segons penalitzen la no-contribució), es destaca per exemple que els incentius positius afavoreixen la participació d'un nombre reduït de membres del grup, mentre que els negatius tendeixen a afavorir una cooperació més general (sobretot si existeix la capacitat d'exclusió), però no són de tanta eficàcia per assegurar la estabilitat<sup>92</sup>.

Existeix força literatura sobre incentius organitzacionals que mostra una gran acumulació de dades amb refutacions a l'èmfasi d'Olson en què l'oferta de béns selectius (o privats) és essencial per l'acció col·lectiva, i que entre aquests, els incentius de caràcter econòmic són els més determinants. En conjunt, els resultats de les investigacions sociològiques sobre els incentius -de caire empíric- tenen un caràcter

---

<sup>91</sup> Una aportació que intenta trencar aquesta relació exògena és el treball de D. Heackthorn ("Collective Sanctions and the Creation of Prisoner's Dilemma Norms", American Journal of Sociology, vol. 94, 1988). Aquest intenta incorporar dins un model de teoria de jocs la provisió de sancions col·lectives (com un tipus específic de incentius col·lectius) com una variable endògena.

<sup>92</sup> P. Oliver, "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations", American Journal of Sociology, vol. 85, 1980; i de la mateixa autora, "Rewards and Punishments as Selective Incentives. An Apex Game", Journal of Conflict Resolution, vol. 28, 1984.

ambivalent, i mostren les seves dificultats per a integrar teòricament en models d'origen 'olsonià' (basats en la capacitat predictiva que faciliten uns pocs supòsits fixos de comportament) la diversitat de motivacions que s'observen en les anàlisis empíriques<sup>93</sup>. Així, es destaquen en aquestes investigacions diversos resultats, força inconnexes: que els incentius afecten, però no determinen la participació en l'acció col·lectiva, que les normes socials poden funcionar també com a incentius -sancionant el 'free-rider' o premiant la participació-, que de vegades els incentius selectius no serveixen per desenvolupar una acció col·lectiva, amb o sense organització formal al darrera.

La resposta a aquesta situació des de la perspectiva sociològica ha estat formular models que recullin l'amplia diversitat d'incentius, tot relacionant aquesta amb una possible diversitat de motivacions de participació. En general, recollint la tradició de la literatura sobre organitzacions que també havia estat parcialment utilitzada per Olson, s'han definit tres tipus generals d'incentius: utilitaris, normatius i afectius, cadascun vinculat a un procés de motivació diferent, amb l'objectiu d'explicar una varietat més àmplia de processos d'acció col·lectiva voluntària. Un possible esquema interpretatiu seria entendre la decisió d'implicació personal com a resultat d'una funció que integrés tres factors (cada un dels quals es relaciona amb un incentiu organitzacional): elecció racional (sobre béns selectius), afectivitat (relacions internes de grup) i conformitat normativa (sobre els béns públics)<sup>94</sup>.

---

<sup>93</sup> Des de la sociologia es critiquen els supòsits d'Olson i la limitació dels seus incentius, donat que no poden recollir incentius selectius no materials (poder, influència, prestigi, contactes,...), o la pròpia atracció dels béns públics per la participació (vid. H. Yillock i D. Morrison, "Group Size and Contributions to Collective Action: An Examination of Olson's Theory Using Data from Zero Population Growth, Inc.", Research in Social Movements, Conflicts and Change, vol. 2, 1979); però generalment no es planteja la dependència entre els supòsits d'Olson i la capacitat explicativa del seu model.

<sup>94</sup> D. Knoke i C. Wright-Isak, "Individual Motives and Organizational Incentives Systems", Research in the Sociology of Organizations, vol. 1, 1982; i D. Knoke, "Incentives in Collective Action Organizations", American Sociological Review, vol. 53, 1988.

Una altra possibilitat seria distingir tres tipus d'incentius: selectius, solidaris i expressius, en un model en el qual els membres dels grups d'interessos rebessin incentius que es relacionessin amb les circumstàncies individuals, per produir diferents formes d'implicació individual<sup>95</sup>. Per exemple, els membres més apàtics són atrets per uns determinats incentius selectius, i amb els recursos que proveeixen, es desenvolupen altres incentius més refinats pels membres més actius, o amb un potencial de contribució superior<sup>96</sup>.

### 3.5.3. Líders i empresaris com a impulsors de les intervencions de segon ordre.

Les intervencions de segon ordre, com les sancions i els incentius, no recullen l'origen de l'acció col·lectiva, sinó que serveixen per a analitzar processos ja estables. A més, presenten com a dificultat que les relacions entre els factors motivacionals i les intervencions poden arribar a ser molt complexes i no integrables en un model de caràcter predictiu. Així, el seu interès es concentra en la capacitat d'explicar les diferències entre les implicacions dels participants com a funció de constel·lacions concretes de motivacions i incentius organitzatius. Aquests models de decisió consideren que els sistemes d'incentius poden ser molt elaborats i diferenciats, però generalment

---

<sup>95</sup> Aquesta tipologia, proposada per J.-H. Hansen ("The Political Economy...", art.cit., 1985) suggereix que la informació, les preferències i els recursos que hi ha en una particular situació interactuen amb els actuals incentius que l'organització ofereix, amb la qual cosa es forma la convicció subjectiva dels beneficis i costos que entra en els càlculs personals. Es tracta de cercar un equilibri entre demanda i oferta, integrant amb una gran complexitat les motivacions que influeixen en les funcions de decisió. En aquesta mateixa línia, un model que accentua el paper dels incentius no materials es pot trobar a K.-D. Opp, "Soft-Incentives and Collective Action: Participation in the Anti-Nuclear Movement", *British Journal of Political Science*, vol. 16, 1986.

<sup>96</sup> Vid. D. Knoke, "Incentives..", art. cit., 1988, pàg. 326-7; i del mateix autor, *Organizing for Collective Action. The Political Economies of Associations*, De Gruyter, Nova York, 1990. Vid. també G. Marwell, P. Oliver i P. Pahl, "Social Networks...", art. cit., 1988. Aquests autors destaquen el rol crucial dels «organitzadors» per tenir l'habilitat de concentrar esforços organitzatius de caire selectiu -per distints tipus de participants-.

no incorporen la dinàmica interactiva que es produeix entre els possibles participants, conceptualitzant la lògica de l'acció col·lectiva com un procés de decisió -o autocompensació- individual. Cal afegir finalment que aquests models sovint trenquen també la linialitat proposada per Olson sobre el comportament de l'individu (el qual considerava que el mateix supòsit de comportament servia per a explicar l'efectivitat dels incentius selectius i l'aprofitament del bé públic sense participar). D'aquesta manera, doncs, es suposa en general que pel bé públic continua funcionant la lògica defensada per Olson (basada en el supòsit d'interès propi) i que per la participació existeix una complexa i diversa combinació d'incentius i motivacions<sup>97</sup>.

Una possible resposta a l'origen dels incentius selectius pot estar en la figura de l'empresari polític. Aquesta ha estat definida com una possibilitat exògena per explicar la fundació o inici de l'acció col·lectiva. Pot ser una solució pel problema de l'acció col·lectiva: s'argumenta que els empresaris organitzen els incentius selectius pels altres (poden apel·lar també a incentius no materials), i en treuen avantatges privats (poder, diner, fama, carrera política, etc.)<sup>98</sup>. Encara que aquesta explicació es formula en termes de comportament racional i interès propi, també es podrien argumentar altres motivacions per a explicar l'aparició d'iniciatives empresarials. Però tots els intents comparteixen la característica d'ésser explicacions exògenes, que afronten l'explicació de la possibilitat d'emergència espontània de cooperació.

---

<sup>97</sup> Evidentment, aquesta contradicció es destaca més o menys segons els diversos models proposats. En alguns casos -per evitar la contradicció esmentada- es suposa que el propi bé públic funciona com a incentiu selectiu, i que els participants són els individus que tenen preferències no auto-interessades (vid. K.-D. Opp, "Soft-Incentives..", art. cit., 1986); però no es possible matisar les diferents formes possibles d'implicació en l'acció col·lectiva. D'altra banda, la major part d'aquests models destaquen indirectament que la lògica dels incentius selectius no ha d'estar necessàriament vinculada amb el comportament racional i el supòsit d'interès propi, sinó que poden ésser efectius en molts altres supòsits de comportament (en aquest sentit, el tractament d'Elster sobre els incentius selectius és relativament confús, ja que els situa dins la lògica del comportament racional, egoista i orientat pels resultats, Cf. The Cement of..., op. cit., 1989, pàgs. 37-46).

<sup>98</sup> Vid. R. Hardin, Collective Action..., op. cit., 1982, pàg. 34; i T. Moe, The Organization of..., op. cit., 1980.

A l'apèndix de 1971 del seu llibre, Olson reconeix que entre les crítiques més destacades que va rebre la seva obra, la més important és la manca de discussió sobre el possible paper d'un empresari polític, que ajudi a un grup a obtenir el bé públic que necessita. Olson destaca que la incorporació del concepte d'empresari en la provisió de béns col·lectius enriqueix el seu model, tot matisant que l'èxit d'aquest empresari en els grups grans dependrà de la seva capacitat d'innovació en formular incentius selectius, ja que oportunitats de grups en estat de suboptimalitat n'hi ha moltes (LCA, pàgs. 175-178). D'altra banda, recentment han sorgit intents de definir més acuradament el comportament dels líders en l'acció col·lectiva, encara que no s'ha afrontat directament el problema de com justificar la seva intervenció inicial. Es pressuposa que els líders són jugadors amb control sobre la distribució de beneficis generats per l'acció col·lectiva. La seva efectivitat per estructurar els incentius als seguidors està en funció de les capacitats inicials d'aquests: informatives -crear reputació- i estratègiques -per condemnar o recompensar selectivament-, així com la capacitat d'alterar l'estructura de pagaments del joc en relació al bé col·lectiu<sup>99</sup>.

Un altre tema és la possibilitat de cooperació racional impulsada per líders o empresaris sense necessitat d'imposar sancions o incentius. Els empresaris poden estimular l'adopció d'actituds participatives de tipus 'kantià' entre un grup potencial<sup>100</sup>, com a resposta exògena al problema de l'emergència espontània de col·laboració. Però en general, cal pensar que s'ha d'actuar sobre les condicions dels béns col·lectius i les situacions d'interacció que són necessàries per a la seva emergència.

---

<sup>99</sup> Vegeu per exemple, W. Bianco i R. Bates, "Cooperation by Design: Leadership, Structure, and Collective Dilemmas", *American Political Science Review*, vol. 84, 1990.

<sup>100</sup> Com ho mostra l'exemple de la mobilització política al Vietnam (Cf. S. Popkin, *The Rational Peasant*, University of California Press, Berkeley, 1979). També Balme argumenta en un sentit similar "la coopération spontanée est favorisée par la répétition des interactions et par l'entreprise politique, instrumentale ou vocationnelle" ("L'action collective..", cit., 1990, pàg. 282).



El problema dels empresaris o líders que impulsen l'acció col·lectiva, així com el de l'anàlisi del refinat instrumental d'incentius selectius que poden aplicar -en funció de cada situació concreta-per fomentar la provisió del bé col·lectiu, presenten encara moltes qüestions no resoltes. De tota manera, en termes generals, es pot afirmar que si les preferències de la població tendeixen a ser similars (encara que amb motivacions diverses) respecte a un bé públic en particular, el bé es pot produir per coordinació espontània<sup>101</sup>. Si les preferències sobre aquest bé entre la població potencialment interessada són disperses, és més necessari que els empresaris polítics intervinguin per proveir el bé. Aquests no han de modificar el comportament bàsic de la gent en relació a l'estructura de provisió del bé, però poden motivar la participació oferint incentius a canvi d'obtenir beneficis de la seva intervenció.

Per recapitular, es pot concloure que el problema de l'acció col·lectiva no té una solució única. Segons els tipus de béns públics que es desitgin, l'aparició d'una o altra forma d'acció col·lectiva serà més probable. Béns col·lectius diferents poden ésser produïts de formes diferents, però també béns de les mateixes característiques es poden produir amb mètodes diferents, segons les distribucions de preferències dels individus implicats i els processos de decisió que adoptin per a escollir les seves estratègies.

---

<sup>101</sup> M. Taylor, The Possibility of..., op. cit., 1987.

### 3.6. Una discussió com a exemple: classe social i lògiques de l'acció col·lectiva.

#### 3.6.1. La tesi dels privilegis estructurals en l'acció col·lectiva.

El darrer apartat d'aquest capítol es dedica a examinar una discussió recent sobre la influència de l'estructura de classes en les possibilitats de desenvolupar formes d'acció col·lectiva. La discussió es pren com a exemple per a il·lustrar amb el cas d'una teoria específica alguns dels problemes esmentats en els apartats anteriors, especialment els que fan referència a les formes de modelitzar l'acció col·lectiva en connexió amb altres teories socials.

El nucli de la discussió consisteix en analitzar si per diferents classes socials es corresponen distintes condicions a l'hora de desenvolupar processos d'acció col·lectiva. Evidentment, sobre aquesta qüestió sorgeixen discrepàncies entre els diferents enfocaments teòrics sobre el fenomen de les classes socials. Sintèticament, es destaquen dos grans enfocaments. Per un costat, la perspectiva marxista, que considera que l'element de divisió bàsic és la posició dels individus en relació al procés de producció, i que aquest determina també l'activitat política<sup>102</sup>; i d'altre costat els enfocaments d'origen 'weberian', que assenyalen l'existència d'un trencament entre l'acció de classe que té lloc en el procés productiu i aquella que es produeix en el camp polític, essent

---

<sup>102</sup> Les discussions sobre classes socials i la seva interpretació teòrica dins el marxisme s'han desenvolupat força en els darrers anys, especialment a partir dels treballs de E.O. Wright (per exemple, *Classes*, Verso, Londres, 1985; i una revisió crítica a M.A. Cañzos "Explotación, dominación y estructura de clase", *Política y Sociedad*, núm. 5, 1990). Cal afegir que l'element central que es considera en aquesta relació és el concepte d'interès objectiu (sobre els problemes d'explicar l'acció a partir d'aquesta noció, vegeu C. Behan McCullagh, "How Objective Interests Explain Action", *Social Science Information*, vol. 30, 1991).

aquesta darrera possible d'articular a través de diferents eixos de conflicte<sup>103</sup>.

Aquí no s'entra en la discussió sobre el concepte de classe, ni tampoc es vol argumentar que l'acció col·lectiva serveixi per a caracteritzar la formació de les classes socials, a partir de les seves condicions estructurals<sup>104</sup>. Deixant a banda aquestes qüestions, l'elecció de la discussió sobre la influència de les estructures de classe en la producció d'acció col·lectiva vol ser un exemple sobre la importància de les nocions teòriques externes en l'especificació de determinades situacions d'acció col·lectiva. En altres termes, mostrar com determinades concepcions de classe social porten a definir caracteritzacions distintes de l'acció col·lectiva per a fenòmens similars, tot i mantenir una mateixa aproximació metodològica.

Especialment dins el que s'ha denominat 'marxisme analític', però també externament a aquest corrent, s'ha produït durant la dècada dels vuitantes un debat sobre les característiques essencials dels comportaments col·lectius de les distintes classes socials en el capitalisme, mantenint com a denominador comú l'aplicació dels procediments analítics esmentats en aquest capítol. L'article que va donar inici a aquestes disputes fou "Two Logics of Collective Action", elaborat per C. Offe i H.

---

<sup>103</sup> Per una revisió d'aquests enfocaments, vegeu V. Burris, "Classes in Contemporary Capitalist Society: Recent Marxist and Weberian Perspectives"; i R. Blom i M. Kivinen, "Analytical Marxism and Class Theory", ambdós a S. Clegg (Ed.) *Organization Theory and Class Analysis. New Approaches and New Issues*, De Gruyter, Berlin-Nova York, 1990.

<sup>104</sup> Aquesta perspectiva es defensada per J. Elster, quan assenyala que les classes socials es poden definir a partir de "*activities in which they are compelled to engage by virtue of the endowment structure*" (a "Three Challenges to Class", J. Roemer (Ed.), *Analytical...*, *op. cit.*, 1986, pàg. 145). No obstant això, tot i reconèixer que l'anàlisi dels processos d'acció col·lectiva té una gran importància per a entendre les relacions entre estructura de classe i comportaments individuals, sembla més encertada la tesi de què cal referir-se a altres principis per a la seva definició, degut a què, com diu M. Caínzos "*las condiciones de posibilidad de la acción colectiva son enormemente complejas y desbordan el marco estricto de las relaciones de producción*" (cf. "Explotación..", art. cit., 1990, pàg. 92).

Wiesenthal<sup>105</sup>. Aquest treball pretenia formular una crítica a la suposada igualtat institucionalitzada de la democràcia liberal, argumentant teòricament que les associacions d'interessos no competien en les mateixes condicions. Aquestes condicions depenien, segons els autors de l'article, de la realitat de classes, en la qual existeix una gran asimetria de poder i llibertat entre les diverses situacions possibles. Seguint la tradició marxista, definiren dues grans classes socials a la societat -treball i capital-, assenyalant que les relacions de producció entre ambdues impliquen força desigualtat estructural, i que això és determinant per a les pràctiques d'acció col·lectiva.

Adoptant el principi de la desigualtat estructural, o d'existència d'interessos privilegiats estructuralment, la tasca que es proposen explicar aquests autors és com es transmet aquesta desigualtat en els processos de formació d'acció col·lectiva. L'argument defensat és que els interessos individuals de treballadors i empresaris són col·lectivitzats de formes diferents a través de les seves organitzacions. Per un costat, l'empresari, a nivell individual, pot implementar estratègies sobre el mercat i la força de treball, per a assegurar la reproducció del seu capital. D'altra costat, els treballadors, quan volen respondre a les accions individuals d'un empresari, han d'afrontar els problemes d'agregació de preferències que es produeixen per a establir una acció col·lectiva. A més, la situació de desigualtat també es pot observar en què els empresaris poden alterar els factors del seu procés de producció afectant els treballadors (amb la introducció de tecnologia), però aquests no poden alterar de forma exògena el seu procés de reproducció (que depèn en gran part del nivell de salaris), amb la qual cosa es produeix una tendència a l'excés d'oferta de treball<sup>106</sup>.

---

<sup>105</sup> Publicat a Political Power and Social Theory, vol. 1, 1980.

<sup>106</sup> Aquesta afirmació és matisada pels autors, en relació a situacions i èpoques concretes, i sintetitzada amb l'afirmació que a partir d'un determinat nivell d'atur, el cost de l'acció col·lectiva és cada cop més gran pels treballadors que no pas pels empresaris (Cf. "Two Logics..", art. cit., 1980, pàg. 91 i ss.).

D'aquestes condicions estructurals, els autors dedueixen que existeixen unes pautes especials per a l'organització dels treballadors, ja que només amb el seu càlcul individual en molts moments es produirien uns costos massa elevats per la participació, i degut a la competència entre els propis treballadors per obtenir un salari, no es produiria l'acció col·lectiva per mantenir llocs de treball i aconseguir majors salaris. Els empresaris, en determinades circumstàncies, podrien tenir un problema similar per associar-se, però aquest problema comença en un nivell superior al de l'empresa, el qual també el pateixen, a més a més, els sindicats.

En aquest sentit, els autors, per evitar haver d'explicar l'acció col·lectiva a partir de l'acció racional individual, incorporen la noció d'«identitat col·lectiva»<sup>107</sup> per a explicar com, malgrat una situació de desigualtat que gairebé implicaria l'absència d'acció col·lectiva, aquesta es produeix entre els treballadors a les empreses i entre els treballadors de tots els sectors productius. Per Offe i Wiesenhal, la formació d'aquesta «identitat col·lectiva» es més complexa que la lògica instrumental que fan us els empresaris, inclús quan s'uneixen amb la finalitat de defensar els seus interessos -ja que només excepcionalment, aquests estan disposats a actuar-. Això és degut a que aquesta «identitat», que es considera necessària per a motivar als treballadors per l'acció col·lectiva, s'ha de formar a partir d'interessos complexos i parcialment divergents dels treballadors, que s'han d'integrar i definir constantment a través d'un procés comunicatiu entre aquests. També consideren que les organitzacions de treballadors són més complexes de gestionar que les patronals, degut a que presenten un conflicte constant entre els processos burocratitzadors per concentrar recursos (que augmenten quan creix en grup) i els processos de formació d'«identitat col·lectiva» per motivar la provisió de recursos (que es compliquen amb un augment de la grandària del grup),

---

<sup>107</sup> La noció va ser emprada prèviament en aquest context per A. Pizzorno (a "Political Exchange and Collective Identity in Industrial Conflict", C. Crouch i A. Pizzorno (Eds.), The Resurgence of Industrial Conflict in Western Europa since 1968, vol. 2, MacMillan, Londres, 1978).

mentre que les organitzacions d'empresaris tan sols integren interessos, definits fora de l'organització.

### 3.6.2. Les interpretacions sobre l'existència de dues lògiques en l'acció col·lectiva.

En l'argumentació d'Offe i Wiesenthal hi ha dues tesis importants. La primera és que la divisió de la societat en classes comporta diferències estructurals als individus que desenvolupen accions col·lectives, i que aquestes diferències són degudes a la posició de cada agent en el procés productiu. La segona és que per a interpretar algunes formes d'acció col·lectiva, cal sortir del pressupòsit de racionalitat i adoptar la noció d'«identitat col·lectiva». Les dues tesis no són interdependents, però a les discussions posterior sovint s'han presentat integrades com una sola tesi.

Sobre la primera tesi, el que hom pot discutir d'entrada és la noció de 'classe social', en el sentit esmentat anteriorment. Ara bé, si s'accepta la interpretació marxista, és necessari discutir el grau de determinació de l'acció col·lectiva per l'estructura: fins a quin punt les creences que indueixen a l'acció són determinades de forma purament estructural<sup>108</sup>?. En altres termes, com es pot considerar que els actors, donades unes condicions estructurals, interpretaran unívocament la situació com una determinada estructura d'interacció derivada de l'existència d'unes determinades característiques del bé públic en joc?

L'argument d'Offe i Wiesenthal incorpora indirectament el supòsit que la

---

<sup>108</sup> Aquesta és la pregunta de M. Taylor, qui introdueix el suggeriment que "*changes in desires and beliefs can be (at least partly) the direct result of intentional actions -of their own hearers or of other actors- and sometimes the result may even be intended*" (a "Structure, Culture and Action in the Explanation of Social Change", *Politics and Society*, vol. 17, 1989, pàg. 116).

desigualtat estructural impedeix als treballadors defensar els seus interessos racionalment. Aquests no es poden enfrontar individualment a l'empresari i, per tan, des del primer moment ho han de fer col·lectivament (no així els empresaris). Degut a la seva competència interna -si es considera un comportament egoista-, l'estabilitat de la cooperació basada en la racionalitat és excessivament fràgil. En el fons, és la perspectiva d'Olson, sobre la impossibilitat d'acció col·lectiva sense incentius -idealitzada amb el joc del dilema del presoner-; i aquests autors no investiguen si les condicions estructurals són compatibles amb altres tipus de comportaments<sup>109</sup>, o com les creences sobre les condicions per l'acció, mantenint el supòsit de racionalitat, poden portar a altres situacions d'interacció. Degut a que la interpretació d'Olson sobre la realitat sindical a partir dels incentius i sancions no es compartida per Offe i Wiesenhal, aquests darrers rebutjen per extensió tot intent d'ampliar l'anàlisi dels efectes de la desigualtat estructural sobre l'acció col·lectiva a través d'elements de segon ordre. Per explicar la cooperació entre els treballadors introdueixen el concepte de la «identitat col·lectiva», que els hi permet suposar possible canvis de creences.

Hom pot considerar que en la formació de les creences i en la definició dels interessos s'incorporin factors culturals de tota mena, i inclús que en determinades circumstàncies se'n faci un ús instrumental d'aquests valors culturals, per tal de millorar la capacitat d'un grup per a obtenir béns col·lectius<sup>110</sup>, però el que és més difícil d'acceptar és la solució del problema mitjançant la incorporació d'un nou concepte extern que en el fons té un paper molt similar a la noció d'utilitat. Amb connotacions diferents, però amb les mateixes característiques de pressupòsit «ad-hoc» que la

---

<sup>109</sup> La necessitat dels treballadors de defensar col·lectivament els seus interessos, quan el capitalista els pot defensar individualment, és una condició estructural del sistema productiu; però la consideració que aquests interessos han d'ésser egoistes -degut a l'existència d'uns interessos objectius derivats d'una situació d'explotació- implica una doble hipòtesi: que els treballadors són conscients dels seus interessos objectius, i que una situació d'explotació porta «objectivament» a una resposta egoista de comportament.

<sup>110</sup> M. Taylor, "Structure,..." , art. cit., 1989, pàg. 147.

perspectiva teòrica utilitzada. No obstant això, el rebuig complet de qualsevol formulació sobre valors -o inclús identitats- aconduïx a uns problemes anàlegs als que presenta la teoria de la preferència rellevada. Com s'ha exposat al primer capítol, les preferències no es poden deduir exclusivament de les accions, si es volen evitar les vies tautològiques.

Les tesis d'Offe i Wiesensthal han sofert algunes crítiques importants, que són interessants d'assenyalar perquè evidencien distintes concepcions sobre els entorns interpretatius en els quals inserir l'anàlisi dels processos d'acció col·lectiva. Per un costat, sobre la influència de l'estructura de classes, Elster<sup>111</sup> argumenta que malgrat que els treballadors tinguin una diversitat d'interessos més amplia que els empresaris, aquests darrers són més heterogenis, per la qual cosa no creu que els treballadors tinguin més problemes per afrontar l'acció col·lectiva que els empresaris. Amb una recerca empírica sobre organitzacions d'empresaris, Streeck<sup>112</sup> segueix aquest argument i arriba a proposar una tesi inversa: l'acció col·lectiva és més complexa pels empresaris que pels treballadors, degut a què per la seva gran heterogeneïtat és més difícil resoldre el dilema entre identitat col·lectiva i nivell d'organització.

Aquestes crítiques no reconeixen els dos nivells d'acció col·lectiva dels treballadors -dins i fora de l'empresa- i les dificultats de coordinació que impliquen, donat que presenten només els problemes d'organització «política» de cada classe per defensar els seus interessos. Tant Elster com Streeck no consideren que l'estructura de classes de la societat tingui una gran influència sobre la forma com les diferents classes

---

<sup>111</sup> a The Cement of... op. cit., 1989, pàgs. 166-168.

<sup>112</sup> W. Streeck, "Interest Diversity and Organizing Capacity: Two Class Logics of Collective Action?" a R. Czada i A. Windhoff-Heritier (Ed.), Political Choice..., op. cit., 1991. Sobre les dificultats d'articulació d'interessos entre els empresaris, vegeu també J.R. Bowman, "The Logic of Capitalist Collective Action", Social Science Information, vol. 21, 1982.



social poden realitzar accions col·lectives. Aquest autors assenyalen que la manca de connexió entre classes i acció col·lectiva es deguda a dos factors. El primer, que pressuposen una situació inicial d'igualtat individual -o que la propietat no determina els condicionants de l'acció-, en aquest sentit estarien més a prop d'una visió relativista dels conflictes socials, i no tindrien interès en buscar determinants específics<sup>113</sup>. El segon, més instrumental, és que en la seva anàlisi de l'acció col·lectiva de treballadors i empresaris, aquests autors només consideren la variable d'heterogeneïtat en els interessos, la qual cosa provoca conflictes interns, però no introdueixen com a conseqüència les possibles modificacions de les estructures d'interacció a partir de les quals sorgeix l'acció col·lectiva, ni consideren les capacitats -iguals o desiguals- per sancionar de què disposen els agents que hi intervenen<sup>114</sup>.

El segon tema que ha estat discutit a l'article d'Offe i Wiesenthal és la utilització del concepte d'«identitat col·lectiva» com a característic de l'acció col·lectiva dels treballadors. Elster es considera escèptic sobre l'aportació explicativa que incorpora el concepte<sup>115</sup>, mentre que Streeck critica l'exclusivitat d'Offe i Wiesenthal, assenyalant que la formació d'una identitat també és necessària per a resoldre el problema de l'acció col·lectiva dels empresaris<sup>116</sup>. Elster destaca que la noció d'«identitat col·lectiva» pot

---

<sup>113</sup> Des d'aquesta perspectiva, l'estructura de classes no hauria de determinar necessàriament el conflicte social, ja que abans caldria analitzar com alguns actors i formes d'acció tenen un reconeixement estratègic: "*Struggle between classes presupposes political and ideological struggle about class, struggle over whether the political real is to be constituted in terms of class or ethnicity, race, religion, nationality and so forth*" (J. Johnson "Symbolic Action and the Limits...", art. cit., 1988, pàg. 211).

<sup>114</sup> Per aquests motius, la resposta de Wiesenthal ("Two Logics of Collective Action' Revisited. Oder: Ein zweiter Blick auf die 'strukturelle Machtposition des kapitals'", Ms., Bremen, 1990) destaca que el seu model introdueix moltes altres variables, no considerades pels seus crítics, i es reafirma en que "*das Privileg begründet im gegebenen Institutionensystem keine überlegene Handlungsposition gegenüber solchen Akteuren, die sich ebenfalls eine Beschränkung auf 'lokale' Interessen leisten (können)*".

<sup>115</sup> Elster, *The Cement of...*, op.cit., 1989, pàg. 168.

<sup>116</sup> P. Streeck, "Interest Diversity...", art. cit., 1991.

assimilar-se a un procés induït de canvi de preferències per a assolir objectius difícilment assequibles amb les preferències inicials<sup>117</sup>.

A més de la discussió sobre l'aplicació concreta del concepte d'«identitat» per explicar les conseqüències sobre l'acció col·lectiva de la desigualtat estructural -no vinculada necessàriament a la pròpia tesi-, el que és més interessant per analitzar és la seva introducció dins el camp de les interpretacions de l'acció col·lectiva basades en l'individualisme metodològic i en la preeminència de l'elecció racional. La noció de la «identitat col·lectiva» no trenca amb l'individualisme metodològic, però té molt escassament en compte el supòsit de comportament racional, tendint a presentar-se com una alternativa a aquest<sup>118</sup>. Es presenta com alternativa global al comportament racional i centrat en l'interès propi, malgrat que la seva especificació interna es manté en força aspectes com una caixa negra, referida a conjunts d'individus als quals s'assigna interessos comuns no egoistes en les seves relacions internes. Malgrat tot, la presentació d'una estructura d'interacció amb característiques similars es pot realitzar mitjançant l'aplicació d'un model d'acció col·lectiva que contempli la capacitat conjunta de pressió com un bé públic dels treballadors -afectada per variables culturals i d'altres distintes a les que afecten als empresaris-. Pel que fa a l'anàlisi de la formació d'aquest bé, així com de l'aparició de creences associades, cal retornar a les discussions sobre les distintes línies d'investigació esmentades al llarg del capítol; recordant -d'altra

---

<sup>117</sup> J. Elster,, The Cement of... op. cit., 1989. En aquest sentit, l'autor precisa: "*If workers, retaining their utilitarian interests in higher wages, develop a collective identity that also includes an interest in solidarity and discipline, their utilitarian interests are more likely to be satisfied*" (pàg. 168).

<sup>118</sup> Sense entrar en les aportacions de la psicologia social, que resten allunyades d'aquesta discussió, A. Pizzorno, l'introducció d'aquesta noció en l'anàlisi de l'acció col·lectiva, l'ha refinat força més als darrers anys. Un exemple en el qual s'introdueix en l'anàlisi de les organitzacions polítiques és l'article "Algun otro tipo de alteridad: Una crítica a las teorías de la decisión racional" (a Sistema, núm. 88, 1989). Un altre cas en el que es fa un extens us d'aquesta noció es l'article de P. Dunleavy "Group Identities and Individual...", art. cit., 1988. Dunleavy presenta el concepte més orientat cap a significar les formes d'expressió externes i internes de les preferències dels distints individus que formen part d'un grup (del mateix autor, vid. també Democracy, Bureaucracy... op. cit., 1991).

banda- que la noció d'«identitat col·lectiva» tampoc aporta més claredat sobre el problema de la seva formació.

Sintetitzant doncs, el problema no és la noció d'«identitat col·lectiva» i les seves connotacions, sinó la seva presentació com a alternativa a l'enfocament de l'«elecció racional». Això es poc coherent perquè alguns dels suggeriments inclosos dins aquesta noció poden ésser expressats mitjançant una anàlisi més acurada dels processos de formació de preferències i dels interessos que expressen; d'altra banda, altres suggeriments que recull la noció poden ser analitzats com influències culturals en els processos de decisió per la realització d'accions col·lectives, que alterin o modifiquin l'elecció racional (sense necessitat d'excloure completament aquesta en l'anàlisi de l'acció).

**CAPÍTOL QUART:**

**La teoria de jocs com a nucli  
explicatiu de l'acció col.lectiva.**

#### 4.1. Cap a una caracterització teòrica de l'acció col·lectiva.

Si a l'anterior capítol s'analitzaven les interpretacions dels processos d'acció col·lectiva des del paradigma de l'«elecció racional», en aquest capítol es discuteixen específicament les distintes formes d'utilització del concepte 'interacció estratègica' en el context dels fenòmens d'acció col·lectiva. Com que la teoria de jocs pot adoptar formes diferents d'escollir estratègies -a partir d'unes preferències establertes prèviament-, la seva aplicació a l'anàlisi de l'acció col·lectiva obliga a una «resituació» de la hipòtesi de racionalitat: cal justificar perquè s'adopta aquesta hipòtesi, que ha de dependre d'un context interpretatiu determinat. L'examen de les solucions «externes» i les «internes» als problemes de l'acció col·lectiva realitzat anteriorment, ha mostrat la impossibilitat de generar solucions «internes» plausibles -en termes dels seus supòsits de partida-, mentre que les solucions «externes» presentaven força dificultats d'integració teòrica amb els postulats que provenen directament de l'economia. Per aportar elements que resolguin les dificultats de les solucions «externes», en aquest capítol es proposa una pregunta que es podria plantejar de forma preliminar com: *si la teoria de jocs pot combinar mecanismes de decisió racional amb altres formes de decisió, és possible utilitzar aquesta teoria com a instrument formal que faciliti la construcció de teories polítiques o socials amb contingut fàctic?*

En el cas de poder respondre afirmativament aquesta qüestió, podria aparèixer una via per eliminar algunes de les dificultats que presenten les solucions «externes» per a explicar els problemes de l'acció col·lectiva. Això seria possible degut a la disponibilitat d'un nucli homogeni d'anàlisi i interpretació, que assenyalaria uns requeriments mínims per a les teories amb contingut fàctic que l'han de complementar. Però hi ha dos punts claus que cal afrontar per aclarir aquest assumpte. Primer, si es pot simplificar de forma convenient la teoria de jocs per a formular aplicacions

versemblants en relació als fenòmens d'acció col·lectiva. Per aquest motiu, tot i que els problemes analítics de la teoria de jocs en el marc de la teoria social ja han estat discutits en el segon capítol, aquí s'aprofundeix en aquesta discussió, en relació a les condicions especials que implica l'existència de situacions de 'N' jugadors, habituals en els problemes de l'acció col·lectiva. Segon, si es poden aïllar amb claredat els resultats formals que ofereix la teoria de jocs dels resultats interpretatius que aporta la teoria amb contingut fàctic. L'objectiu és poder distingir en cada fenomen d'acció col·lectiva dos nivells d'anàlisi diferents. El primer, possible de modelitzar utilitzant només la teoria de jocs, seria referent als «resultats de les eleccions» (un cop s'han determinat les preferències dels individus); mentre que el segon nivell, referent a l'anàlisi de les «situacions d'elecció», hauria d'incorporar elements explicatius externs a la teoria de jocs, que justificuessin les preferències adoptades pels individus<sup>1</sup>, així com mecanismes que integressin els efectes de successius «resultats d'eleccions».

La distinció entre «situació d'elecció» i «resultat d'elecció» serveix per a il·lustrar que la teoria de jocs, sense ocupar un paper exclusiu, pot tenir una gran interès per ésser aplicada a la interpretació de l'acció col·lectiva. En el seu estat de desenvolupament actual, la teoria de jocs té utilitat en l'anàlisi dels «resultats de eleccions». A partir de les característiques pròpies dels béns col·lectius i el nombre d'electors (com variables provinents de la «situació d'elecció»), que configuren uns models de jocs específics, apareix el nucli bàsic amb el qual s'explica el resultat d'una acció col·lectiva. Però seguidament cal desenvolupar línies d'anàlisi per connectar aquest nucli bàsic amb les formes d'elecció d'estratègies, que generen les decisions individuals i seguidament els resultats de les eleccions. No obstant això, si aquest procediment és vàlid, disposar d'un potent instrument formal capaç d'integrar-se

---

<sup>1</sup> Com ho fa T. Schelling, quan reconeix que els seus models es refereixen als resultats de les eleccions binàries, però no a les «situacions d'elecció» binària, sobre les quals influeixen molts altres factors no inclosos en els seus models (cf. *Micromotives and Macrobehavior*, Norton and Co., Nova York, 1978, pàgs. 241-243).

fàcilment en formulacions teòriques més àmplies, és ja un pas important en relació a la qüestió plantejada anteriorment, que permet afrontar més fàcilment la resolució de les solucions «externes» dels fenòmens d'acció col·lectiva.

Abans d'analitzar les possibilitats de simplificació de la teoria de jocs, cal plantejar una qüestió conceptual que en els estudis d'aquest àmbit resta sempre força ambigua. Quina és la vinculació entre l'anàlisi de l'acció col·lectiva i la teoria política, si s'entén que la teoria de jocs pot servir com a nucli per a interpretar l'acció col·lectiva? Aquest nucli no es refereix pròpiament a cap àmbit concret de les ciències socials, ja que en tots ells apareixen processos d'eleccions interactives en presència de béns públics. Les distincions específiques dins les ciències socials, que no configuren camps absolutament tancats, pel que fa a l'explicació de fenòmens d'acció col·lectiva, s'han de referir als factors que influeixen en les «situacions d'elecció». Les adaptacions teòriques sobre com interpretar aquestes situacions, així com els propis referents empírics a què aquestes situacions es refereixin, són els que poden caracteritzar les diverses adaptacions teòriques sobre l'acció col·lectiva en la política, l'economia o la sociologia. Pel cas de la política, l'element analític caracteritzador pot ser la interpretació dels processos de coordinació i l'anàlisi del seu grau de consciència per part dels actors, així com la discussió sobre les institucions en què els fenòmens d'acció col·lectiva estan inserits.

Una discussió global sobre el paper de la racionalitat i la teoria de jocs a les ciències socials es deixa per a les conclusions finals de la tesi, mentre que en aquest capítol es tracten bàsicament les dificultats existents per determinar els «resultats de les eleccions». La relació entre les eleccions i la interpretació de la «situació d'elecció», un punt en el qual divergeixen les distintes perspectives analitzades en el segon capítol, té també les seves controvèrsies en el camp de l'acció col·lectiva. Presentant caracteritzacions distintes del paper de la racionalitat, temes com la relació entre regles

i teoria de jocs (vid. apt. 2.2.3.), la transformació de les funcions d'utilitat (vid. apt. 2.2.4.) o la introducció de complexos supòsits sobre el comportament interactiu dels individus en relació al seu entorn de decisió (vid. apt. 2.2.5.), competeixen per obrir vies convincents que aclareixin aquesta relació. Degut a la seva generalitat, els procediments d'anàlisi de les «situacions d'elecció» en l'acció col·lectiva (vid. seccions 3.5. i 3.6.) han de ser capaços d'integrar paral·lelament diversos mecanismes de formació de preferències i de pressa de decisions, definint combinacions específiques per diversos àmbits temàtics (vid. secció 4.4.). El capítol es centra especialment en discutir les diverses consideracions simplificadores que són necessàries d'adoptar per fer aplicable la teoria de jocs a l'anàlisi de l'acció col·lectiva, que han de ser compartides per totes les aplicacions. També s'analitzen les conseqüències que se'n deriven d'aquestes simplificacions per la capacitat predictiva de la teoria de jocs, especialment quan diverses variables són definides en entorns teòrics més àmplis.

En primer lloc, a la secció 4.2., es qüestiona la introducció de supòsits simplificadors sobre les condicions dels jocs amb 'N' jugadors, en especial respecte als requisits d'informació i de determinació d'equilibri. A la secció 4.3. s'introdueixen supòsits simplificadors sobre les característiques dels jugadors i les seves estratègies, que limiten el nombre dels resultats possibles. Així, s'analitzen les condicions per a reduir el nombre d'eleccions -estratègies- possibles als actors, amb models que formalitzen preferiblement eleccions binàries quan es consideren 'N' jugadors amb idèntiques formes de comportament<sup>2</sup>. La simplificació màxima -però poc convincent- d'una situació en què intervenen 'N' persones és proposar una matriu de dos per dos, considerant com a supòsit que tothom juga un joc contra tots els altres. A la secció 4.4. s'introdueix una altra perspectiva: fins el moment -per simplificar- s'havia considerat

---

<sup>2</sup> A. Colman, *Experimental Games...*, op. cit., 1982, pàgs. 156-166. Per no complicar excessivament la discussió dels problemes que es plantegen, es manté en aquest apartat el supòsit que existeix en aquests models una preeminència d'un comportament que maximitza l'interès propi, amb poques possibilitats de variació.



que tothom tenia la mateixa orientació d'interès davant un problema d'acció col·lectiva; però l'observació empírica mostra que existeixen moltes diferències entre els individus, constituint aquesta situació una dificultat més per a l'aplicació plausible dels models en moltes situacions<sup>3</sup>. La secció analitza com es poden reduir les diferències de comportament entre els jugadors, agrupant aquests entre grups que tenen el mateix ordre de preferències (i per tant, la mateixa percepció sobre els resultats possibles), concentrant l'atenció en els problemes de modelitzar els processos que s'estableixen entre orientacions d'interès i de càlcul interactiu distints.

Les formes de modelització simplificada permeten proposar diferents aplicacions sobre situacions d'acció col·lectiva de natura diversa. *No existeix una forma preeminent de modelitzar l'acció col·lectiva, en el sentit que des del paradigma de l'«elecció racional» no existeix una teoria general satisfactòria sobre l'acció col·lectiva*, almenys des de la perspectiva metodològica proposada en aquesta tesi i, per tant, no es pot desenvolupar una correspondència directa amb un model aplicable a tots els casos. L'únic model general possible en termes de teoria de jocs, amb totes les variables definides externament, és inaplicable empíricament i intractable formalment. Per aquest motiu, en cada cas, i depenent de la finalitat explicativa, pot tenir més interès emprar una o altra estratègia de simplificació, suggerida per un coneixement teòric del problema i limitacions de cada procediment. La possibilitat d'establir proposicions teòriques parcials sobre les «situacions d'elecció» en els processos d'acció col·lectiva pot sorgir de l'aïllament de variables, durant la realització del procés de modelització (amb les possibilitats instrumentals que ofereix la teoria de jocs), degut a la necessitat d'escollir entre les diverses possibilitats de simplificació existents.

---

<sup>3</sup> Com assenyalava J. Elster, "The assumption that the actors are homogeneous and interchangeable, while useful for expositional purposes, is clearly unrealistic" (*The Cement of...*, op. cit., 1989, pàg. 49).

## 4.2. Questionant l'aplicació de la teoria de jocs a l'anàlisi de l'acció col·lectiva.

### 4.2.1. Interdependència i acció col·lectiva.

Els matemàtics que treballen en teoria de jocs no estan excessivament preocupats pels problemes de la seva aplicació a les constel·lacions del món real; en canvi, aquests problemes són de vital importància per als científics socials que fan recerques empíriques i volen emprar models derivats d'aquesta teoria. Mantenint el supòsit de racionalitat, cada jugador ha d'anticipar la resposta dels altres jugadors als seus moviments esperats, per tal d'escollir finalment l'estratègia òptima. En principi, les limitades capacitats computacionals dels éssers humans fan implausible ràpidament aquest supòsit quan el nombre de jugadors i d'estratègies augmenta<sup>4</sup>, si es considera qualsevol mecanisme de càlcul que no sigui el seguiment de pautes d'elecció d'estratègies fixades d'antuvi. Això implica greus dificultats per la teoria matemàtica dels jocs quan és aplicada directament en la interpretació de fenòmens socials, ja que difícilment els actors poden complir amb les prescripcions que la teoria assenyala pel seu comportament. La crítica a la manca de plausibilitat de la teoria de jocs té com a punt de partida les elevades capacitats cognitives i computacionals que generalment la teoria de jocs imputa als seus actors «idealitzats». Així, si generalment en l'anàlisi de

---

<sup>4</sup> Cal pensar que un joc entre X jugadors, amb Y estratègies per jugador, obliga a cada jugador -seguint la teoria estàndard- a considerar, comparar i ordenar  $(Y)^X$  resultats possibles (per exemple, amb 3 jugadors i 4 estratègies a disposició de cada jugador, serien possibles 81 resultats distints). Aquesta simple mostra indica les dificultats per considerar que els individus calculen realment les seves accions de forma tan complexa. La capacitat discriminatòria dels humans -sense un gran entrenament previ específic- no pot superar més de set nivells diferents. Sobre les limitacions computacionals humanes encara hi ha relativament poc conegut, especialment pel que fa referència als processos d'elecció en contextos socials. Un dels articles seminals d'aquest assumpte és el de G.A. Miller, "The Magical Number Seven, Plus or Minus Two. Some Limits on Our Capacity of Processing Information", *Psychological Review*, vol. 63, 1956. Una recent aproximació a la qüestió, integrant els problemes de la hipòtesi de racionalitat, ha estat realitzada per H.Simon, "Invariants of Human...", art. cit., 1990.

l'acció col·lectiva es suposa un nombre d'actors força elevat, aquesta crítica encara podria ser més adequada, tant pel que fa als propis actors, com també pels investigadors que han de reconstruir posteriorment les seves interaccions.

Tanmateix, els fenòmens que són recollits sota el rètol d'acció col·lectiva tenen en comú que, tot i existir un nombre elevat de subjectes, existeix interdependència, tant en el sentit d'un observador extern, com també en el sentit que els subjectes calculen les seves accions de forma conscient considerant que es tracta d'eleccions interdependents. Les situacions d'acció col·lectiva impliquen que els individus poden ésser conscients que les seves eleccions modifiquen els paràmetres. Encara que no necessàriament la consciència s'ha d'expressar en forma de càlcul racional, ni aquesta ha d'ésser explícitament present perquè es produeixi una situació d'acció col·lectiva. Així, cal discutir si la consciència de la interdependència es pot modelitzar completament en termes de teoria de jocs, o si cal vincular-la a la introducció d'elements de reforç de segon ordre<sup>5</sup>. L'anàlisi de la consciència es pot relacionar també amb la necessitat d'establir «políticament» la coordinació per aconseguir la producció de molts béns públics, a través de diversos mecanismes i institucions vinculants. Aquestes situacions són contràries a les estructures de competència perfecta, on s'estableix la coordinació entre individus a través dels preus, com a informació que agrega els efectes de la interacció; i els subjectes no poden -mitjançant les seves eleccions- afectar o modificar els preus. Malgrat tot, poden aparèixer altres mecanismes de coordinació en processos d'acció col·lectiva en els quals existeix interdependència però no ajust interactiu. Per exemple l'ajust seqüencial, en què els individus segueixen un pla o un ordre de decisió prèviament establert<sup>6</sup>. Cal notar que en tot els casos, pot

---

<sup>5</sup> Es podria considerar -encara que no necessàriament- l'existència d'un cert grau de consciència indirecta a través de les actituds davant sancions, requisits d'entrada, expectatives sobre respostes a accions pròpies.

<sup>6</sup> Cf. M. Granovetter, "Modelos de umbral de conducta colectiva", *Zona Abierta*, nùm. 54-55, 1991 (ed. or. 1978). Vid. també l'apartat 4.3.3..

incorporar-se -com habitualment es fa- el supòsit de comportament racional, però no és estrictament necessari per denotar la classificació. A tots els casos de coordinació, tant el mercat, l'ajust interactiu o l'interdependència seqüencial, els mecanismes de decisió aplicats pels individus poden basar-se en regles, normes, rutines o càlcul racional; i la seva aplicació com a supòsit de comportament depèn de la interpretació sobre els subjectes que s'estableixi<sup>7</sup>.

A partir d'aquesta caracterització, cal discutir primer quin tipus de relació existeix entre els individus en un context determinat i, seguidament, considerar si és necessària, malgrat els seus problemes d'aplicació, la modelització del fenomen en termes de teoria de jocs. Si es tracta d'un fenomen d'acció col·lectiva, pot ser que per la seva interpretació siguin útils els models formals de teoria de jocs, però també és possible que amb modelitzacions més simples es puguin desenvolupar explicacions convincents. Si s'entén que els fenòmens d'acció col·lectiva són aquells que impliquen una interdependència, així com la consciència d'aquesta interdependència per part dels subjectes, al capítol es considerarà que -excepte per alguns casos particulars- els models de la teoria de jocs són els més aptes per recollir aquests fenòmens. No obstant això, reconeixent la crítica esmentada a les impossibilitats de càlcul dels agents, hom podria qüestionar aquesta proposta, defensant que s'ha de rebutjar completament l'aplicació de la teoria de jocs. En aquest sentit, com destaca F. Scharpf, no cal rebutjar la teoria de jocs, sinó tan sols limitar les seves capacitats formals en la utilització que es proposi dins les ciències socials:

*"In empirical applications of game theory, by contrast, the direct correspondence*

---

<sup>7</sup> Per exemple, si es considera l'existència d'entorns socials molt estables, es poden aplicar més fàcilment que quan la dinàmica social és intensa, els models d'aprenentatge que seleccionen regles de comportament mitjançant la valoració dels resultats anteriors (vid. per exemple, M. Macy, "Walking Out of Social Traps", *Rationality and Society*, vol. 1, 1989; i U. Witt, "The Evolution of Economic Institutions as a Propagation Process", *Public Choice*, vol. 62, 1989).

*between the mental operations that are used by the analyst, and those that are imputed to the actors, forces us to work with only a subset of the analytical procedures developed by mathematical game theorist*"<sup>8</sup>.

Així, la qüestió és saber quina part de la teoria de jocs pot utilitzar-se, i quins supòsits cal establir per limitar les demandes computacionals que imposa la teoria de jocs. És precisament aquesta tensió la que serveix per articular el capítol, analitzant les diverses línies de combinar plausibilitat i capacitat explicativa. En el cas que no es pugui pressuposar cap tipus de consciència sobre la interacció, o que no existeixi cap coordinació política, cap mecanisme de reforç, ni cap procés deliberatiu sobre l'entrada o la sortida, etc., potser cal considerar que l'individu es troba també en un problema de decisió davant la natura -modelitzable a través dels mecanismes de competència perfecta-, encara que no es tracti de cap activitat econòmica, sinó de conflictes socials, o adquisició de béns públics. En aquests casos, l'aplicació de models que pressuposen la interdependència pot portar a complicacions innecessàries, sobretot quan als models es suposa també el càlcul dels individus sobre tots els efectes de les seves decisions<sup>9</sup>. Així, es pot resumir la qüestió assenyalant que els processos d'interacció considerats dins l'àmbit de l'acció col·lectiva no inclouen els models d'equilibri en una estructura de competència perfecta. Existeixen diversos fenòmens socials en què són aplicables els esquemes del mercat competitiu, els quals, potser no del tot adequadament, de vegades també es nomenen com fenòmens d'acció col·lectiva.

---

<sup>8</sup> F. Scharpf, "Games Real..", art. cit., 1989, pàg. 13.

<sup>9</sup> L. Rothenberg ("Organizational Maintenance...", art. cit., 1988) defensa que per a explicar la contribució dels individus en els grans grups, l'aplicació de teoria de jocs és poc apropiada, defensant l'estructura analítica de les decisions davant la natura per aplicar en aquestes situacions (pàg. 1130). Davant aquesta perspectiva, es pot argumentar que l'aplicació d'una estructura de competència perfecta implica renunciar a analitzar la situació en termes d'acció interdependent.

#### 4.2.2. La teoria de jocs no és la teoria de l'acció col·lectiva.

Recentment, i dins el propi enfocament de l'«elecció racional», s'ha produït una interessant discussió sobre la conveniència d'aplicar models de teoria de jocs per explicar processos d'acció col·lectiva. El principal argument contra aquesta utilització destaca que després de força anys de discussions i elaboracions teòriques, les aplicacions de teoria de jocs per l'explicació de l'acció col·lectiva tenen una escassa capacitat d'aplicació empírica i, a més, que en els seus propis desenvolupaments interns mostren molts problemes teòrics pendents. Sovint, aquestes crítiques coincideixen amb aquelles contra els defensors d'una solució «interna» bàsica. La preeminència d'una solució «interna» és defensada argumentant que la introducció d'elements explicatius externs pressuposa un problema de segon ordre en la lògica de l'acció col·lectiva i que, per tant, l'única solució completa ha de ser la que expliqui l'origen de l'acció col·lectiva a partir d'un conjunt d'axiomes inicials sobre el comportament individual (vid. apt. 3.4.). Aquest argument pot semblar correcte, però hom pot qüestionar-se si la solució «interna» ha d'ésser expressada exclusivament amb el supòsit de comportament racional i modelada a través de la teoria de jocs<sup>10</sup>. Per aquest motiu, donades les dificultats per evadir l'arbitrarietat d'aquest o un altre supòsit, és més adient proposar la construcció d'interpretacions teòriques sobre el procés de decisió interactiu, en bona part externes a la teoria de jocs, que justifiquin l'adopció de supòsits de comportament i expliquin la «situació d'elecció».

La teoria dels jocs repetits, en les seves distintes versions (vid. apt. 2.1.2.), ha estat l'instrumental bàsic emprat per aportar explicacions «internes» al problema de l'acció col·lectiva. Però d'aquestes aplicacions es critica la seva manca de plausibilitat, deguda especialment a dos factors: l'existència d'equilibris múltiples i el manteniment

---

<sup>10</sup> Vid. K. Basu, "On the Non-Existence..", art. cit., 1990. Sobre el problema en general, vid. apt. 2.2.4.

del supòsit d'informació completa sobre el comportament dels jugadors. Pels jocs repetits, el supòsit d'informació completa implica que al final de cada etapa del joc, cada jugador obté informació de l'estratègia que han adoptat tots els altres jugadors en aquesta etapa<sup>11</sup>. Cada jugador, en el moment de fer la seva elecció en una etapa concreta, disposa d'informació sobre les estratègies adoptades pels altres jugadors en totes les etapes prèvies, amb la qual cosa pot calcular la resposta «racional» dels altres jugadors als seus propis moviments. Un supòsit addicional és que no es consideren els costos d'informació per tenir coneixement d'aquestes estratègies (vid. apt. 1.1.4.). És possible relaxar els supòsits sobre la disponibilitat d'informació dels jugadors en els jocs repetits, introduint el factor «cost», però el resultat d'aquest procediment és que augmenta la dificultat per trobar punts d'equilibri (eficients en el sentit de Pareto). Com afirma encertadament Hechter,

*"If the players have zero monitoring costs, then there is a much greater chance that they can reach a cooperative equilibrium"*<sup>12</sup>.

Els costos de la informació sobre les estratègies anteriors dels jugadors depenen en part del nombre de jugadors implicats en el joc. Si aquest nombre és molt elevat, també seran molt elevats els costos i, per tant, impossibles d'assumir en l'explicació. Sense conèixer les estratègies anteriors dels altres jugadors, les possibilitats d'arribar a solucions cooperatives -en termes de la teoria- es reduiran. En els casos de jocs repetits amb dos jugadors, on és més possible de considerar de forma realista els costos d'informació, les anàlisis realitzades il·lustren que també augmenta la dificultat per

---

<sup>11</sup> El model concret de referència és proposat per M. Taylor, *The Possibility of...*, op. cit., 1987. El motiu d'escollir aquest és que els requisits sobre informació són comparativament poc exigents.

<sup>12</sup> M. Hechter, "On the Inadequacy of Game Theory..", art. cit., 1990, pàg. 243.

trobar equilibris eficients, quan es considera costosa l'adquisició d'informació<sup>13</sup>. Aquests problemes, donat l'elevat nombre d'actors inclosos en les situacions d'acció col·lectiva, porten a cercar altres solucions. Una possibilitat és introduir a la definició dels jocs elements contextuals per simplificar la seva complexitat formal, que permetin suposar l'existència de informació completa amb uns costos molt reduïts de recerca i obtenció d'aquesta (i redueixin també l'arbitrarietat en la definició dels supòsits de comportament). Això es realitza a partir d'interpretacions sobre la «situació d'elecció» que, per exemple, permeten considerar que els actors utilitzen uns factors de decisió limitats en l'elecció d'estratègies, perquè es mouen en entorns que ja coneixen força bé. En aquest sentit, també es poden proposar interpretacions sobre els nivells de raonament o sobre els entorns constitucionals dels individus que realitzen les eleccions, per assenyalar com estan limitats els processos de decisió dels actors en el món real.

Per resoldre el problema de la manca d'informació completa requerida per a augmentar la capacitat de determinar els equilibris d'un joc, hi ha dues alternatives analítiques. La primera, considerar que degut al seu cost normalment només existeix informació incompleta i que, per tant, cal introduir el supòsit que els individus utilitzen elements probabilístics (per exemple, que aquests apliquen la teoria de Bayes). La segona, suposar que existeix informació completa, però que el cost d'assolir aquesta és molt reduït, degut a què la «situació d'elecció» ja incorpora la major part de la informació necessària per una elecció.

Com que s'ha observat que els individus utilitzen molt poc la probabilitat en les seves eleccions<sup>14</sup>, només resta explorar l'alternativa de resoldre el problema de la

---

<sup>13</sup> M. Hechter, "On the Inadequacy of Game Theory..", art. cit., 1990, pàg. 244.

<sup>14</sup> Així ho argumenta J. Elster a "Taming Chance: Randomization in Individual and Social Decisions" (dins Solomonic..op. cit., 1989; trad. cast. Paidós, Barcelona, 1991).



informació completa suposant la seva existència amb un cost molt reduït. Una possible via que permet reduir aquest cost és desestimar un dels supòsits més restrictius que fa servir la teoria de jocs per refinar el seu concepte d'equilibri: la hipòtesi de «coneixement comú»<sup>15</sup>, que considera que cada actor disposa d'informació sobre l'ordenació de preferències dels altres actors. Sense el «coneixement comú», resten com supòsits informacionals del joc el coneixement de cada actor sobre les seves pròpies preferències, les estratègies possibles de tots els altres actors, i el conjunt de resultats possibles del joc. Amb aquests elements, els actors no poden anticipar amb seguretat les eleccions dels altres per refinar les pròpies i, per tant, es perden probabilitats de trobar equilibris òptims. Malgrat això, encara segueixen essent possibles els equilibris basats en les estratègies maximin<sup>16</sup>. Una segona via apareix quan es considera que existeixen eleccions en les que hom pot suposar l'existència d'un conjunt d'informacions compartides sobre el joc i el seu context, a partir de les quals els actors poden preveure amb relativa facilitat les preferències dels contraris. Aquesta via per reduir els costos d'informació que requereix el coneixement de les preferències dels altres és proposada per F. Scharpf quan esmenta les «situacions d'elecció»<sup>17</sup>. Aquest autor parteix de la idea de «sentit comú»<sup>18</sup> entès com una analogia social -poc costosa- de la hipòtesi de «coneixement comú» sobre la racionalitat dels jugadors que existeix a la teoria de

---

<sup>15</sup> Definida explícitament per R. Aumann, a "Agreeing to Disagree", Annals of Statistics, vol. 4, 1976. La hipòtesi expressa el supòsit segons el qual tots els agents són conscients de la mútua racionalitat i de la igualtat en la valoració de les seves funcions d'utilitat.

<sup>16</sup> F. Scharpf, "Games Real Actors Could Play. The Problem of Mutual Predictability", Rationality and Society, vol. 2, 1990, pàgs. 481-483.

<sup>17</sup> Referint-se a "influences of the preexisting societal and institutional environment in which players and their games are embedded" (F. Scharpf, "Games Real Actors...", art. cit., 1990, pàg. 477).

<sup>18</sup> Connexió desenvolupada per J.-P. Dupy, "Common Knowledge, Common Sense", Theory and Decision, vol. 27, 1989.

jocs<sup>19</sup>. L'existència d'una situació de «sentit comú» entre els jugadors, pot servir a l'analista per discriminar una gran part de les possibles ordenacions subjectives de preferències, al mateix temps que facilitaria també als actors arribar a «convencions» o «ordres espontanis», equilibris no sempre òptims ni completament estables que explicarien el resultat d'una situació d'interacció. L'existència d'un espai de «sentit comú» implica l'aparició d'orientacions de comportament compartides i inserides en contextos socials fàcilment reconeguts pels actors que hi participen, que també poden configurar en bona part regles de decisió o de raonament per la selecció d'estratègies. No cal insistir que l'explicació de la formació d'aquestes xarxes de «sentit comú» resta fora de l'àmbit de la teoria de jocs i que la seva utilització té sentit com a construcció provisional, que hauria de deixar pas a proposicions teòriques més ben fonamentades. No obstant això, l'existència de contextos interpretatius sobre les relacions socials, que generin espais de «sentit comú» entre els jugadors, possibilita una limitació de les estratègies pensables. Aquests contextos faciliten als actors la possibilitat d'una mútua predictabilitat -socialment construïda-, a més que simplifica la natura dels jocs d'interacció en què aquests estan immersos, la qual cosa també fa més fàcil la seva anàlisi.

Malgrat les estratègies de simplificació i la introducció de supòsits sobre la informació assequible als actors dels jocs, pot ésser que el problema de determinar una solució entre els equilibris múltiples continuï més o menys present. Sovint apareixen diversos punts d'equilibri, tant en jocs simples com en superjocs, i no tots aquests punts d'equilibri són eficients en termes de Pareto. També en poden aparèixer d'ineficients, que deixin sense resoldre el problema de l'acció col·lectiva. Així, a continuació de la identificació dels equilibris per part dels actors sorgeix un problema de negociació no-

---

<sup>19</sup> R. Aumann, "Agreeing to..", art. cit., 1976; però sense implicar en aquest cas expectatives mútues de comportament racional.

cooperativa sobre com escollir entre els punts d'equilibri que siguin eficients. En aquestes situacions, existeixen dos tipus de dificultats. Per una banda, de quins mecanismes disposen els actors per a distingir entre equilibris eficients i equilibris ineficients. D'altra banda, com es poden posar d'acord els agents per a escollir un equilibri entre els diversos possibles que són al seu abast.

El problema dels equilibris múltiples és un dels més importants que cal afrontar per l'aplicació de la teoria de jocs. En el cas dels jocs repetits, la introducció de taxes de descompte sobre el futur és un mètode que s'ha emprat per a reduir el nombre dels equilibris ineficients, i inclús per a proposar l'existència d'un únic equilibri eficient<sup>20</sup>; però, deixant a banda la discussió sobre la possibilitat de càlcul d'aquest equilibri per part dels jugadors, resta la qüestió de com saber quina és la taxa de descompte que s'ha d'escollir perquè el joc repetit tingui un únic equilibri. Tot i que la taxa pugui ésser calculada formalment, ben segur que no és aquest el procediment que segueixen els individus per adoptar-la. El problema de explicar la seva adopció es important, perquè s'ha de calcular arbitràriament una actitud front al risc, sense disposar de cap forma d'explicar les preferències envers aquest<sup>21</sup>. Per evitar aquestes dificultats en la determinació d'equilibris, alguns autors expressen la necessitat d'incorporar més elements derivats de la «situació d'elecció» -justificant, per exemple, la taxa de descompte del futur-. Aquests elements han de facilitar la interpretació de les eleccions dels actors, de forma externa a la teoria de jocs. Aquests autors proposen la utilització combinada de mecanismes de reforç amb altres elements de segon ordre sobre la lògica de l'acció col·lectiva per a explicar els resultats que s'observen. Ara bé, en aquestes

---

<sup>20</sup> Els models analitzen casos de jocs repetits entre dos jugadors, o bé de competicions entre parelles de jugadors. Vid. M. Taylor, "Cooperation and Rationality..", art. cit., 1990, pàg. 228; i R. Axelrod, La evolució de..., op. cit., 1986.

<sup>21</sup> Hechter presenta paradoxalment el problema en aquests termes: "In what way can an internal solution to the collective action problem rely upon such a gratuitously exogenous assumption?"; a "On the Inadequacy of Game Theory..", art. cit., 1990, pàg. 244.

situacions torna a aparèixer el problema de la connexió:

*"the task is to find a path from internal solutions of the collective action problem to external ones"*<sup>22</sup>.

Una proposta concreta per solucionar aquest problema és formulada per Hechter. Aquest autor intenta explicar endògenament l'aparició de regles i sancions, basant-se en la premissa d'una alta visibilitat en la producció i distribució del bé col·lectiu. Donada aquesta situació, Hechter considera que és racional pels individus establir controls formals en les institucions cooperatives per tal de preservar la integritat de les seves inversions i evitar amb les sancions les actituds de «free-rider». Malgrat tot, si no es donen les condicions d'alta visibilitat en la producció i consum dels béns públics, que permetin la distinció dels equilibris eficients, tampoc funcionarà aquesta proposta. Encara que l'autor argumenta que els participants ja no hi volen entrar, hom pot trobar molts casos en què es produeix acció col·lectiva, amb condicions que presenten un cert grau d'opacitat. Per aquest motiu les condicions d'Hechter no poden ésser generalitzades per a explicar tots els problemes de l'acció col·lectiva, tot i que s'adaptin a determinats contextos. L'intent d'explicació -precisament a causa de la seva parcialitat- mostra que no hi ha una solució general del problema de la connexió entre les solucions «internes» i les «externes». Així, si la incorporació d'elements contextuals a les condicions del joc (per exemple, la noció de «sentit comú» o la taxa de descompte del futur) no és suficient per determinar solucions d'equilibri, cal incorporar en l'explicació de cada acció col·lectiva factors externs, com poden ser especificitats físiques o processos de producció, que indiquen la necessitat de construcció teòrica per precisar les limitacions d'acció dels individus en diferents àmbits socials. L'explicitació de l'estructura

---

<sup>22</sup> M. Hechter, "On the Inadequacy of Game Theory...", art. cit., 1990, pág. 245. En aquest sentit, vegeu també altres proposicions a la secció 3.5.

d'interacció pot ésser necessària, però molts cops insuficient per explicar el sorgiment de l'acció col·lectiva; per aquest motiu cal també analitzar els fenòmens d'entrada i sortida del joc (i, a partir d'aquests, les condicions de la interacció).

Per concloure aquest apartat, cal remarcar les tres qüestions més significatives que s'han esmentat. En primer lloc, la relació existent entre la teoria de jocs i l'acció col·lectiva: no totes les aplicacions de teoria de jocs tracten sobre acció col·lectiva, sinó que *l'acció col·lectiva, en referir-se als processos d'interacció entre individus en presència de béns públics, ha d'emprar en el seu nucli bàsic la teoria de jocs (com a formulació més general); tot i que la teoria de jocs no pot incloure suficients factors per a ser equivalent a una teoria de l'acció col·lectiva*. Això és degut a què la teoria de jocs té capacitat per a modelitzar diferents formes de coordinació d'eleccions (per exemple, les decisions polítiques que impliquen consciència sobre valors col·lectius), amb graus distints de consciència, esforç i implicació. La segona qüestió que s'ha discutit és la limitació del paper de la hipòtesi de racionalitat. La discussió important per al desenvolupament d'una teoria de l'acció col·lectiva ha d'ésser al voltant de la forma d'utilització d'aquest concepte. La utilització de la teoria de jocs és necessària per integrar la hipòtesi de racionalitat -com un mecanisme de decisió, entre d'altres- a l'anàlisi de l'acció col·lectiva, seguint la perspectiva de les solucions «externes». La tercera qüestió es centra en els problemes de connexió entre els models de la teoria de jocs -amb una aplicació limitada de la hipòtesi de racionalitat- i la «situació d'elecció»; com hom pot sospitar, darrera de la noció de «situació d'elecció» cal situar-hi diversos enfocaments de teoria social. Aquest és un problema que queda al marge de la tesi. Tanmateix, sí es pot afirmar que, en el seu conjunt, les ciències socials encara estan molt lluny de formular teories integrades amb suficient rigor, capacitat analítica i consistència per a tot tipus de «situacions d'elecció».

En la mesura que la racionalitat és l'únic supòsit de comportament, això obliga

a la teoria de jocs a resoldre tots els problemes de l'acció col.lectiva mitjançant una sofisticació -cada vegada més allunyada de les possibilitats d'ésser assumits com a supòsits de comportament real- del càlcul racional interactiu. Però, com s'ha argumentat detalladament a la primera part de la tesi, existeixen possibilitats analítiques perquè el model estàndard de racionalitat no sigui completament hegemònic. Seguint les discussions dels dos primers capítols, a la secció 4.4. es plantegen els problemes d'adoptar diverses línies de transformació de la hipòtesi de racionalitat a la teoria de jocs aplicada a resoldre el problema de l'acció col.lectiva. Sense ésser la «teoria» de l'acció col.lectiva, la teoria de jocs és l'eina fonamental per a modelitzar els processos d'acció col.lectiva, i no tant la hipòtesi de racionalitat, que a través dels models concrets d'interacció, ha de prendre la seva plaça en el marc de successives aproximacions a una teoria general de l'acció col.lectiva.

### **4.3. Béns públics i lògica interactiva: estratègies de modelització.**

Aquesta secció està dedicada a analitzar com els models de la teoria de jocs poden ajudar a interpretar els fenòmens d'acció col·lectiva, entesos com els processos de provisió voluntària i col·lectiva de béns públics. S'argumenta que independentment de la forma com es realitzin les eleccions, les estructures d'interacció formalitzades a través de la teoria de jocs poden mostrar algunes propietats dels béns públics, en relació al seu consum o a la seva producció, com per exemple la no-exclusivitat o la rivalitat. Així, segons la natura del bé públic, el resultat del conjunt d'eleccions individuals pot ser molt diferent, tot i que els jugadors escullin la mateixa estratègia. Alhora, es pot suggerir que al darrera de diferents estructures d'interacció poden haver també «situacions d'elecció» distintes, que permetin explicar les característiques institucionals dels béns col·lectius que els individus desitgen, tot i que per les especificitats físiques d'aquests béns no sigui possible fer distincions. Un punt a destacar és que tradicionalment les estructures d'interacció modelitzen el problema de la provisió de béns públics a partir del supòsit bàsic d'una orientació del comportament basat en l'interès propi. Però si hom pressuposa l'existència d'altres orientacions d'interès, cal pensar que per un mateix bé col·lectiu, l'estructura d'interacció que s'articula per la seva provisió ha de ser necessàriament distinta. Deixant el desenvolupament d'aquestes qüestions per a la següent secció (4.4.), en aquesta es manté el supòsit d'interès propi, per concentrar l'atenció a les distintes possibilitats de simplificació dels jocs que modelitzen els processos d'acció col·lectiva.

#### **4.3.1. No-unicitat en la interpretació de l'acció col·lectiva.**

En primer lloc, cal assenyalar que existeixen diverses propostes de simplificació que intenten resoldre els problemes d'aplicació de la teoria de jocs. Bona part dels

esforços dels respectius defensors d'aquestes propostes han estat encaminats a reclamar l'exclusivitat de la seva simplificació com a forma bàsica d'analitzar el problema de l'acció col·lectiva. Tanmateix, aquests esforços han estat en gran part improductius, degut a què la definició d'acció col·lectiva no inclou cap criteri per a discriminar entre les diverses simplificacions dels models que representen resultats d'eleccions. En tot cas, els criteris s'han d'extreure de l'anàlisi de les situacions d'elecció existents en diferents processos d'acció col·lectiva.

La proposta que aquí es realitza és inversa. Assumint la impossibilitat de bastir una teoria general sobre l'acció col·lectiva exclusivament en termes de teoria de jocs, es considera que els esforços realitzats amb la finalitat de modelitzar amb teoria de jocs formes distintes d'interacció voluntària en presència de béns públics no són inútils, però sí que cal situar aquests en la seva posició correcta dins el que hom podria considerar com l'anàlisi teòrica de l'acció col·lectiva. En aquest sentit, convé recordar que els models de teoria de jocs tracten essencialment d'explicar els resultats d'eleccions, i no les «situacions d'elecció»<sup>23</sup>. Sobre aquestes darreres, les aplicacions actuals dels models tan sols assumeixen supòsits molt simplificats sobre racionalitat, interès i adquisició d'informació. Però realment, no es pot pretendre que donin raó del conjunt de variables que intervenen en una «situació» d'acció col·lectiva. Així, aquests supòsits simplificadors es poden entendre també com indicadors que serveixen per a establir una

---

<sup>23</sup> En referència a una teoria general de l'acció col·lectiva, es pot considerar que aquesta serà menys formalitzada, però més completa per explicar les «situacions d'elecció» degut a què caldria introduir moltes altres variables. En aquest sentit, cal afegir que *"cooperation is not a simple 'additive' effect of various conditions with regard to interdependencies, but the result of a relative complicated, though explicitly specifiable 'interaction' of these conditions [...] if they are implicit in and can be derived from a more general theory (...) the explanation of cooperation in problematic social situations has more 'depth' in that it specifies the mechanisms which lead from certain features of strategic interdependencies to cooperative behavior and efficiency"* (W. Raub i Th. Voss, "Conditions for Cooperation in Problematic Social Situations", a A. Diekmann i P. Mitter (Eds.), *Paradoxical Effects of Social Behavior*, Springer Verlag, Heidelberg-Viena, 1986, pàgs. 99-100). El problema d'endogenitzar les variables dins el model de resultats d'elecció és quelcom que provoca força dificultats i distorsions quan es vol fer de forma aïllada un enfocament més general sobre la cooperació.



connexió entre l'anàlisi del resultat de les eleccions, i les interpretacions -específiques o generals- sobre l'orientació d'interès, els contextos institucionals o la formació de decisions. Sobre els resultats de les eleccions, acceptant ja d'entrada que aquests només poden constituir una part de l'explicació del fenomen de l'acció col.lectiva, cal analitzar com s'ajusten les diverses modelitzacions proposades a les diferents formes en què fenòmens d'acció col.lectiva poden ser possibles.

Com s'ha destacat al capítol anterior, es poden interpretar diferents situacions d'acció col.lectiva a partir de les característiques del bé públic que els individus volen aconseguir. Però apart de les característiques del bé públic, hi ha altres variables, relacionades amb el nombre i condicions dels individus que participen en la interacció, que també es poden incorporar -sense excessius problemes- en la modelització per a analitzar el resultat de les eleccions en una acció col.lectiva (en termes de teoria de jocs). La incorporació d'aquestes variables porta a què augmentin les possibilitats de modelització. Dins aquesta dinàmica, cal tenir sempre present que allò que interessa especialment és mantenir la màxima simplificació possible que permeti el seu tractament analític sense desestimar les característiques bàsiques. Perquè el model general, que podria integrar hipotèticament tots els casos possibles és inaplicable; però a més, moltes vegades també és innecessari.

*La forma més senzilla que permet representar les propietats dels béns públics (que reflecteix la interacció entre individus encaminada a la seva provisió) és l'elaboració de models de dos jugadors, en els quals es suposa que cada jugador disposa de dues estratègies. Per aplicar a les situacions convencionals d'acció col.lectiva és necessari generalment desenvolupar models més complexes, que considerin l'existència de més de dos jugadors en la interacció. Mantenir models de dos jugadors i dues estratègies només pot servir per presentar situacions idealitzades que no permeten recollir la totalitat dels efectes que apareixen en els models que, per exemple,*

incorporen com una variable més el nombre de jugadors que intervenen en l'acció col·lectiva. Hi ha diverses possibilitats de modelització d'un procés d'acció col·lectiva en termes de teoria de jocs, però en darrer terme totes són formes distintes de resumir i simplificar l'enorme complexitat del que es podria anomenar els models «fotogràfics» d'una situació d'acció col·lectiva. La composició d'aquesta forma general tindria 'N' jugadors i 'N' estratègies, i les orientacions dels jugadors podrien ésser distintes, així com les estratègies, que també serien distintes per cada jugador (pels casos asimètrics d'interès i racionalitat, vegeu la secció 4.4.). Això produiria valoracions dels resultats diferents per a cada jugador i ordres de preferències (o funcions d'utilitat) individualitzades. Només la definició i presentació de l'esquema d'interacció ja significaria un esforç gairebé inabastable, perquè caldria elaborar de forma aproximada un mapa cognitiu de cada individu implicat en l'acció col·lectiva.

La definició de les condicions d'informació que disposen els jugadors -com 'input' bàsic per a l'elecció d'una estratègia- deixa també obertes diverses possibilitats analítiques. L'observació empírica sobre la forma d'elecció d'estratègies podria conduir a escollir alguna fórmula entre d'altres, igual que la relació amb la grandària també podria possibilitar -segons els casos- establir supòsits plausibles sobre la capacitat dels agents d'obtenir i processar la informació sobre les estratègies dels altres<sup>24</sup>. En tot cas, la variabilitat de jugadors, estratègies i informació hauria de conduir a l'anàlisi individualitzada dels equilibris possibles en cada interacció i, posteriorment, en el cas d'existir més d'un equilibri, a introduir en el tractament dels models fórmules per discriminar -endògena o exògenament- entre els equilibris possibles, aquell equilibri que previsiblement s'adopti.

---

<sup>24</sup> Tant pel cas referit al coneixement de les estratègies escollides en les jugades anteriors -en un joc iterat- com en el cas de conèixer les estratègies que escolliran els altres en el moment present (que implica disposar de la informació sobre les preferències dels altres en cada moment i de les seves regles de càlcul).

Degut a les impossibilitats pràctiques i teòriques d'aplicar aquesta forma general a la teoria de l'acció col·lectiva, el que cal és cercar fórmules de simplificació. En aquest sentit, es poden destacar tres grans línies de propostes, totes parcialment desenvolupades fins al moment, amb alguns intents d'aplicació empírica, que poden ser tractades de forma més global i sistemàtica. La perspectiva defensada és que no hi ha una fórmula de simplificació superior a una altra, ja que totes resten molt lluny de la «forma general». En tot cas, la superioritat d'una o altra proposta es limita a ser més o menys aplicable i adequada pel cas d'un determinat procés d'acció col·lectiva.

#### 4.3.2. L'aplicació del cas particular d'eleccions binàries entre dos jugadors.

La primera línia de simplificació possible consisteix en els models de jocs més elementals, els quals presenten dos jugadors i dues estratègies per cada jugador. Mantenint els supòsits convencionals de la teoria de jocs -interès propi i racionalitat- apareixen algunes de les distincions bàsiques dels béns públics, a través de diferents models d'interacció típics, com es pot observar en el quadre I (canviant els supòsits de comportament sorgirien altres models d'interacció). El significat d'aquests jocs il·lustra la conducta dels individus per la producció del bé col·lectiu segons la pròpia natura del bé en què estiguin interessats. Creuant dues característiques del bé col·lectiu sorgeixen quatre possibles models d'interacció diferents. Aquestes característiques són l'existència de rivalitat (o no) en el consum del bé, i la forma de la seva provisió, si aquesta es realitza de forma contínua o bé de forma escalar (en quantitats determinades, s'aconsegueix una unitat addicional dels producte)<sup>25</sup>.

---

<sup>25</sup> La classificació proposada s'inspira en els models de dos per dos proposats per M. Taylor i H. Ward, (a "Chickens, Whales, and Lumpy Goods: Alternative Models of Public-Goods Provision", *Political Studies*, vol. 30, 1982, en especial pàgs. 351-354). Cal advertir que tercers factors no considerats poden modificar l'estructura dels jocs que s'assenyalen (per comentaris en aquest sentit, vid. J. Hampton, "Free-Rider Problems...", art. cit., 1987).

Quadre I

	No-rivalitat	Rivalitat
Provisió contínua	Joc del «Gallina»	Dilema del Presoner
Provisió escalar	Joc de l'Heroi	Joc de la Assegurança

El primer cas, quan existeix la provisió contínua d'un bé col·lectiu i la no rivalitat en el seu consum, pot ser formalitzat gràcies al joc del «gallina». Aquest joc il·lustra una situació en què la producció del bé col·lectiu pot ser realitzada per un dels dos jugadors, no essent necessària la participació de tots dos. Però el bé es produeix mentre es participa, i desapareix quan s'acaba la participació. En aquest sentit, el joc mostra l'existència d'un problema de coordinació -que pot ser intertemporal-, ja que el jugador que no participa s'aprofita de l'esforç de l'altre jugador. Si cap jugador participa en la producció del bé col·lectiu -esperant que l'altra ho faci-, aquest bé pot restar sense produir-se indefinidament. En aquest joc, pot ser que la participació sigui preferida a la no participació, però només quan l'altre jugador no participa<sup>26</sup>. El segon cas és el joc de l'heroi. Aquest joc difereix del joc del gallina perquè serveix per

<sup>26</sup> Una major precisió sobre aquest joc i les seves variants en les condicions assenyalades es pot trobar a I. Lipnowski i S. Maital, "Voluntary Provision of a Pure Public Good as the Game of 'Chicken'", *Journal of Public Economics*, vol. 20, 1983.

modelitzar una situació en què el bé públic és produït mitjançant una sola acció. L'heroi o voluntari no pot treure cap compensació del seu acte, del qual se n'aprofita l'altre jugador. El dilema del presoner s'aplica a la producció de béns que requereixen una provisió contínua i els quals presenten rivalitat en el seu consum. Per aquest motiu, tot els consumidors han de participar en la producció de bé. Quan no ho fan i intenten adoptar l'actitud de «free-rider», la producció contínua del bé col·lectiu depèn del nombre de jugadors que coincideixin en aquesta actitud (es pot produir un procés de regulació per ajustar el nombre mínim de cooperadors necessari). Però si només es modelitzen dos jugadors, cal inferir que el joc no es produiria, perquè la no participació és preferida a la participació, independentment del que faci l'altre jugador. El cas del joc de l'assegurança no deixa lloc als jugadors per actuar com a «free-rider», ambdós prefereixen la participació, independentment de l'actitud del contrari. Com que els jugadors són conscients de la seva estructura d'interacció, en produir-se el bé col·lectiu en un sol moment i existir rivalitat en el consum, només es produiria si tots els jugadors participen en la producció d'aquest bé.

La limitació a dos jugadors en la interpretació d'una situació d'acció col·lectiva implica l'existència de dues possibilitats analítiques: considerar que és un joc d'un individu davant tots els altres (que actuen igual), o considerar que les interaccions són sempre per parelles, dins una població més àmplia. En relació al primer cas, és fàcilment acceptable per determinades situacions el supòsit d'un comportament individual basat en la creença subjectiva que la resta del món té unes preferències similars, tot i que cal analitzar amb més detall si l'estructura del comportament agregat es manté també amb la mateixa lògica. Per exemple, un conductor a la carretera pot escollir el carril a seguir, coneixent les conseqüències de cada opció, perquè suposa que els altres jugadors tenen preferències estables i coincidents. En el cas de poder afirmar que tots els individus implicats en aquesta acció col·lectiva segueixen la mateixa conducta, seria possible fer prediccions sobre els resultats de la interacció. És possible

que per algunes situacions estables sigui factible establir un mateix model d'interacció subjectiu, i que la lògica de la grandària no hi tingui cap influència<sup>27</sup>. En tot cas, per mantenir aquesta modelització, la més simple entre les existents, cal acceptar altres supòsits restrictius, com l'existència d'estratègies binàries o la igualtat dels ordres de preferències de tots els agents. Si la situació d'acció col·lectiva que s'analitza no pot acceptar tots aquests supòsits, cal emprar alguna de les altres fórmules de modelització. Un camp molt desenvolupat dins aquesta forma de modelització ha estat el dels jocs iterats, que permeten l'anàlisi de situacions que es perllonguen en el temps. Les condicions d'aparició i estabilitat de la cooperació en aquests casos s'han analitzat especialment entorn l'anàlisi de l'estructura del dilema del presoner i les condicions per l'aparició de punts d'equilibri estables i únics<sup>28</sup>.

Els jocs que modelitzen les interaccions per parelles són la segona possibilitat analítica en les simplificacions de dos jugadors amb dues estratègies. Si es donen les condicions per suposar tal procediment en una situació d'acció col·lectiva (com per exemple les relacions de parelles en jocs com el tennis o els escacs, o un avançament entre dos vehicles en una carretera), poden aportar una visió més plausible d'aquestes interaccions de dos jugadors, degut especialment a què es suposa que les parelles poden anar canviant, fins arribar al cas hipotètic que tothom s'ha relacionat amb tothom en una població. No obstant això, les interaccions per parelles tenen una utilització més

---

<sup>27</sup> En aquest sentit, el model pot funcionar tant amb una estructura de preferències ordinal com amb una cardinal, però en aquest darrer cas, les intensitats s'han de suposar idèntiques per a tots els jugadors. R. Hardin fa una proposta en termes d'individu contra col·lectiu, a "Collective Action as an...", art. cit., 1971. Una proposta similar referida al dilema dels comuns (un dilema del presoner on es tracta de 'prendre', i no de 'donar') es pot trobar a H.V. Muhsam, "An Algebraic Theory of the Commons", a G. Hardin i J. Baden (Eds.), Managing the Commons, Freeman and Co., San Francisco, 1977.

<sup>28</sup> J. Sobel, "Utility Maximizers in Iterated Prisoner's Dilemmas" (a R. Campbell i L. Sowden (eds.), Paradoxes of Rationality and Cooperation, University of British Columbia Press, Vancouver, 1985); i K.-D. Opp, "The Evolution of a Prisoner's Dilemma in the Market" (a Paradoxical Effects... op. cit., 1986); aquest darrer és també un intent de vincular el model amb una teoria més amplia de la cooperació, basada en els supòsits de l'elecció racional. Vid. també sobre el tema l'apartat 2.2.4.

destacada quan junt al canvi de parelles, s'admet l'existència d'iteracions en el joc (una seqüència de repeticions entre la mateixa parella). En aquests casos es pot considerar que els actors disposen de diverses «regles estratègiques» per tal de comportar-se al llarg de la iteració, i que es produeix una competició entre «regles estratègiques», perquè cada jugador desitja millorar el seu resultat seguint una determinada regla contra les altres regles. Aquesta modalitat de jocs iterats de cooperació per parelles de dos jugadors, ha estat inicialment desenvolupada en la seva aplicació a les ciències socials per R. Axelrod<sup>29</sup>. Inspirant-se en l'obra de Maynard Smith sobre l'aplicació de la teoria de jocs a la biologia evolutiva (vid. apt. 2.2.3.), Axelrod formula una anàlisi de diverses «regles estratègiques» en un context de dilemes de presoner iterats, en competicions de dos jugadors amb dues estratègies. La competició s'estableix entre les estratègies seqüencials estables que guien el comportament dels jugadors al llarg de la iteració. Les estratègies que escullen els jugadors a cada jugada són la «cooperació» o la «no cooperació». Sobre l'elecció d'una d'aquestes, apart de la «regla estratègica» adoptada, influeix de forma important -com a tot tipus de jocs iterats- la taxa de descompte del futur que apliquen els jugadors. L'anàlisi d'Axelrod i dels seus continuadors es concentra en investigar la competició entre «regles estratègiques», per a cercar l'aparició de regles col·lectivament estables, o sigui, regles de comportament que no poden ésser superades (en termes de resultats per l'individu que les aplica al llarg de la iteració) per altres regles presents en la competició. Seguidament, el supòsit que es considera és que les regles col·lectivament estables poden ésser adoptades, com un fenomen de difusió del comportament exitós, entre la població d'individus que competeixen<sup>30</sup>.

---

<sup>29</sup> Vid. d'aquesta autor, "The Emergence of Cooperation among Egoists", *American Political Science Review*, vol. 75, 1981; i el seu llibre *La evolución de...*, op. cit., 1986.

<sup>30</sup> R. Axelrod, *La evolución de...* op. cit., 1986, pàgs. 139-180.

L'èxit del model d'Axelrod i la seva regla estratègica preferida, l'«ull per ull»<sup>31</sup>, ha portat a què posteriorment s'elaboressin models similars de competició per parelles amb altres estructures d'interacció, diferents al dilema del presoner, com el joc del «gallina»<sup>32</sup>. Mantenint el supòsit d'interès propi, s'han investigat també els efectes d'alterar l'estructura de pagaments per cada un dels diferents jocs, per tal de comprovar la fermesa d'una determinada regla estratègica. Els resultats han mostrat que les regles amb èxit depenen en gran manera de les estructures d'interacció, i que el seu grau d'èxit depèn de variables com la distància entre guanys i pèrdues, la taxa de descompte del futur o la durada de les repeticions. Com a conclusió, s'ha deduït que no és possible proposar regles universals de cooperació vàlides en totes les situacions, o inclús referides a un mateix joc<sup>33</sup>. Una possibilitat a explorar seria proposar regles de validesa general per a categories ben determinades de béns públics, que en la seva definició, a més de la seva estructura d'interacció ideal, incorporessin com a variables de la «situació d'elecció» el descompte de futur o la quantificació de les pèrdues de la no cooperació.

---

<sup>31</sup> Axelrod defensava aquesta regla estratègica (en anglès «tic for tac») com una constant universal de comportament per afavorir la cooperació. Estudis posteriors han mostrat que les condicions perquè aquesta regla triomfi no són tan generals inclús dins l'estructura del dilema del presoner (J. Hirshleifer i J.C. Martínez Coll, "What Strategies Can Support the Evolutionary Emergence of Cooperation?", Journal of Conflict Resolution, vol. 32, 1988).

<sup>32</sup> Vid. B.L. Lipman, "Cooperation among Egoists in Prisoner's Dilemma and Chicken Games", Public Choice, vol. 51, 1986; i J.C. Martínez Coll i J. Hirshleifer, "The Limits of Reciprocity: Solution Concepts and Reactive Strategies in Evolutionary Equilibrium Models", Rationality and Society, vol. 3, 1991. Dels mateixos, "Selección y mutación en juegos evolutivos", Ms., 1991. En aquestes propostes no s'ha considerat directament la seva possible aplicació a la relació entre bens públics i acció col·lectiva.

<sup>33</sup> En aquest sentit, s'ha afirmat que "no se pueden establecer proposiciones en el sentido de considerar OJO POR OJO una especie de estrategia universalmente superior. Cualquier valor de supervivencia que posea OJO POR OJO, está limitado al contexto del dilema de los presos." a J. Hirshleifer i J.C. Martínez Coll, "Selección y Mutación...", art. cit., 1991, pàg. 22. Dins el joc del Dilema del Presoner, tampoc aquesta estratègia és sempre dominant, com diu Ch. Donninger, "there is no single winning principle in the iterated Prisoner's Dilemma. It depends heavily on the payoff structure, external influences and of course from the set of competing strategies" (a "Is it Always Efficient to be Nice?. A Computer Simulation of Axelrod's Computer Tournament", a A. Diekmann i P. Mitter, (Eds.), Paradoxical Effects...., op. cit., 1986, pàg. 134).



En relació a la teoria de l'acció col·lectiva, la proposta d'Axelrod pot ésser aplicada per a interpretar l'expansió de regles (les estratègies que apliquen els individus al llarg de tota la interacció) que facilitin o dificultin la producció de béns públics de forma voluntària, al mateix temps que per analitzar com unes determinades regles dominen i s'estenen sobre d'altres. No obstant això, l'aplicació del model d'Axelrod és bastant limitada, ja que l'interès que aporten les seves explicacions teòriques no inclou els béns públics que han de ser produïts conjuntament per 'N' persones, la situació d'elecció més general d'acció col·lectiva, i només es pot aplicar als casos de béns públics, sempre que la provisió es realitzi de forma descentralitzada, amb interaccions de dos a dos<sup>34</sup>. La limitació d'aquests models és deguda a que només poden integrar els béns públics que presenten un procés de producció que és possible de representar mitjançant parelles de jugadors.

Deixant els models de dos jugadors, que serveixen per a explicar alguns resultats d'elecció en situacions d'acció col·lectiva de forma força simple -i, per tant, recomanable quan sigui suficient-, apareixen els models que consideren 'N' jugadors. Aquests models són els que es poden aplicar a la majoria de situacions d'acció col·lectiva, però presenten el problema de generar una excessiva complexitat, que impossibilita l'anàlisi aplicada. Per aquest motiu, cal cercar les estratègies de simplificació possibles, i determinar en quins casos és més adequada cadascuna d'elles. Les diferències específiques entre els models de dos jugadors i els de 'N' jugadors es poden trobar bàsicament en els següents factors:

i) l'efecte d'una defecció o una col·laboració individual en els models de 'N' jugadors

---

<sup>34</sup> Com un exemple d'aquests casos es podria considerar la propagació de rumors. Un nivell superior d'informació pot ésser el bé públic en qüestió, i les estratègies a considerar serien la distorsió o no de la informació que es transmet. (vid. M.L. Rouquette, *Les Rumeurs*, PUF, Paris, 1975). Segons el tipus d'informació, el model d'interacció hauria de canviar i, vinculat amb aquest, distintes «regles estratègiques» podrien tenir major o menor predomini.

queda difuminada entre els comportaments de molts altres jugadors, i no concentrada en un de sol, com en els jocs de dos jugadors. Com a conseqüència, en el joc de 'N' jugadors, els efectes de la conducta individual resten relativament anònims, a no ser que s'apliquin esforços per contrarestar aquesta tendència;

ii) cap jugador té una influència determinant sobre el comportament dels altres (no poden imposar-se regles estratègiques dominants, si no existeixen prèviament) i generalment tampoc sobre el resultat final de la interacció (l'única excepció seria en el cas d'existir una gran asimetria de recursos). Això és degut a que el resultat és determinat per moltes decisions individuals, en les que es pressuposa algun nivell de consciència sobre la situació d'interdependència.

iii) contràriament als models anteriors, en les interaccions amb 'N' jugadors sí poden analitzar-se realment els resultats d'eleccions que es deriven de situacions simultànies de cooperació i explotació, amb la qual cosa apareix un ampli conjunt de relacions possibles, tant segons el respectiu nombre de cooperadors i d'explotadors, com sobre les condicions d'informació i d'equilibri del joc (vid. apt. 4.3.3.).

*Excepte la tercera, que sí té implicacions substancials, aquestes diferències només deriven de la influència que el nombre de jugadors exerceix sobre la transmissió d'informació: en augmentar el nombre de jugadors, és més difícil -en general- reconèixer quins són els que no participen. En aquest sentit, la influència d'una determinada situació d'elecció podria contrarestar les tendències esmentades.*

#### **4.3.3. Els models convencionals d'eleccions binàries amb 'N' jugadors.**

La segona línia de simplificació per modelitzar l'acció col·lectiva consisteix en no limitar el nombre de jugadors, però sí mantenir el supòsit que cada jugador només pot emprar dues estratègies. En aquestes condicions, la formulació més simple és

assumir que tots els jugadors escullen unànimament una de les dues estratègies que tenen al seu abast. Per exemple, en una estructura del dilema del presoner, malgrat que tothom prefereixi participar, l'estratègia dominant per cada jugador és la defecció, i com a resultat la no cooperació és l'únic resultat d'equilibri previsible<sup>35</sup>. Però apart d'introduir un major nivell de realisme que els models de dos jugadors, la capacitat analítica d'aquesta formulació resta escassa o, en tot cas, es distancia poc dels models esmentats a l'apartat anterior.

Mantenint els requisits d'aquesta línia de simplificació, la possibilitat d'ampliar la capacitat analítica dels models d'elecció binària amb 'N' jugadors es va desenvolupar durant els inicis de la dècada dels setantes mitjançant els anomenats «diagrames de Schelling»<sup>36</sup>. Aquests models es diferencien dels primers perquè admeten que no tots els jugadors han d'escollir la mateixa estratègia. Així, és possible explicar com la cooperació d'una part dels jugadors pot provocar externalitats, positives o negatives, pel conjunt dels 'N' jugadors o un subconjunt d'aquests, alhora que els «pagaments» que es deriven d'escollir una o altra estratègia depenen sempre del nombre d'individus que opten per una d'ambdues estratègies. Aquests models es construeixen a partir de dues funcions de «pagaments» (ordinals o cardinals), una referida a cada estratègia, que varien els seus resultats en funció del nombre d'individus (en relació al total de participants) que escullen seguir una determinada estratègia. Cada funció assenyala uns

---

<sup>35</sup> La primera proposta en aquest sentit, referida al dilema del presoner, va ser realitzada per A. Sen (a "A Game-Theoretic Analysis of Theories of Collectivism in Allocation", dins T. Majumdar (ed.), Growth and Choice, Oxford University Press, Londres, 1969).

<sup>36</sup> Proposats per primera vegada per T. Schelling a "Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities", Journal of Conflict Resolution, vol. 17, 1973 (article reproduït de forma reduïda a l'obra del mateix autor Micromotives and... op. cit., 1978). El model de Schelling va ésser proposat també en el mateix temps i de forma independent per H. Hamburger (a "N-Person Prisoner's Dilemma", Journal of Mathematical Sociology, vol. 3, 1973). Posteriorment, aquests models han estat difosos i utilitzats molt àmpliament, com per exemple ho fa -entre molts altres- J. Elster quan exposa la seva concepció de la teoria de l'acció col·lectiva, a The Cement of... op. cit., 1989.

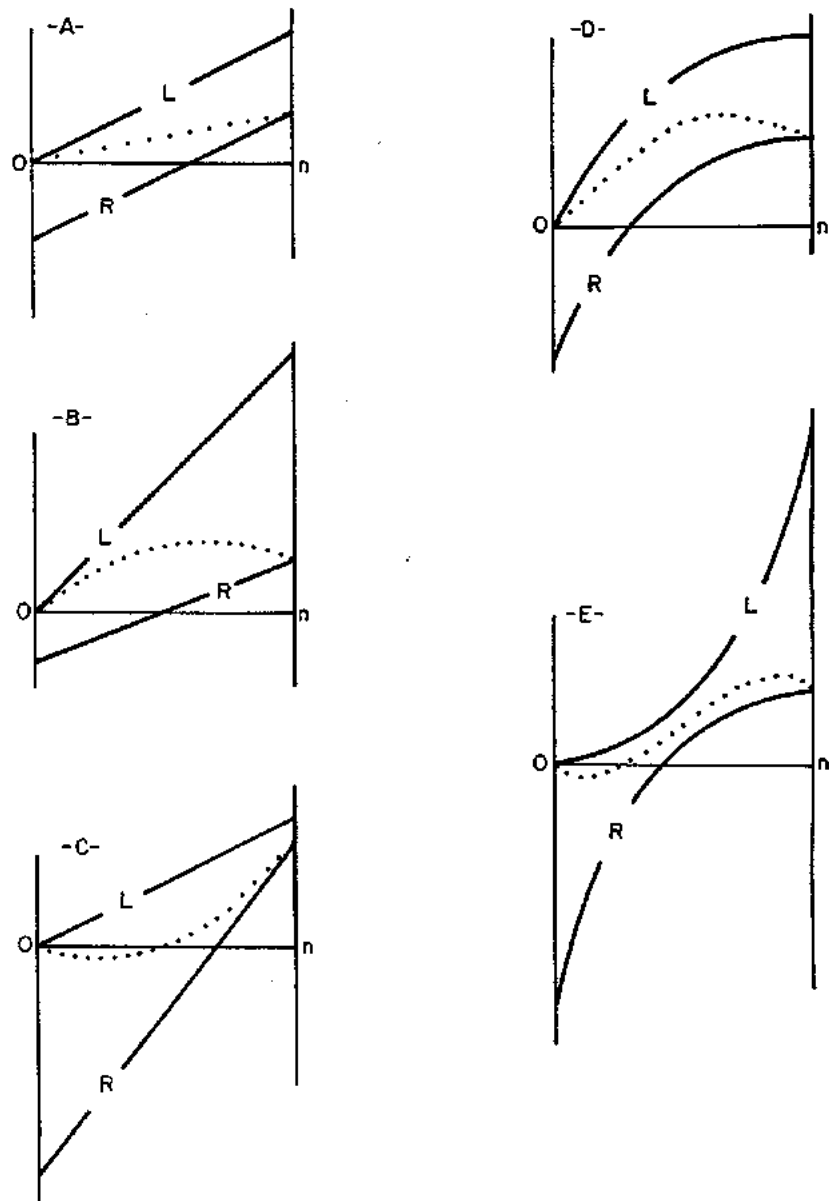
«pagaments» individuals pels jugadors que han escollit l'estratègia que representa. En aquest sentit, un dels supòsits més restrictius que han d'assumir aquests models és que tots els individus que realitzen una mateixa elecció reben també una idèntica retribució. Les primeres situacions d'acció col·lectiva que es modelitzaren amb preferències variables sobre una elecció binària foren aplicacions del dilema del presoner, a les quals la cooperació provoca externalitats positives a partir d'un nucli mínim de cooperadors. La presentació gràfica proposada per T. Schelling d'aquestes modelitzacions de resultats d'eleccions és molt suggeridora perquè permet comprendre la dinàmica de la interacció des d'un punt de vista «subjectiu». El seu principi és establir un eix de coordenades en el qual les dues funcions quedin definides pels «pagaments» i pel nombre d'individus 'N' que escullen una de les dues estratègies, la mateixa que ha adoptat l'«ego» que observa el resultat de la seva elecció<sup>37</sup>.

Per facilitar la comprensió d'aquestes representacions, es reproduïxen les gràfiques proposades per Schelling sobre variacions del dilema del presoner. Les gràfiques suposen l'existència de 'N' jugadors, cadascun dels quals pot escollir entre dues estratègies possibles ('L', no contribuir; i 'R' contribuir). Els eixos verticals dels gràfics (a dreta i esquerra) mostren una escala dels nivells de satisfacció que es poden assolir. Donades aquestes configuracions, es suposa que un individu 'N+1', per optimitzar la seva satisfacció, escull entre l'estratègia 'L' i l'estratègia 'R' a partir dels seus càlculs o supòsits sobre el nombre de jugadors que escolliran al mateix temps una i altra estratègia.

---

<sup>37</sup> T. Schelling defineix un model amb 'n+1' jugadors, però conceptualment és més clara la fórmula de 'n-1' jugadors (a "Hockey Helmets...", art. cit., 1973). Davant aquesta versió «subjectiva» de la interacció, també es pot presentar una versió «objectiva», en la qual es consideren els 'n' jugadors (sobre un total de 'N' possibles) que adopten una estratègia, però les funcions reflecteixen en aquesta versió els resultats conjunts de cadascuna d'ambdues coalicions. Més que la seva lògica interna, l'objectiu a modelitzar pot imposar un o altre estil de presentació. Per una anàlisi de les diferències entre ambdues presentacions, vegeu A. H. Carling, "The Schelling Diagram: On Binary Choice with Externalities", a *Behavioral Sciences*, vol. 32, 1987.

Gràfica 1



(Font: T. Schelling, Micromotives and... op. cit., 1978, pàg. 220)

En aquests dilemes de presoner, un jugador sempre aconsegueix més satisfacció si individualment escull l'estratègia 'L', però l'augment del seu nivell de satisfacció depèn el nombre de jugadors que escullin l'estratègia contrària, 'R'. Per exemple, si un

jugador té la informació -o la creença més o menys fonamentada- que a la gràfica B tots -o casi tots- els 'N' jugadors escolliran l'estratègia 'R', pot calcular que seguint l'estratègia 'L' els seus guanys relatius seran molt grans. Per tant, és probable que tots els jugadors pensin el mateix i que com a resultat el nivell de satisfacció per tots sigui zero (en tot cas, també depèn de si les condicions d'informació són iguals per tots els jugadors). Però si aquest jugador es troba en la situació de la gràfica C, aquests guanys relatius seran molt petits, i la decisió de seguir l'estratègia 'R' pot ser més fàcil d'adoptar, si està afectada per altres factors externs al càlcul racional. Així, les distintes funcions de satisfacció que es poden integrar en un dilema del presoner poden alterar la decisió de contribuir o no contribuir, tot i que sempre la decisió racional per beneficiar el propi interès és la no contribució.

La qüestió important en aquests models de la teoria de jocs és saber quanta gent opta per una o altra elecció, i si la distribució resultant configura un punt d'equilibri. Seguint amb el dilema de presoner multi-persona uniforme, es considera que existeix sempre un nombre 'k' de jugadors, entre 1 i 'N', els quals poden conduir a la provisió del bé col·lectiu i assolir una millor situació que si tothom hagués rebutjat la provisió del bé col·lectiu. Però d'aquest també se n'aprofiten igualment els 'N-k' jugadors que no han contribuït (aquesta és l'externalitat positiva que generen els 'k' jugadors). Si les coalicions de jugadors que contribueixen són més petites que 'k', aquestes no es podrien formar (o serien molt inestables) degut a què individualment seria més satisfactori no contribuir; si les coalicions fossin més grans, molts individus podrien sentir-se temptats de no contribuir i seguir gaudint del bé públic (perquè aquest es seguiria produint amb 'k' jugadors)<sup>38</sup>.

---

<sup>38</sup> Per aquest motiu proposa M. Taylor (a *The Possibility of...*, op. cit., 1987, pàgs. 92-94) la idea d'un joc del «gallina» insert en una estructura del dilema del presoner, ja que aquesta és la lògica que s'imposa per formar coalicions amb capacitat de contribució.

No totes les combinacions de 'k' individus que contribueixen arriben a constituir punts d'equilibri estables, ja que l'aparició d'equilibris depèn de la forma de les funcions (i pot aparèixer més d'un equilibri en una estructura d'interacció). Les formes concretes de les funcions són també les que configuren quina ha de ser la dimensió de 'k', així com el tipus de joc que s'estableix; a més, cal recordar que normalment 'k' no és un nombre fixat d'individus, sinó que aquest depèn de la grandària de la població total 'N'. En referència als béns públics, com s'ha esmentat, el joc del dilema del presoner reflecteix unes determinades propietats d'aquests béns, que també es poden trobar en aquestes modelitzacions (amb possibilitats de variació interna dins l'estructura del dilema del presoner). Altres tipus de béns públics, poden portar a altres famílies de funcions, i especialment a formes de relacions distintes entre les funcions, la qual cosa implica la possibilitat que es formin altres tipus d'equilibris<sup>39</sup>. Un intent de sistematització proposat per Th. Schelling inclou tres tipus de solucions en els jocs d'eleccions binàries i 'N' jugadors. El primer tipus de solució és el que produeix un resultat que és preferit per tots els jugadors; el segon tipus és la situació contrària, quan ningú vol el resultat a què aconduïx l'equilibri. El tercer tipus de solució el constitueixen les solucions intermèdies, que només són preferides per una part dels participants. Cadascun d'aquests tres tipus de solució poden aparèixer a través de tres categories d'equilibris diferents: i) Un equilibri únic (cooperar o no cooperar), en què tots els jugadors escullen la mateixa estratègia, ii) Un equilibri únic, en què un nombre 'k' de jugadors escull una estratègia, i la resta de jugadors escull l'altra opció, iii) Hi ha dos equilibris possibles (cooperar i no cooperar), quan tots els participants escullen la mateixa estratègia<sup>40</sup>.

La noció d'equilibri que proposa Schelling, que és adoptada generalment en

---

<sup>39</sup> T. Schelling, *Micromotives and...op. cit.*, 1978, pàgs. 224-237.

<sup>40</sup> T. Schelling, *Micromotives and...op. cit.*, 1978, pàgs. 241-242.

aquests jocs simplificats, consisteix en una aplicació de la regla d'elecció racional implícita en el teorema del minimax: quan ningú pot augmentar el seu guany individual canviant la seva elecció. Ara bé, en tractar-se d'una elecció binària, només es requereix la consistència per aplicar el principi de maximització<sup>41</sup>. Aquesta noció té també implícits alguns supòsits sobre com s'arriba a la formació de l'equilibri. No són supòsits especialment «implausibles», però sí plantegen alguns problemes. En primer lloc, es considera que tots els actors han de conèixer la forma de les dues funcions, amb els seus respectius «pagaments» segons la seva elecció del nombre de jugadors que escullin una i altra estratègia. El segon supòsit és que els jugadors han de conèixer (o estimar) quines seran les eleccions dels altres jugadors, per saber els «pagaments» que obtindran en cada resultat d'elecció, i poder decidir la seva pròpia elecció. Així, el problema per aplicar aquests jocs consisteix bàsicament en justificar les consideracions de simplificació que cal adoptar perquè els jugadors disposin de la informació completa requerida sense uns costos massa elevats.

El concepte de «sentit comú», proposat en un apartat anterior (vid. apt. 4.2.2.), pot ésser aplicat en aquest cas, almenys pel primer supòsit, sense excessius problemes. Si són 'N' jugadors, seran funcions contínues difícilment calculables, però els trets generals d'aquestes funcions poden ser intuitivament reconeguts, especialment si es tracta de situacions de producció de béns públics habituals pels participants. El segon supòsit introdueix algunes especificitats pròpies dels jocs amb 'N' jugadors, ja que elements variables, com la visibilitat de l'acció, la grandària del grup potencial, poden fer més o menys plausible que els individus aconseguixin informació per estimar les eleccions dels altres. Quan manca -total o parcialment- la informació, depèn dels supòsits de comportament i de decisió que s'incorporin des de la «situació d'elecció»,

---

<sup>41</sup> Així ho argumenta Th. Schelling, "Equilibrium occurs where there is no gain in switching from a choice of L to a R or from R to L" [R: dreta, L: esquerra] (a *Micromotives and...op. cit.*, 1978, pàgs. 226, not. 15).



que la formació d'una situació d'equilibri no es produeixi o que, en tot cas, s'hagi de transformar en una situació probabilística<sup>42</sup>. Una possibilitat addicional per a reduir el problema de la manca d'informació és considerar innecessari el coneixement de les preferències de tots els 'N' jugadors. Coneixent sols les de 'k' jugadors que coincideixin en la mateixa estratègia (a més de saber aproximadament quin és el nombre total -'N'- d'individus en el joc), en molts jocs pot ser factible arribar a un equilibri.

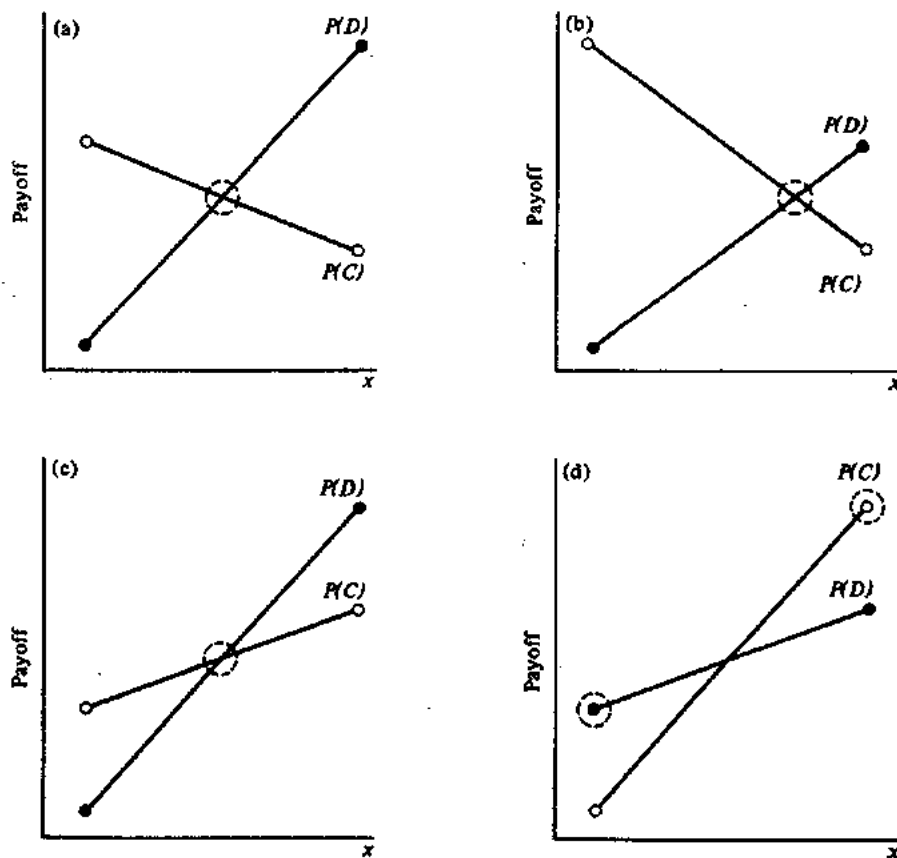
Apart del dilema del presoner, altres estructures de la teoria de jocs amb dues estratègies i 'N' jugadors uniformes han estat també formalitzats i desenvolupats pel propi Schelling i altres investigadors<sup>43</sup>. Un exemple de sistematització és el proposat per A. Colman, similar formalment als diagrames de Schelling. Aquest autor estableix quatre formules bàsiques amb funcions sense curvatures, a partir de les quals elabora la representació gràfica dels jocs multipersonals -amb dues estratègies per jugador- del «gallina», la «batalla dels sexes», l'«heroi» i la «maximització de diferències». A la gràfica 2, el cercle indica el punt d'equilibri de cada joc representat, i les dues rectes indiquen les retribucions de les actituds de contribució i no contribució per cada jugador en funció del nombre de jugadors que adopten l'estratègia contrària. No obstant això, res no impedeix que es puguin desenvolupar molts altres models d'interacció amb similars condicions (especialment si es considera la introducció de funcions contínues):

---

<sup>42</sup> L'argument que es defensa a la següent secció és que la solució a aquest problema es troba en bona part quan es disposa de la capacitat de modificar els supòsits d'orientació d'interès i de racionalitat.

<sup>43</sup> Schelling, amb distints tipus de funcions, proposa l'anàlisi de diverses propietats de béns públics que es poden representar amb aquests models, com la complementarietat o la uniformitat ( a Micromotives and... op. cit., 1978, pàgs. 234-239). Aquestes dues propietats es poden observar en funcions corbes que tinguin màxims o mínims en punts intermedis. Es considera que si eleccions no unànims aconduïxen a un mínim de la funció, el bé col.lectiu es satisfà si té la propietat d'uniformitat. Quan les eleccions no unànims aconduïxen a un màxim succeeix el fenomen invers, i apareix la complementarietat com a característica del bé col.lectiu (com en el cas d'una possible elecció del sexe dels infants).

## Gràfica 2



(Font: A. Colman, *Experimental Games...*, op. cit., 1982, pàg. 165)

Per la seva banda, A. Carling proposa set jocs elementals d'elecció binària amb preferències ordinals<sup>44</sup>. Primer, destaca que només hi ha dotze jocs diferents (amb els seus corresponents simètrics), que són les distintes formes d'ordenar les quatre possibles situacions extremes: 1. «Tots cooperen», 2. «Cap coopera», 3. «Tots cooperen menys un», 4. «Cap coopera menys un». Seguidament, afirma que d'entre els dotze jocs, n'hi ha set d'essencialment distintes, i els ordena en tres grups. El primer, és el dels jocs en

<sup>44</sup> A. Carling, "The Schelling Diagram...", art. cit., 1987.

què la cooperació (o no cooperació) és sempre preferida a la no cooperació (o cooperació) i, per tant, no és necessària la informació sobre les estratègies dels altres per l'equilibri cooperatiu, en existir una estratègia dominant per tots els jugadors<sup>45</sup>. Els altres jocs, per contra, són jocs on l'equilibri és un punt intermedi o bé existeix més d'un equilibri. En aquests darrers la defecció pot ser preferida -en ocasions- a la cooperació. Això és degut a què no existeixen estratègies dominants i, per tant, saber les estratègies dels altres té un important paper, sigui a través de processos de comunicació o d'inferència racional a partir del coneixement de les seves preferències (els jocs proposats per Colman serien d'aquest darrer grup). Com ja s'ha esmentat, resoldre els requisits d'informació constitueix un problema bàsic per aplicar aquests models en un entorn d'interacció real.

La utilització de models que permeten representar les eleccions de 'N' individus fa aparèixer aspectes que no sorgien amb els models que només n'acceptaven dos. Per exemple, si el bé és escalar, és indispensable la contribució de 'k' individus per la seva producció; mentre que si la provisió del bé és de natura contínua, encara que hi contribueixen un nombre de jugadors menor que 'k', aquest pot produir-se, tot i que els que hi contribueixen obtinguin una «retribució» menor que aquells que no ho fan. Aquesta «retribució» potser és compensada per altres aspectes, no contemplats en el model, però presents en la «situació d'elecció»; això es podria introduir en el model canviant d'orientació d'interès. Altres aspectes referits a la relació d'aquests models amb les característiques dels béns públics també han estat estudiats, considerant-se especialment la influència de la grandària del grup i els seus efectes sobre els equilibris possibles. La conclusió generalitzada és similar a la dels jocs de dos per dos: el dilema del presoner és tan sols un tipus d'interacció entre altres possibles. Especialment jocs

---

<sup>45</sup> Aquest jocs, són equivalents al procés de decisió individual davant la provisió d'un bé públic, quan els seus beneficis són sempre superiors amb la contribució a la situació de no contribució (vid. apt. 3.4.1).

com el del «gallina», o el joc de l'«assegurança», poden caracteritzar força acuradament resultats d'eleccions que es produeixen en la provisió voluntària de molts tipus de béns públics. De vegades, segons les característiques de la situació d'elecció, es poden aplicar senzills jocs de coordinació<sup>46</sup>.

Un dels aspectes remarcables d'aquests models d'eleccions binàries és que sembla més senzilla d'obtenir la provisió de béns públics escalars que la provisió de béns continus, ja que en aquest darrer cas, l'equilibri a llarg terme, com ho caracteritzen les anàlisis del dilema del presoner que parteixen del supòsit de racionalitat, és la situació de no col.laboració general<sup>47</sup>. No obstant això, aquesta situació pot estar distorsionada per l'elecció binària, ja que si els jugadors disposessin d'un continuum d'estratègies, s'hi podria introduir el supòsit de contribució inferior a l'òptima, amb la possibilitat de produir aquest tipus de bé públic, almenys parcialment. Aquest seria un fenomen distint, ja analitzat en el cas de dos jugadors, i que complicaria força més la determinació dels equilibris possibles. Una complicació dels diagrames de Schelling que s'acosta a l'anàlisi del fenomen esmentat es pot introduir considerant tres estratègies, amb 'N' jugadors; per a tal fi, cal elaborar dos diagrames connectats. Però aquest procediment ja no funciona quan el nombre d'estratègies és superior a tres, almenys la seva presentació gràfica<sup>48</sup>.

Una variant d'aquests models és la proposta de M. Granovetter sobre els models d'elecció binària de llindar. En part es tracta d'una sobre-simplificació, ja que no té en

---

<sup>46</sup> Vid. per una anàlisi general, J. Hovi, "Binary Games as Models of Public Goods Provision", *Scandinavian Political Studies*, vol. 9, 1986. En particular, sobre el joc del «gallina», vid. M. Taylor i H. Ward, "Chickens, Whales, and...", art. cit., 1982, i J. Hampton, "Free-Rider Problems in...", art. cit., 1987. Sobre el joc de l'«assegurança», vid. C. Ford Runge, "Institutions and the Free-Rider...", art. cit., 1984.

<sup>47</sup> J. Hovi, "Binary Games as...", art. cit., 1986, pàgs. 349-50.

<sup>48</sup> L'intent és realitzat per J. Hovi, "Binary Games as...", art. cit., 1986.

compte la conducta interactiva entre els jugadors respecte a les dues estratègies (vid. apt. 4.2.1.). Per aquest motiu, amb aquesta proposta es poden debilitar els supòsits d'informació, en considerar-se que cada individu pren la seva decisió en funció de les decisions que prèviament han adoptat els altres individus -seqüencialment en el temps-. Segons el nombre de cooperadors, hi ha un nivell de retribucions (o probabilitats d'èxit en la consecució del bé públic), i cada individu, coneixent aquesta dada, la valora i escull una opció -o pren actitud davant el risc que comporta<sup>49</sup>.

Sobre les possibilitats per iterar els models d'elecció binària entre 'N' jugadors, hi ha una proposta desenvolupada per M. Taylor. Aquest autor proposa una formulació de «superjoc» (o iteració indefinida) del dilema del presoner amb 'N' jugadors i dues estratègies, suposant que els jugadors tenen informació sobre totes les jugades anteriors dels altres jugadors. Així, destaca la possibilitat de formar equilibris en un dilema del presoner si els jugadors estableixen estratègies condicionals, generades a partir de la reputació (cooperativa o no cooperativa) dels diversos jugadors implicats en la interacció<sup>50</sup>. En el cas del joc del «gallina» es pot produir una situació inversa, ja que en augmentar al llarg de la iteració la reputació de «tossuderia» d'un jugador, augmenten les possibilitats de què els altres jugadors no contribueixin<sup>51</sup>. La noció d'interacció que adopta aquesta proposta no requereix un càlcul sobre les possibles estratègies dels altres

---

<sup>49</sup> Vid. M. Granovetter, "Modelos de Umbral...", art. cit., 1991 (ed. or. 1978). Com el mateix autor assenyala, "los modelos de umbral toman los dos elementos de la conducta colectiva que la teoria de juegos maneja sólo con dificultad, haciéndolos centrales: heterogeneidad sustancial de las preferencias e interdependencias de las decisiones en el tiempo" (pàg. 156). Cada actor té un vector unidimensional de llindars, però no hi ha càlcul interactiu simultani: la interdependència es «simula» a través del factor temps.

<sup>50</sup> M. Taylor proposa una sèrie d'equacions d'equilibri, en les que intervé també el descompte del futur, que pot diferir entre els jugadors. En aquest context, M. Taylor defensa que la cooperació es pot produir entre subgrups de jugadors, sigui quina sigui la dimensió del grup (The Possibility of... op. cit., 1987, pàgs. 83-106).

<sup>51</sup> H. Ward, "The Risks of a Reputation for Toughness: Strategy in Public Goods Provision Problems Modeled by Chicken Supergames", British Journal of Political Science, vol. 17, 1987.

en una jugada simultània, sinó sols la formació d'una actitud estratègica condicional a partir de la informació sobre el comportament dels altres jugadors a les jugades anteriors.

Malgrat les variacions de modelització que impliquen les diverses propostes ressenyades, resten alguns problemes per integrar completament els resultats de les eleccions en els models de teoria de jocs d'una forma plausible. El que es planteja és si hi ha factors de la «situació d'elecció» que no són possibles d'integrar en els models de la teoria de jocs. Per aclarir aquesta dificultat, cal analitzar quin és el problema: la identificació del factor i la creació d'un indicador adequat, o bé la manca d'articulació teòrica amb la teoria de jocs. Per exemple, com s'assenyala més endavant, sobre la grandària del grup, apareixen en les observacions empíriques efectes que no són explicables amb els models presentats. Es pot argumentar que això és una mostra de l'efecte de factors no introduïts en el model, i que la seva possible definició i integració tampoc és possible de realitzar afegint simplement una variable més a les funcions de decisió, sinó que requereixen una integració a partir d'un desenvolupament teòric sobre la influència en els processos de decisió del nombre de persones implicades en una acció col·lectiva.

En el dilema del presoner, amb 'N' jugadors i dues estratègies, s'ha comprovat experimentalment que la cooperació es redueix en augmentar el nombre d'individus<sup>52</sup>. Però cal pensar que aquest fenomen és degut a les propietats dels béns públics modelitzades a través d'aquest joc, i no a una característica general de tots els jocs. Precisament perquè els jocs representen propietats dels béns públics, es pot argumentar que amb altres béns públics podrien produir-se altres comportaments, que no

---

<sup>52</sup> Per exemple, així ho mostren H. Hamburger, M. Guyer i J. Fox, a "Group Size and Cooperation", *Journal of Conflict Resolution*, vol. 19, 1975; i també J. Fox i M. Guyer, a "Group Size and Others' Strategy in an N-Person Game", *Journal of Conflict Resolution*, vol. 21, 1977.

signifiquessin una reducció de la cooperació en augmentar la grandària del grup. Tornant al dilema del presoner, resta la qüestió de com explicar el fenomen que s'observa, ja que en termes formals no es pot pronosticar la reducció de la cooperació<sup>53</sup>. De les diverses explicacions que s'exposen<sup>54</sup>, potser la més prometedora és la basada en el concepte de «desindividualització», que assenyala que la identitat i la personalitat dels individus es difumina en grups poc estructurats i força desorganitzats; produint que el comportament d'aquest sigui més egoista, o menys probable que es segueixin normes socials de cooperació. Així, com assenyalen Hamburger, Guyer i Fox,

*"deindividuation in a group increases as the size of the group increases since, everything equal, an individual appears more anonymous in a large group than in a smaller one. Given this assumption, it would be predictable that the players in our seven-person groups would be the players in the three-person groups"*<sup>55</sup>

Aquest argument s'escapa de la teoria de l'«elecció racional» per a connectar-se en les interpretacions psico-socials sobre la identitat personal<sup>56</sup>. No obstant això, cal recordar que és el mateix problema que sorgia a l'anàlisi de l'obra d'Olson i la seva teoria de l'acció col·lectiva. En aquest capítol es pot formular amb major claredat en disposar d'un instrumental analític -la teoria de jocs amb 'N' jugadors- que ofereix més precisió en la identificació dels problemes de l'acció col·lectiva. Finalment, cal senyalar que s'han identificat altres factors que poden reforçar o debilitar la cooperació en grups

---

<sup>53</sup> És un fenomen similar al tractat a la secció 3.5., sobre l'augment de les actituds de «free-rider» en créixer la grandària del grup.

<sup>54</sup> Vid. Colman, *Game Theory and...op. cit.*, 1982, pàgs. 181-182.

<sup>55</sup> H. Hamburger, M. Guyer i J. Fox, "Group Size..", art. cit., 1975, pàg. 524.

<sup>56</sup> La idea de «desindividualització» sorgeix d'un article de L. Festinger, A. Pepitone i T.M. Newcomb ("Some Consequences of De-Individuation in a Group", *Journal of Abnormal and Social Psychology*, vol. 47, 1952).

interactius, entorn als quals es reproduïx el problema de la connexió entre la «situació d'elecció» i la teoria de jocs. Entre d'altres, hom pot esmentar les oportunitats per a la comunicació prèvia a l'acció<sup>57</sup>, el nivell d'anonimat o la visibilitat de les accions<sup>58</sup>, i el reconeixement de la informació sobre l'estructura del joc<sup>59</sup>. A més, amb la introducció d'aquest procediment analític es poden descobrir efectes paradoxals, com que els jugadors actuïn sempre més cooperativament en els grups més petits, inclús en els jocs que per l'eficàcia en la producció del bé col·lectiu (degut a les seves característiques), seria preferible col·laborar més quan augmenta la grandària del grup<sup>60</sup>.

#### 4.3.4. Aproximacions parcials al cas general de 'N' estratègies i 'N' jugadors.

La tercera línia és la que més s'aproxima a la forma general, ja que desenvolupa models que inclouen 'N' estratègies i 'N' jugadors, encara que estableix molts requisits

---

<sup>57</sup> Per als béns públics escalars, que impliquen jocs de coordinació o el joc del «gallina», quan la comunicació prèvia augmenta, incrementen les probabilitats de cooperació (vid. Van de Kragt et al., "The Minimal Contribution Set..", art. cit., 1983). Si el procés de comunicació és seqüencial, i no simultani, s'ha observat experimentalment que el nivell de cooperació és més elevat (cf. I. Erev i A. Rapoport, "Provision of Steep-Level Public Goods", *Journal of Conflict Resolution*, vol. 34, 1990).

<sup>58</sup> Vid. la discussió sobre aquest factor a partir de la proposta d'Hechter, a l'apartat 4.2.2.

<sup>59</sup> Un aspecte important és la informació que tenen els actors sobre les característiques del joc en què es troben, com major informació existeixi (o experiència), més probable és que es produeixi la cooperació. Però aquesta tendència no és igual per a tots els jocs, ja que és més accentuada en els jocs que més dificultats presenten per la cooperació, com el dilema del presoner (cf. Anatol Rapoport, "Experiments with N-Person Social Traps", *Journal of Conflict Resolution*, vol. 32, 1988).

<sup>60</sup> Aquesta paradoxa es podria explicar assenyalant que els subjectes tenen la tendència a generalitzar per a totes les situacions l'estructura d'interacció del dilema del presoner. En aquest, els subjectes percebeixen que els grups petits són més eficaços que els grans, i extrapolen el principi a jocs (que reflecteixen béns públics distints). Indirectament, aquesta constatació experimental mostra la importància de les variables esmentades (i la seva apreciació subjectiva) per analitzar l'actitud cooperadora dels individus (Cf. N.L. Kerr, "Illusions of Efficacy: The Effects of Group Size on Perceived Efficacy in Social Dilemmas", *Journal of Experimental Social Psychology*, vol. 25, 1989).



addicionals de simplificació per a poder tractar els models. Aquesta forma de modelització, més complicada formalment, només s'imposa quan no és possible de representar la contribució que realitzen els jugadors de forma binària (o triària en tot cas), sinó que per la natura del bé col·lectiu, aquesta contribució ha de ser formalitzada de forma contínua. Per exemple, quan hi ha més de dos jugadors i s'ha de precisar cardinalment la contribució a realitzar en quantitats monetàries o en quantitats d'esforç físic a realitzar.

La primera proposta per modelitzar el problema de l'acció col·lectiva amb un model de 'N' jugadors i 'N' estratègies fou realitzada per R. Hardin<sup>61</sup>. En realitat, el seu model no considera 'N' estratègies, sinó tantes com el nombre de jugadors que existeixi. Però cada estratègia és binària: es decideix separatament si cooperar o no cooperar amb cada jugador determinat. Aquesta és la diferència específica amb els models anteriors de 'N' jugadors amb dues estratègies, en els quals la decisió de cooperar o no cooperar d'un jugador afectava a tots els altres jugadors indiscriminadament. En aquest sentit, es pot entendre millor aquest model com una generalització dels diagrames de Schelling. Hardin proposa com a model del problema de l'acció col·lectiva un dilema del presoner jugat una sola vegada; i assenyala que encara que l'estratègia dominant és la no contribució, si la intensitat dels possibles beneficis és elevada, es poden produir solucions cooperatives. La proposta de Taylor<sup>62</sup> es diferencia en força aspectes de la de R. Hardin. Primer, Taylor insisteix que la provisió de béns públics pot ésser modelitzada a través de diversos jocs, o inclús de combinacions de jocs. Destaca especialment la rellevància del joc del «gallina» per moltes situacions d'acció col·lectiva, i considera que el dilema del presoner és sols una

---

<sup>61</sup> R. Hardin, "Collective Action as an ...", art. cit., 1971.

<sup>62</sup> Realitzada primer en el seu llibre Anarchy and Cooperation (Wiley, London, 1976), i ampliada posteriorment en l'edició revisada del llibre The Possibility of Cooperation.

possibilitat entre d'altres<sup>63</sup>. Però en relació al cas del dilema del presoner, Taylor argumenta que el supòsit d'Hardin -el joc només es realitza una vegada- és poc plausible, ja que no introdueix el factor temps. Per aquest motiu, Taylor elabora un model de dilema del presoner amb 'N' jugadors i dues estratègies, similar al d'Hardin, però amb la diferència de ser iterat. El model ja s'ha descrit anteriorment, perquè és una variació dels jocs binaris de 'N' jugadors<sup>64</sup>. Taylor modifica la proposta d'Hardin en un altre sentit. En relació a les estratègies, introdueix la possibilitat de què aquestes siguin contínues (o quasi-contínues) per a cada jugador. Amb aquesta modificació, el model que proposa Taylor pot ser aplicat per analitzar la formació de béns públics en els quals la contribució individual de béns privats (com el diner) és completament divisible<sup>65</sup>. Això no obstant, en el model el comportament estratègic s'infereix de les propietats del bé públic, i no és pròpiament integrat, perquè no es considera la capacitat estratègica dels altres jugadors, cosa que implica que el model no reflecteixi realment la interacció entre els jugadors.

En aquests models més complexos, també es pot observar com un augment de

---

<sup>63</sup> H. Ward desenvolupa una modelització similar a la de Taylor pel cas del joc del «gallina», a "The Risks of a Reputation...", art. cit., 1987.

<sup>64</sup> Els seus supòsits són que els jugadors escullen les estratègies simultàniament a cada jugada, i només tenen la informació sobre les estratègies escollides per tots els actors en les jugades anteriors. Taylor destaca que la cooperació pot sorgir a través d'estratègies condicionals, basades en la història de les interaccions anteriors. En aquest model la possibilitat de cooperació condicional (independentment de l'estratègia concreta que s'adopti) depèn de la taxa de descompte del futur en els jocs iterats. Si aquesta és suficientment baixa, al llarg de la iteració es pot produir la cooperació (The Possibility of... op. cit., 1987, pàgs. 68-81).

<sup>65</sup> Taylor, The Possibility of... op. cit., 1987, pàgs. 49-56. El model assumeix una funció de transformació entre la producció de bé públic i la contribució del bé privat individual, considerant donades les contribucions dels altres. La contribució individual és fixa en el punt de coincidència entre la funció d'utilitat individual i la funció de transformació. Diferents tipus de jocs poden formular-se, en relació als béns públics, modificant les funcions de transformació. Per exemple, un cert grau d'indivisibilitat es pot introduir en modificar-se la funció de transformació en augmentar el nombre d'individus, o la rivalitat apareix amb la modificació de les funcions d'utilitat quan augmenta el nombre de jugadors. També H. Ward aplica el mateix model al cas del joc del «gallina» ("The Risks of a Reputation..", art. cit., 1987, pàgs. 38-45).

la grandària del grup provoca un increment de la tendència dels seus membres a la defecció. Tanmateix, sorgeixen igualment problemes per establir una explicació d'aquest fenomen a través del model. Taylor argumenta que és més difícil mantenir les estratègies basades en la condicionalitat quan augmenta el grup, degut a l'increment del cost de la informació sobre la reputació dels altres jugadors<sup>66</sup>. Un exemple de variables provinents de la «situació d'elecció» que s'intenten integrar en els models és la introducció -com una regla auxiliar- de la possibilitat d'exclusió d'individus dins un grup. Si existeix aquesta capacitat d'«imposició forçada» en una acció col·lectiva, la seva aplicació afavoreix que un bé públic que presenti rivalitat pel seu consum, modelitzat a través d'un dilema del presoner, sigui més factible d'aconseguir a través de la cooperació<sup>67</sup>.

Aquesta tercera línia de simplificació, força heterogènia, és la que presenta un menor desenvolupament, i gairebé es pot senyalar que sols s'han estudiat alguns punts parcials o certs aspectes específics, no contemplats, per la seva complexitat, en les altres aproximacions. Malgrat que un desenvolupament d'aquests models constituiria l'aproximació més propera a una possible interpretació general de l'acció col·lectiva, no vol dir que sigui aquesta la via per aproximar-se a formular una teoria general (ja que sols es tracta de possibilitats formals, però no d'un desenvolupament teòric). En realitat, la recerca d'aquesta teoria general no té massa sentit. El que realment té importància és l'adequació entre els fenòmens observats i el model proposat per a

---

<sup>66</sup> Taylor no modelitza l'efecte del cost de l'informació en el seu model, deixant la qüestió oberta (a The Possibility of...op.cit., 1987, pàg. 105). Una possible aproximació alternativa seria considerar que a partir d'un determinat cost de la informació, els individus apliquen altres processos de decisió que no estan correlacionats amb l'augment de la complexitat informativa (vid. secció 2.2.).

<sup>67</sup> D. Hirshleifer i E. Rasmusen modelen aquesta situació, amb els supòsits convencionals, introduint la possibilitat d'ostracisme per als individus que no cooperen. Amb el seu model, argumenten que la possibilitat de cooperació es reforça ("Cooperation in a Repeated Prisoners' Dilemma with Ostracism", Journal of Economic Behavior and Organization, vol. 12, 1989).

explicar una categoria de fenòmens (els que acompleixen un conjunt de propietats). Malgrat tot, cal considerar que les adequacions estan mediatitzades per un factor molt important: la capacitat de recollir amb simplicitat la complexitat dels resultats de les eleccions, ja que aquests models han de permetre establir punts de contacte i facilitar els desenvolupaments teòrics d'integració amb les variables importants que intervenen en la constitució de les «situacions d'elecció». De la tensió entre la «situació d'elecció» i el model de la teoria de jocs que analitza els resultats de les eleccions, en surt un problema que ha tingut força influència en l'aplicació de la teoria de jocs a la ciència social. El punt de partida és que només amb l'anàlisi dels resultats d'eleccions ja poden sorgir en determinats casos explicacions sorprenents, malgrat integrar un mínim de variables de la «situació d'elecció» i disposar d'una molt precària estructura teòrica sobre el funcionament de l'acció col·lectiva. El problema és que aquests èxits poden alterar el sentit dels models sobre els resultats d'una elecció, en augmentar ràpidament les expectatives sobre aquests models per explicar en exclusivitat els processos d'acció col·lectiva. Aquesta situació ha dificultat la possibilitat de desenvolupar models teòrics més complets, tot i que el desenvolupament teòric exclusivament en termes de teoria de jocs, malgrat els seus resultats inicials, té cada cop més dificultats per produir resultats addicionals significatius. En tots aquests models resta força indeterminació. Excepte en algunes situacions molt concretes, els models pateixen el problema d'haver de definir de forma molt detallada, i bàsicament externa, la situació d'elecció: la qüestió és discutir com es poden afavorir interpretacions que incloguin una major determinació dels models amb menors dependències externes.

#### 4.4. Problemes en la dissolució dels supòsits de racionalitat i interès propi per l'explicació de l'acció col·lectiva.

##### 4.4.1. La variabilitat de les solucions als problemes de l'acció col·lectiva.

Aquesta secció es dedica a analitzar les implicacions que té per l'estructura dels jocs que analitzen l'acció col·lectiva el canvi dels seus pressupòsits convencionals - interès propi i racionalitat-. Aquests pressupòsits no són entesos com principis inqüestionables, sinó com supòsits que s'han de justificar des de la «situació d'elecció». La definició d'un bé col·lectiu en termes de teoria de jocs no és una definició basada en les seves característiques materials o contextuals (les formes de producció i consum); sinó que és una definició bàsicament «relacional». La definició es formula a partir de la relació que s'estableix entre els individus i el bé públic. És en aquest context que pren sentit una afirmació com la de D. Snidal:

*"given any particular problem of collective action, a solution might be achieved in any number of different ways and be reflected in correspondingly different institutional development"*<sup>68</sup>

Hom pot trobar la clau d'aquesta variabilitat de solucions a l'acció col·lectiva en els indicadors derivats de la «situació d'elecció» (o estructures institucionals, en el sentit de Snidal) que s'incorporen a la modelització dels jocs a través dels pressupòsits sobre les motivacions (orientacions d'interès) i la forma de presa de decisions dels actors en el joc. Els individus configuren un joc perquè a partir d'un conjunt de resultats possibles

---

<sup>68</sup> D. Snidal, "Coordination versus Prisoners' Dilemma: Implications for International Cooperation and Regimes", *American Political Science Review*, vol. 79, 1985, pàg. 993.

(relacionat amb les accions factibles) hi apliquen unes determinades funcions per a valorar la satisfacció que els poden produir aquests resultats. Els resultats possibles són analitzats pels subjectes com dependents de les seves estratègies, les quals són seleccionades a través de mecanismes de decisió, generalment connectats amb les funcions de satisfacció. En aquest sentit, els jugadors poden valorar idènticament o no els resultats. Si tots els jugadors ho valoren igual apareix la simetria en el joc. Malgrat tot, és la natura del bé col·lectiu la que determina quins són aquests resultats possibles, com un aspecte extern a l'heurística individual. Per exemple, considerant la intensitat de la diferència de satisfacció entre els resultats de l'estratègia de cooperació o la de no cooperació: en el cas de ser necessari un voluntari per impedir una catàstrofe col·lectiva, el mecanisme de decisió que s'adoptaria serà probablement molt diferent si aquest acte voluntari es refereix a qüestions de vida o mort, o bé a qüestions de diferències retributives.

En les relacions entre accions i resultats, hom pot distingir dos nivells d'anàlisi sobre la intervenció dels individus. Així, en relació al primer nivell, si canvien els pressupòsits sobre l'orientació d'interès dels individus (vid. seccions 1.2. i 2.2.), en modificar-se la relació entre aquests i la funció de producció del bé col·lectiu, es modifica també l'estructura del joc. Els resultats possibles de la interacció segueixen essent els mateixos, perquè les característiques del bé col·lectiu que s'intenta aconseguir no sofreixen modificacions, però no passa el mateix amb la valoració dels resultats i amb el procés d'establiment de preferències sobre aquests, ja que es guien per orientacions de comportament diferents a les de l'interès propi. El segon nivell d'anàlisi es concentra en el procés d'escollir unes determinades estratègies d'acció, normalment de forma consistent amb les preferències establertes. Però aquest procés també pot diferir de les normes de racionalitat aplicades tradicionalment, i cal subratllar que aquest nivell d'anàlisi és analíticament independent del primer. Aplicant algun dels diversos mecanismes de decisió, alguns ja destacats en pàgines anteriors (vid. seccions 1.3. i

2.2.), poden ésser escollides estratègies diferents i, per tant arribar a distintes situacions d'equilibri. Finalment, a partir de la valoració dels resultats possibles, i de la interacció entre les estratègies adoptades per cada individu, el joc produeix un resultat determinat. L'estabilitat d'aquest resultat depèn de diversos factors, però especialment de si els mecanismes de decisió de cada individu són coincidents en acceptar el resultat produït.

En el sentit exposat, no és necessari suposar que tothom té les mateixes orientacions de comportament. Diferents individus en interacció per l'obtenció d'un sol bé col·lectiu poden estar considerant al mateix temps l'existència de diferents jocs per obtenir aquest bé, sense que cap individu estigui pròpiament equivocat<sup>69</sup>. De la mateixa forma, diferents individus en un mateix procés de producció d'un bé col·lectiu, poden estar aplicant mecanismes de decisió diferents, la qual cosa produeix que encara que tothom consideri que està jugant el mateix joc, apareixen equilibris diferents als convencionals. Per aquest motiu, l'anàlisi dels equilibris d'un joc deriva molt estretament -donades unes ordenacions de preferències (o funcions d'utilitat individuals)- d'aquest mecanisme de decisió. Si els mecanismes són iguals per a tots els actors, l'anàlisi del procés d'interacció es simplifica. Però si aquest no és el cas, el que és important d'establir en l'anàlisi de la interacció és si combinacions de diverses pautes de conducta (adoptades per subgrups d'individus) poden conduir a una situació global que pot millorar individualment sense empitjorar la dels altres.

---

<sup>69</sup> Un altre tema és el de les equivocacions o confusions reals. Formalment han de rebre el mateix tractament, i és tan sols un problema empíric determinar si l'existència de diferències en la interpretació d'una interacció entre diversos actors és fruit de diferències en les seves orientacions, o bé són problemes en l'anàlisi de la informació al seu abast. Si són diferències genuïnes, es pot considerar que probablement en un procés d'interacció continuat tendiran a reduir-se, però l'explicació dels processos de coincidència d'expectatives caldria buscar-la en els canvis de les funcions d'utilitat. D'altra banda, existeix la possibilitat intencional dels subjectes per manipular la informació a l'abast dels altres i introduir confusions sobre la natura del joc, per obtenir avantatges pròpies (en funció de la pròpia orientació de comportament). Sobre aquesta qüestió, hi ha alguns intents de modelització, com una aplicació limitada a la maximització del propi interès en un dilema del presoner iterat (vid. apt. 2.2.5.).

La introducció d'heterogeneïtat en qualsevol tipus de models implica sempre l'aparició de més complexitat i, en aquest sentit, l'argument anterior no s'escapa a aquest problema<sup>70</sup>. Evidentment, hom pot proposar models simples (i simètrics) amb pressupòsits d'orientació i de decisió diferents dels convencionals, però també únics (sense establir combinacions dins un mateix joc). En aquesta situació, hipotèticament els models no tindrien perquè tenir una menor capacitat de predicció que els derivats dels supòsits convencionals, deixant apart el seu major o menor ajust als comportaments observats. De tota manera, si l'adopció d'uns nous supòsits es realitza de forma exclusiva, els problemes sobre els pressupòsits convencionals que s'han criticat abans tornarien a aparèixer amb una força renovada. Per aquest motiu, el que sembla realment interessant justificar des de la «situació d'elecció» són els pressupòsits de comportament adoptats, siguin aquests simples, o bé utilitzant la possibilitat d'introduir combinacions entre diversos pressupòsits, cercant les relacions internes i els seus efectes conjunts sobre l'acció col.lectiva.

Les distincions abans proposades entre els dos nivells analítics d'una elecció poden facilitar la investigació sobre el funcionament de les eleccions i la seva influència en els resultats dels processos d'acció col.lectiva. Habitualment les propostes realitzades sobre els resultats de les eleccions consideraven inseparable l'orientació de l'interès i el mecanisme de decisió, considerant, en tot cas, que havien de variar conjuntament.

---

<sup>70</sup> Per un tractament general del problema, que és força general en el camp de la modelització dins les ciències socials, vid. G. Arminger, "Fallacies and Paradoxes Caused by Heterogeneity", a A. Diekmann i P. Mitter, (Eds.), Paradoxical Effects..., op. cit., 1986. Sense sortir dels models sobre l'acció col.lectiva, hom podria introduir molts altres tipus d'asimetries en els jocs, que complicarien excessivament la discussió teòrica que es realitza en aquestes pàgines, encara que per una aplicació empírica caldria discutir detalladament sobre la conveniència d'incorporar-ne algunes com variables del model que es proposi. Com a mostra, Hardin (a Collective Action, op. cit., 1982, cap. 5) assenyala quatre possibles formes d'asimetria, que a més es podrien combinar entre elles: diferències entre costos i beneficis, desigual percepció dels beneficis, heterogeneïtat dels beneficis i distorsions en les valoracions de guanys i pèrdues. Aquestes asimetries afectarien directament les funcions d'utilitat dels individus (la formació de les seves preferències), modificant, per tant, l'estructura dels jocs (tot i que potser no afectarien tothom d'igual forma), i podrien alterar els efectes sobre el procés d'acció col.lectiva que es pronosticarien amb els models convencionals.



En aquest sentit, un problema freqüent és el de les pretensions de generalitat de força proposades, les quals, en no distingir entre diferents processos de decisió, només enfoquen algunes formes de producció d'acció col·lectiva. Al llarg d'aquest apartat es desenvolupen les distincions esmentades, per situar algunes de les diverses formulacions parcials existents com a parts específiques d'un problema analític comú, que és presentat en la seva forma més general.

Fins a cert punt, els jocs analitzats a les seccions anteriors amb els supòsits convencionals sobre interès i forma de decisió, es poden entendre com a un conjunt de models de referència per a la producció «típica» de béns col·lectius en situacions d'acció col·lectiva. Però cal afegir-hi el convenciment que moltes «situacions d'elecció» comporten eleccions no representables amb el supòsit de racionalitat i que, per tant, en força casos és necessari analitzar els resultats de les eleccions introduint altres consideracions. La modelització de qualsevol situació concreta d'acció col·lectiva difícilment es pot realitzar amb pressupòsits únics, ja que els individus hi poden intervenir amb una àmplia varietat d'interessos, alterables, a més, en el temps. Per aquest motiu, els intents de teorització de les situacions d'acció col·lectiva s'han d'escometre a través de dues línies parcials. La primera, mitjançant la interpretació teòrica de les «situacions d'elecció». En aquest cas, l'especificació dels contextos a través de l'anàlisi institucional i de les limitacions que estableixen les relacions tècniques i socials pot permetre una aproximació a aquest objectiu. La segona, més propera a l'objecte de la tesi, establint generalitzacions parcials sobre resultats d'eleccions amb varietat de supòsits de comportament i de decisió<sup>71</sup>.

---

<sup>71</sup> La proposta de generalitzacions parcials, en un marc interpretatiu més reduït, és realitzada per J. Elster, a *The Cement of...op. cit.*, 1989.

#### 4.4.2. Orientacions d'interès i provisió interactiva de béns públics.

La divisió analítica entre dos «moments» del procés d'elecció -formació de preferències i mecanismes de decisió- permet interpretar millor diverses propostes de generalitzar el problema de l'acció col·lectiva ampliant la diversitat de supòsits que l'enfocament de l'«elecció racional» presenta normalment. Gran part de la confusió existent dins aquest enfocament ha estat deguda a què freqüentment s'ha transmès la identificació tradicional entre racionalitat i interès (vid. apt. 1.1.3.), en els intents d'ampliar els supòsits de comportament individual més enllà del pressupòsit d'interès propi. També s'han proposat sovint canvis de supòsits sobre el comportament, però de forma aïllada, sense considerar les implicacions d'aquests canvis per a la resta de variables. Entre les diverses propostes ja analitzades (vid. seccions 1.2. i 2.2.), se'n presenten dues per ésser reinterpretades en aquest apartat com a formes distintes d'explicar el primer «moment» del procés d'elecció -quan s'estableixen les preferències en funció de les motivacions i valoracions individuals-.

La primera proposta és la realitzada per Schulz i May (vid. apt. 2.2.2.), els quals consideren l'existència de diverses funcions d'utilitat, on la funció de maximitzar els guanys propis solament és un cas particular. Entre d'altres funcions, aquests autors presenten funcions igualitàries, altruistes i competitives per representar possibles orientacions d'interès dels individus. Posteriorment, s'ha assenyalat que diversos entorns d'interacció ('networks') poden generar formes particulars de comportament, com a orientacions generals comunes a tots els individus vinculats a un entorn, que ja són conegudes familiarment pels individus que es relacionen regularment<sup>72</sup>. Sense modificar la hipòtesi de racionalitat, la transformació de les funcions d'utilitat ja implica la formació de jocs distints per un mateix procés de producció d'un bé públic. Les

---

<sup>72</sup> Cf. F. Scharpf, "Games Real...", art. cit., 1990.

implicacions per l'anàlisi de l'acció col·lectiva d'aquestes transformacions gairebé no han estat analitzades<sup>73</sup>. Com a continuació de l'apartat anterior, es suggereixen algunes possibles línies de desenvolupament. Formalment, aquesta proposta podria aplicar-se a les tres línies distintes de simplificació de la teoria de jocs analitzades abans, cosa que conduiria a modelitzacions de l'acció col·lectiva significativament diferents. Sobre la seva aplicació als models de dos estratègies per 'N' jugadors, i suposant que existeix una mateixa orientació per tothom, l'objectiu seria poder arribar a formular generalitzacions del tipus «un comportament igualitari en la producció d'un bé col·lectiu caracteritzat anteriorment per un dilema del presoner, implica l'aparició d'un joc de coordinació». La possibilitat de combinar diverses orientacions de comportament en un mateix joc de 'N' jugadors complica força l'anàlisi, ja que formalment només es pot representar a través d'una combinació de diversos jocs elementals, perquè cada orientació de comportament implica suposar una determinada estructura d'interacció per a la producció d'un mateix bé col·lectiu.

L'altra proposta que s'analiza en l'àmbit de les orientacions d'interès (el primer «moment» d'una elecció) és elaborada per J. Elster. Aquest autor suposa l'existència de diverses orientacions de comportament en una societat, considerant que és bàsicament a través de la seva interrelació que es produeixen les situacions d'acció col·lectiva que generen una alta taxa de participació voluntària. Elster proposa l'existència de cinc «tipus ideals», que representen diferents formes de comportament: «Egoistes», «Kantians», «Utilitaristes», «Seguidors», «Elitistes-Massius» (vid. apt. 1.2.1.) -sense assenyalar si són representables o no a través de funcions d'utilitat-. Aquest autor no

---

<sup>73</sup> M. Taylor, a *The Possibility of...*, op. cit., 1987, pàgs. 39-40, proposa jocs «híbrids». Per exemple, el jugador de fila té preferències d'un dilema del presoner, i mentre que el de la columna presenta les preferències d'un joc del «gallina». Si es pressuposa que hi ha una interpretació subjectiva de la provisió del bé públic, un jugador pot valorar de forma diferent que l'altre la provisió del bé públic, i considerar que la interacció és un altre joc. En una situació com aquesta, l'aparició d'un punt d'equilibri pot ésser molt menys problemàtica, en augmentar les probabilitats d'existir estratègies dominants i un equilibri òptim en el sentit de Pareto.

aclareix fins a quin punt tots els tipus de comportament segueixen també les pautes de la racionalitat, encara que és possible deduir en la seva proposta l'intent de presentar com a «tipus ideals» aquestes orientacions de comportament, que es corresponen a determinades funcions d'utilitat (que reflecteixen orientacions d'interès) combinades amb algunes propietats del càlcul racional per «gestionar» les preferències (vid. apt. 1.2.2.). Per exemple, es podria interpretar que el tipus «seguidor» seria un individu que adoptaria funcions d'utilitat igualitàries i amb un mecanisme de decisió basat en una relació de consistència i el seguiment de convencions. Aquesta proposta pot ser entesa com l'establiment de «tipus ideals», aplicable a la interpretació d'observacions freqüents d'un comportament en situacions d'acció col·lectiva. La definició és suggerent, perquè permet reduir a uns pocs tipus l'infinít nombre de combinacions possibles (d'orientacions d'interès i de formes de decisió). Per tant, mitjançant la formulació de proposicions sobre les interaccions entre els diferents «tipus ideals» definits, és possible d'establir relacions causals sobre el desenvolupament de processos d'acció col·lectiva. En aquest sentit, Elster desenvolupa una primera aproximació a les possibilitats de la seva proposta, destacant dues generalitzacions parcials:

*"Every day Kantianism and the Norm of Fairness interact to produce much more cooperation than either could do by itself (...) the strength of utilitarianism and that of fairness vary inversely with each other"<sup>74</sup>*

No obstant aquests objectius, Elster presenta aquestes generalitzacions parcials de forma excessivament simplificada, mitjançant una aplicació dels diversos «tipus ideals» a una funció de retribucions d'un bé col·lectiu que depèn del nombre de participants. El problema és que la formulació de la funció implica que tothom té la mateixa funció d'utilitat, diferint els «tipus ideals» -segons Elster- en els límits de participació dels

---

<sup>74</sup> Elster, J., The Cement of..., op. cit., 1989, pàgs. 205-206.

altres sobre els que els individus decideixen seqüencialment la pròpia participació<sup>75</sup>. La formalització no és completament incorrecta, però no serveix per a desenvolupar l'argument que es proposa. Per a tal fi caldria -almenys- definir acuradament les funcions d'utilitat de cada «tipus ideal» i, seguidament, conèixer a partir de les propietats que presenta el bé col·lectiu el tipus de joc a què s'enfronta cada «tipus ideal». Com a pas següent, es podrien buscar els tipus d'equilibri existents, tant interns dins cada grup de subjectes d'un «tipus ideal» determinat, com entre els conjunts d'individus amb comportaments caracteritzats pels diversos «tipus ideals» existents a l'escenari que s'analitza. Finalment, l'anàlisi de les diferents possibilitats d'equilibri mostraria els possibles nivells de contribució del bé col·lectiu (estables o inestables) tant en un joc simple com en un joc iterat, en el qual el factor temps també afecta les possibilitats d'assolir equilibris<sup>76</sup>.

Un intent de resoldre els problemes que presenta l'aplicació de «tipus ideals» per a l'anàlisi de l'acció col·lectiva proposada per Jon Elster, els quals combinen diferents orientacions d'interès i mecanismes de decisió, es desenvolupa per T. Kondo a través de jocs d'aprenentatge amb dilemes del presoner iterats entre dos jugadors<sup>77</sup>. Kondo assumeix només l'existència de «tipus ideals» dins els individus, amb distribucions de probabilitat sobre la seva possibilitat de predomini, i no entre els individus, tal com

---

<sup>75</sup> Aquest model no és pròpiament interactiu, sinó que és molt aproximat als models de llindar de M. Granovetter (vid. "Modelos de umbral.", art. cit., 1991), per la qual cosa manca d'introduir els càlculs estratègics de cada «tipus ideal». (vid. apt. 3.3.2.).

<sup>76</sup> Sobre els casos asimètrics en els jocs de dues estratègies amb 'N' jugadors hi ha poques propostes realitzades. En la darrera part del seu article, J. Hovi ("Binary Games as...", art. cit., 1986) presenta un cas de diferències de contribució entre els jugadors. Aquesta variable no ha estat considerada en aquestes pàgines, però l'interès de seu model és que elabora dos jocs per a la provisió d'un mateix bé públic. Els actors que realitzen una contribució significativament superior afronten un joc del «gallina», mentre que els que contribueixen menys entenen el joc com un dilema del presoner. D'aquesta forma, una estructura similar també es podria desenvolupar considerant orientacions d'interès distintes.

<sup>77</sup> T. Kondo, "Some Notes on Rational Behavior, Normative Behavior, Moral Behavior, and Cooperation", *Journal of Conflict Resolution*, vol. 34, 1990.

proposa Elster. Així, aquest autor aconsegueix reduir la complexitat que suposaria una aplicació correcta de la proposta d'Elster. A partir d'aquests elements, Kondo aplica en el seu model tres «tipus ideals»: el 'racional-egoista' (sempre trair), el 'normatiu que adopta l'estratègia «ull per ull» (cooperar si l'altre coopera), i el 'moral kantian' (sempre cooperar). Seguidament elabora una iteració del model amb expectatives adaptatives (que es modifiquen seqüencialment segons les accions dels altres) en la qual les estratègies s'adopten per optimitzar la utilitat esperada subjectiva (amb l'orientació d'interès que determini el «tipus ideal»)<sup>78</sup>. Cal notar que Kondo aplica probabilitats en la simulació per ordinador del seu model. Amb aquest procediment pretén arribar a inferir proposicions generalitzables sobre les relacions entre els «tipus ideals»<sup>79</sup>. Aquest no és el procediment clàssic de formular proposicions teòriques, però també s'encamina a aquest objectiu. En aquest sentit, la formulació d'una interpretació teòrica sobre la «situació d'elecció» seria el procediment habitual per establir hipòtesis sobre les relacions entre «tipus ideals», que podria prescindir que l'aplicació de probabilitats (i també podria evitar el problema de l'excessiva complexitat del model proposat per

---

<sup>78</sup> Els seus resultats mostren que el comportament normatiu serveix com a estabilitzador de la cooperació, la qual també apareix amb un comportament racional exclusiu, encara que en aquestes condicions adopta una forma inestable, malgrat sigui racional cooperar. D'altra banda, Kondo argumenta que el comportament normatiu aconsegueix actuar com a catalitzador per passar d'una situació de no-cooperació a una de cooperació, tot i que un excessiu nombre de cooperadors incondicionals podria convertir en racional l'estratègia de no cooperar. L'autor acaba proposant una utilització estratègica del concepte de racionalitat per agents, en el sentit d'intentar controlar les expectatives dels altres sobre el propi comportament, si aquests volen assegurar la cooperació recíproca.

<sup>79</sup> Un model que podria ésser complementari al de Kondo, que mostra la complexitat de les variables existents en aquestes formulacions, és el realitzat per M. Macy (a "Walking Out of...", art. cit., 1989). L'autor també desenvolupa un model d'aprenentatge bayesià, però en aquest cas la seva variable clau és l'estructura del joc: en comptes de definir un sol dilema del presoner, estableix diverses intensitats de diferències entre els resultats cooperatius i no cooperatius. Així, a partir d'unes disposicions inicials dels jugadors sobre cooperar o no cooperar, segons l'estructura del joc es poden produir diverses evolucions: "*Non-cooperative players in an iterated prisoner's dilemma will learn to cooperate through state-conditioned propensities if the symmetric payoffs are sufficiently salient*" (pàg. 211). L'autor suposa que tots els individus són racionals, però que la seva capacitat és bastant limitada. Aquests individus han de saber calcular les probabilitats en funció de la jugada immediatament anterior, dels paràmetres del joc i de la probabilitat assignada a la jugada anterior. No obstant això, l'aparició de la cooperació tan sols pot ser explicada com un procés basat en l'atzar, que no és produït ni intencional ni estratègicament.

Elster). No obstant, els problemes d'aquest procediment habitual són els clàssics de l'inferència a partir de les observacions empíriques.

#### 4.4.3. Mecanismes de decisió i elecció d'estratègies.

En aquest apartat es realitza una revisió de les propostes sobre els mecanismes de decisió analitzades en els capítols anteriors (deixant apart el model de decisió racional convencional, ja extensament analitzat), relacionant-los amb les situacions d'acció col·lectiva. Fora de les varietats de la hipòtesi de racionalitat, es dedica també una especial atenció a l'anàlisi de les convencions, com una fórmula per escollir estratègies força adequada per aplicar a molts processos d'acció col·lectiva. La idea que les convencions podien facilitar l'establiment de punts d'equilibri va ésser introduïda en el context de la teoria de jocs per T. Schelling, en el seu llibre The Strategy of Conflict (1960). L'interès de Schelling es centrava en identificar elements perceptibles que determinessin la formació d'expectatives mútuament coincidents entre els jugadors, per evitar l'establiment d'una regressió 'ad infinitum' de les seves estratègies (els punts focals, vid. apt. 2.1.2.). Aquests elements havien de conformar la creença de cada jugador en què tothom identificava el mateix resultat com el virtualment inevitable. Segons els especialistes en teoria de jocs, aquest element només podia formar-se a través de les «expectatives racionals» de cada jugador, suposant la simetria del joc i que tothom actuava amb les mateixes regles de racionalitat i amb els mateixos interessos<sup>80</sup>. Però a més de la complexitat de càlcul que en ocasions cal imputar als individus perquè arribin a coincidir en un resultat -a partir de l'estructura matemàtica del joc-, cal

---

<sup>80</sup> Aquesta perspectiva, llargament dominant, ha portat a una restricció de les possibilitats d'aplicació de la teoria de jocs a la ciència social. D'altra banda, el que és important per determinar un equilibri és constatar que existeixi una creença generalitzada entre tots els actors sobre la inevitabilitat d'un resultat específic. Es pot arribar a aquesta creença mitjançant un mecanisme de càlcul racional, però també a través d'altres mecanismes de decisió, sense afectar al resultat.

comptar amb la possibilitat latent de què de vegades no sigui possible arribar amb aquest procediment a un punt d'equilibri. Schelling, sense desestimar completament aquest mètode, va proposar una alternativa més simple. Aquest mecanisme de decisió alternatiu el proposava Schelling per a tot tipus de jocs mixtes de conflicte i coordinació, encara que la seva aplicació més exemplar es referia primordialment als jocs purs de coordinació, en els que freqüentment no es podia determinar un equilibri únic en termes de «expectatives racionals»<sup>81</sup>, encara que tots els equilibris fossin favorables, en conjunt, pels jugadors. En aquestes situacions, si els beneficis de la cooperació, encara que amb equilibris que produeixen «pagaments» desiguals pels jugadors, són sempre superiors a la no cooperació, la possibilitat d'aparició de la convenció és més gran, sense que per altra banda, aquesta representi necessàriament cap òptim paretà.

El supòsit bàsic de les convencions com a mecanisme de decisió en els jocs de coordinació és considerar que cada individu té necessitat d'establir contacte amb els altres mitjançant algun procés imaginatiu d'introspecció i investigació de les claus comuns a tots. En aquest context, és molt important introduir el marc interpretatiu (o «situació d'elecció») sobre el qual es defineix el joc, generalment extern a la seva estructura formal, ja que sovint són els detalls generats per la «superfície» d'aplicació del joc els que proveeixen un indicador que és reconegut per totes les parts<sup>82</sup>. Però

---

<sup>81</sup> Els jocs de coordinació són jocs que contemplen diverses situacions de cooperació, que poden ser beneficioses de forma desigual pels participants. Si només hi ha un punt d'equilibri mútuament beneficiós, establir la coordinació no és cap problema; però quan apareix més d'un punt d'equilibri mútuament beneficiós i no n'hi ha cap que sigui estrictament preferit, les convencions poden ajudar a resoldre aquest problema. La definició de Lewis sobre les convencions (a *Convention: A Philosophical Study*, Cambridge University Press, Cambridge, 1969) parteix d'aquest problema: una convenció és l'elecció d'una solució d'equilibri quan hi ha una multiplicitat de solucions. A més, per Lewis, totes les solucions d'equilibri impliquen comportaments que compleixen el principi de maximització.

<sup>82</sup> Per exemple, el típic sorgiment d'un incident entre una massa excitada, que acaba provocant una revolta (Cf. T. Schelling, *La estrategia del...*, op. cit., 1964, págs. 117-120). L'autor insisteix més argumentant que "el proceso psíquico e intelectual esencial es el de participar en la creación de tradiciones; y los ingredientes



acceptar que els actors no es guien sempre pels supòsits convencionals de la racionalitat, amb els que s'ha bastit una teoria normativa-matemàtica sobre les estratègies dels jocs, tampoc vol dir que aquests sempre es comportin de forma contrària. Segons les circumstàncies, pot ser que l'adopció d'una decisió racional sigui l'única forma de resoldre el joc, o bé pot ser que es produeixi una coincidència respecte a l'equilibri entre les estratègies escollides mitjançant diversos mecanismes de decisió (potser degut a la visibilitat o la unicitat de la pròpia solució). La transformació de les convencions en normes socials és una qüestió que queda fora de l'àmbit d'aquesta discussió. De tota manera, si bé l'aparició d'una convenció tan sols deriva d'una coincidència, motivada per algun factor extern, en l'elecció d'estratègies per arribar a un equilibri cooperatiu l'establiment d'una norma social requereix addicionalment altres elements, apart de l'extensió i repetició de la pròpia convenció. A més, tampoc és necessari pensar que totes les regles socials tinguin el seu origen en convencions<sup>83</sup>.

Pel que fa a la interpretació de les situacions d'acció col·lectiva, l'aplicació d'una anàlisi en termes de convencions pot facilitar un ràpid aclariment de la seva

---

*con los que pueden crearse unas tradiciones, o los materiales con los que pueden percibirse unas tradiciones potenciales, no coinciden en absoluto con el contenido matemático del juego" (pàg. 129). Apart de la coordinació cognitiva, basada en compartir intuïcions, també poden apareixer situacions de coordinació en processos d'aprenentatge mitjançant prova i error (associats a funcions d'utilitat), que portin a estratègies cooperatives condicionals en jocs iterats. Però aquests processos ja impliquen mecanismes més complexos de decisió (vid. les referències als treballs de Kondo i de Macy a l'apartat anterior).*

<sup>83</sup> Vid. apt. 2.2.3.. D'altra banda, la creació i manteniment d'institucions socials que proveeixin indicadors clars als individus per coordinar les seves expectatives poden contribuir a la generalització de les convencions en normes, sobretot si els resultats coincideixen amb l'optimització de les funcions d'utilitat dels individus. Per una discussió d'aquesta qüestió, vid. C. Ford Runge, "The Assurance Problem...", art. cit., 1984 i H. Leibenstein, "On the Economics of Conventions and Institutions: An Exploratory Essay", *Zeitschrift für die gesamte Staatswissenschaft*, vol. 140, 1984. Evidentment, el procés seria més fàcil en jocs de coordinació que en jocs mixtes de conflicte-cooperació, però no és necessari que canviïn les funcions d'utilitat -contràriament a l'opinió de Runge- perquè es produeixi un equilibri estable mitjançant les convencions (sobretot si els indicadors proveïts per les institucions són eficaços), encara que, com assenyalava Leibenstein, es pot tractar d'equilibris no òptims. Vid. també el treball d'E. Ullmann-Margalit (*The Emergence of... op. cit.*, 1977) per a un tractament de les convencions com a normes socials.

dinàmica interna, sense necessitat de suposar el càlcul racional per part de cada jugador sobre totes les estratègies possibles dels altres jugadors. En jocs amb 'N' jugadors, sovint és necessari aplicar els supòsits inclosos a la noció de «sentit comú» per eliminar els costos de la informació -i això només és possible quan la interacció i els seus actors ja són familiars entre si-. L'elevat esforç necessari per a obtenir aquesta informació quan la «situació d'elecció» és 'nova', pot portar a la major part dels actors a aplicar estratègies basades en convencions (potser amb l'excepció dels jugadors que estan millor situats per tenir coneixement de les estratègies dels altres) per tal de fer predible el comportament dels altres<sup>84</sup>.

Una alternativa proposada per R. Hardin per a reduir els problemes de la manca d'informació presents en els grups grans consisteix en considerar els processos d'acció col·lectiva com a un "by product" of smaller subgroup interactions<sup>85</sup>. Seguint aquesta proposta, Hardin afirma que en els grups més petits es redueixen els problemes d'obtenció d'informació i augmenten les possibilitats d'assolir equilibris no basats en convencions o normes:

*"The greater distinction between small and large groups is not one of Olson's logic, but merely of the likelihood of their being involved in a thick enough network of mutual interactions"*<sup>86</sup>

---

<sup>84</sup> Precisament en aquest sentit, s'afirma que en augmentar la grandària dels grups es redueix la tendència a la innovació en les convencions, malgrat que existeixin majors beneficis socials en el cas d'adoptar-se una convenció alternativa. Sobre els processos de canvi i manteniment de convencions vid. M. Tietzel, "Virtue, Vice and Dr. Pangloss", *ORDO*, vol. 41, 1990. D'altra banda, quan existeix heterogeneïtat o asimetria entre els membres del grup, és també més difícil el funcionament continuat d'una convenció general (cf. R. Hardin, *Collective Action*, op. cit., 1982, pàgs. 92-96).

<sup>85</sup> Cf. Hardin, *Collective Action*, op. cit., 1982, pàg. 184. Sobre el paper de la informació en els equilibris en jocs de 'N' jugadors, vegeu l'apartat 3.3.2.

<sup>86</sup> R. Hardin, *Collective Action*, op. cit., 1982, pàg. 228.

L'argument d'Hardin és interessant, però cal remarcar que la seva anàlisi reflecteix una tendència i no un principi que es pugui proposar axiomàticament. Tot i que els costos d'informació augmenten molt amb la dimensió del grup, també poden existir situacions de «sentit comú» que contribueixin a solucionar el problema. Cal afegir a més que no es pot fer dependre l'existència d'una acció col·lectiva -basada en càlculs racionals- de l'estructuració precisa de les relacions entre els agents, articulant necessàriament les seves interaccions en grups reduïts. Per exemple, com s'ha esmentat anteriorment, ni en els models més simples d'interaccions de subgrups, les competicions iterades de jocs de dos per dos, tampoc són completament satisfactòries les condicions per assegurar la cooperació en totes les circumstàncies (vid. apt. 4.3.2.).

En conjunt, el propi argument d'Hardin serveix com a exemple de les dificultats de realitzar algunes proposicions teòriques, les quals, per la seva natura, normalment només poden arribar a ser generalitzacions empíriques. Pel cas proposat per Hardin, el que es podria deduir és que quan les interaccions s'observen bàsicament a través de subgrups, és més probable d'observar en els participants l'aplicació de mecanismes racionals, o aproximadament racionals, de calcular les estratègies. D'altra banda, l'aplicació de les convencions a la interpretació de l'acció col·lectiva presenta els seus límits quan els processos d'implicació per a la provisió voluntària de béns públics són considerats més atentament pels subjectes participants. Siguin aquests processos situacions representables amb jocs de pura coordinació, o bé amb jocs de conflicte-cooperació, hom hi podria considerar hipotèticament mecanismes d'elecció d'estratègies basats en convencions -potser més fàcilment en els primers, però tampoc desestimables en els segons-; però no sempre es pot mantenir la hipòtesi que les eleccions d'estratègies són simplement producte de convencions establertes. A més, el mateix argument també pot ser aplicat per l'anàlisi de les normes socials. Els motius d'aquestes limitacions poden ser diversos, des de la mateixa complexitat de les decisions, fins a la pròpia centralitat del procés de participació per la identitat individual, passant per una possible

inexistència de senyals per a l'establiment de convencions. En aquest sentit, el tipus d'acció col·lectiva que es produeixi pot ser determinant perquè els mateixos subjectes escullen fer ús d'un o altre mecanisme de decisió. També, dins una situació d'acció col·lectiva poden aparèixer classes d'individus que apliquen distints mecanismes -per exemple, alguns fent ús de convencions, altres utilitzant càlculs racionals-, ja que no tothom ha d'atorgar un mateix valor a la seva implicació en una acció col·lectiva.

Apart de les convencions, també poden existir altres mecanismes de decisió no basats en la hipòtesi de racionalitat, i la distribució d'aquests mecanismes no ha de ser homogènia en un grup. Una possibilitat freqüent és que la majoria estableixi estratègies basant-se en càlculs racionals, o en convencions; però que una minoria d'altres individus utilitzi altres mecanismes que siguin fonamentals per a la formació de l'acció col·lectiva, en el sentit de facilitar les condicions als altres individus per contribuir o per trobar un «punt focal». Entre les diverses possibilitats existents, hom pot mencionar tant mecanismes aleatoris deguts a la ignorància i la confusió sobre la lògica de l'acció col·lectiva, com normes morals, que apliquen sempre una determinada estratègia<sup>87</sup>. Aquestes fórmules per si soles tenen poc interès, i és poc plausible la seva difusió generalitzada i uniforme. La seva importància consisteix en la possibilitat de combinar-se amb mecanismes racionals o basats en convencions -dins un mateix individu o entre individus distints- per reforçar les contribucions a l'acció col·lectiva<sup>88</sup>.

---

<sup>87</sup> T. Kondo (a "Some Notes on Rational..", art. cit., 1990) introdueix aquesta possibilitat amb individus que tenen funcions d'utilitat que maximitzen l'interès propi i consideren que amb l'aplicació de la norma ho aconsegueixen. En aquest sentit, cal introduir una distinció important i poc reconeguda, entre individus amb funcions d'utilitat altruistes però calculadors racionals, i individus amb funcions d'utilitat egoistes però aplicadors incondicionals -sense càlcul racional- d'una estratègia («Kantians», seguint la terminologia d'Elster).

<sup>88</sup> R. Hardin, *Collective Action*, op. cit., 1982, cap. 7. Un problema important de l'obra d'Hardin és la seva constant identificació entre interès propi i racionalitat, sense distingir analíticament que es tracta de dos processos distints. En aquest sentit, argumenta que el manteniment de les convencions ha deure's degut fonamentalment a la racionalitat i l'interès propi associat; quan aquesta explicació només és una possibilitat, entre d'altres a aplicar segons el joc concret que s'analitzi.

Per acabar amb els mecanismes de decisió, cal introduir de nou les concepcions sobre la racionalitat exposades en els dos primers capítols de la tesi, especialment les que intenten presentar formulacions més plausibles sobre els processos observables de decisió racional dels individus. Les dues gran línies d'aproximació són la proposta de racionalitat limitada formulada inicialment per H. Simon, i les diverses variacions de la Hipòtesi de la Utilitat Esperada, a les que en conjunt s'anomena com variacions de la hipòtesi de racionalitat. Ambdues perspectives pretenen, amb pressupòsits diferents, aconseguir explicar mecanismes de decisió generalitzables, per poder ésser aplicats a processos socials com els que s'inclouen dins les situacions d'acció col·lectiva.

Els efectes descrits en el primer capítol de la tesi sobre les distorsions observades en els processos de decisió racional també s'han detectat en situacions d'interacció modelitzables a través de teoria dels jocs. L'axioma d'independència en la teoria de la utilitat esperada, un dels punts més debatuts per la plausibilitat de l'aplicació d'aquesta teoria de la decisió, tampoc es veu confirmat en els estudis experimentals sobre dilemes socials. En termes generals, els resultats mostren que quan el dilema s'enfoca en termes de donar, els individus són favorables al risc, mentre que quan el dilema es planteja en termes de prendre, els individus són adversos al risc; per tant, la gent tendeix a ser més cooperativa quan el dilema s'enfoca en termes de prendre (que implica un cert guany immediat) que quan s'enfoca en termes de donar; tot i que l'estructura formal del joc sigui idèntica<sup>89</sup>.

---

<sup>89</sup> Per aquests resultats, vid. B. Brewer i R. Kramer, "Choice Behavior in Social Dilemmas: Effects of Social Identity, Group Size, and Decision Framing", *Journal of Personality and Social Psychology*, vol. 50, 1986; i J. Fleishman, "The Effects of Decision Framing and Others' Behavior on Cooperation in a Social Dilemma", *Journal of Conflict Resolution*, vol. 32, 1988. Un típic dilema en termes de prendre es el dels comuns (per exemple, l'explotació en comunitat d'una font de riquesa); i un típic dilema en termes de donar pot ser el dilema del presoner (per exemple, la contribució per assolir un bé col·lectiu inexistent). Un intent d'explicar aquests comportaments aplicant la «Prospect theory» (vid. apt. 1.3.3.) es pot trobar a W. McDaniel i F. Sistrunk, "Management Dilemmas and Decisions", *Journal of Conflict Resolution*, vol. 35, 1991.

Però aquests efectes no actuen sols, i si s'analitzen combinadament amb els efectes del comportament dels altres sobre les decisions pròpies, apareixen noves conseqüències. Així, una tendència ja coneguda a l'augment del nivell de cooperació si els individus observen que els altres també cooperen, pot prendre orientacions diferents segons si el joc s'orienti en termes de prendre o de donar. En concret, s'ha observat que l'actitud de prendre es relaciona força adequadament amb els estàndards del grup, però que l'actitud de donar es mostra contrària, ja que apareix el raonament segons el qual si els altres també cooperen, té menys importància la pròpia cooperació, quan aquesta es formula en termes de decisions morals o d'adopció de convencions<sup>90</sup>.

Aquests exemples experimentals no proposen models formals alternatius, però assenyalen algunes de les distorsions que poden aparèixer si hom aplica en els models d'acció col·lectiva els mecanismes de decisió basats en una concepció de la racionalitat que és incapaç de recollir-les. Per a analitzar aquestes distorsions no cal qüestionar les orientacions de comportament (o funcions d'utilitat) que puguin presentar els individus<sup>91</sup>, ja que cal indicar que els efectes esmentats es refereixen solament al segon dels dos nivells esmentats anteriorment, al dels mecanismes d'elecció d'estratègies una vegada ja valorats els resultats possibles. En aquest cas, per exemple, es tracta de l'adopció de distint nivell de risc en un procés de decisió, segons els termes en què aquesta decisió sigui presentada a l'individu. Per aquest motiu, apareix altre cop la

---

<sup>90</sup> J. Fleishman, "The Effects of Decision...", art. cit., 1988, pàgs. 174-178. L'autor, després d'analitzar el fenomen, realitza propostes en termes d'enginyeria social: "*Framing a social dilemma in terms of giving may facilitate the diffusion of responsibility*" (pàg. 175).

<sup>91</sup> Un cas en què si hi ha implícita una funció d'utilitat diferent de la convencional és quan apareixen béns «relacionals», la utilitat dels quals no depèn només del seu nivell de consum individual, sinó que també depèn de la relació amb els altres, ja que el seu consum conjunt proveeix utilitat (C. Uhlander, "Relational Goods' and participation...", art. cit., 1989; vid. també apt. 2.2.2.). Al contrari del que defensa l'autora, es la natura del bé -definint una estructura del joc-, i no una funció d'utilitat inherent a aquest, la que imposa funcions d'utilitat conjuntes als actors, ja que les individuals serien sempre nul·les. En aquest sentit, es podria donar el cas de no producció del bé perquè els individus no adoptessin funcions d'utilitat més profitoses.

influència de l'entorn institucional, que cal incorporar en el model d'elecció<sup>92</sup>. Un altre exemple es refereix a l'observació de diferents mecanismes de decisió en una situació d'acció col·lectiva entre grups d'individus, depenent de la seva història anterior en el joc. En aquest sentit, els estudis experimentals assenyalen que

*"cooperators do not leave more reality than defectors because (...) they are more optimistic than defectors about the number who will cooperate, such optimistic compensating for the fact that they can expect less there than defectors with constant expectations"*<sup>93</sup>

Així, s'observa en aquests estudis que el càlcul racional està afectat per distorsions derivades de pressions d'aparença normativa sobre el benestar del grup, però que els jugadors continuen aplicant la racionalitat en les seves decisions. Aquesta situació porta a suposar l'existència d'algun factor que limiti les possibilitats d'elecció d'estratègies, amb una influència variable segons les pressions normatives existents. Diverses propostes de mecanismes de decisió, modelitzant variacions de la hipòtesi de la racionalitat han estat discutides en el segon capítol, mostrant la seves dificultats - difícils de solucionar- per a convertir-se en principis de comportament vàlids per a una aplicació general<sup>94</sup>. Pel que fa a la seva aplicació a l'explicació dels processos de

---

<sup>92</sup> Això es pot entendre com que la gent respon passivament a l'entorn (i, per tant, s'han d'incorporar indicadors sobre l'entorn a través dels mecanismes de decisió i cognició). Però no cal oblidar, com diu T. Yamagishi, que els individus també poden actuar sobre l'entorn: "*Structural factors affect people's perceptions and expectations, but people act, on the basis of these perceptions and expectations, to change the structure itself*" (a "The Provision of a Sanctioning System as a Public Good", Journal of Personality and Social Psychology, vol. 51, 1986, pàg. 115). Aquesta dimensió d'anàlisi (els efectes de les eleccions sobre la «situació d'elecció») no considerada aquí, recorda la major amplitud dels processos d'acció col·lectiva.

<sup>93</sup> J. Orbell, P. Schwartz-Shea i R. Simmons, "Do Cooperators Exit More Readily than Defectors?", American Political Science Review, vol. 78, 1984, pàg. 155. Sobre el mateix tema vegeu també P. von Lange, W. Liebrand, i D. Kuhlman, "Causal Attribution of Choice Behavior in Three N-Person Prisoner's Dilemmas", Journal of Experimental Social Psychology, vol. 26, 1990.

<sup>94</sup> vid. apts. 2.2.3. i 2.2.4.

decisió existents en les situacions d'acció col·lectiva<sup>95</sup>, poden tenir una oportunitat d'aplicació si la seva justificació es deriva de factors de la «situació d'elecció», com la natura i característiques del bé públic. Possiblement aquesta aplicació s'ha d'articular a través de propostes teòriques que no entrin en competició, sinó que siguin complementàries, en el sentit d'utilitzar diferents principis de decisió per a interpretar distints contextos.

Donada aquesta situació, com que no és el propòsit de la tesi desenvolupar un model concret, tan sols s'indiquen alguns dels possibles problemes que podrien aparèixer en aquests intents. Primer, en referència a l'exclusió de relacions inherents a la hipòtesi de racionalitat, cal dir que no té perquè existir una fórmula superior a les altres, ja que en diverses situacions potser sigui millor eliminar la transitivitat, mentre que en altres la consistència -per exemple-, depenen de l'observació sobre la forma de decisió dels subjectes. Segon, les possibilitats d'aparició d'equilibris tendeixen a augmentar en tots aquests casos, per la qual cosa és recomanable de combinar en la modelització d'una interacció subjectes amb diferents mecanismes de decisió -que a més, coincideix amb les situacions típiques d'acció col·lectiva-. Tercer, l'augment el nombre de subjectes pot alterar, però no modificar substancialment els problemes ja esmentats per aquestes situacions, degut a què no cal alterar l'estructura del joc, les funcions d'utilitat dels individus i els supòsits sobre la informació que aquests tenen al seu abast. La incorporació de models que utilitzen la teoria de la racionalitat limitada, planteja el problema ja conegut d'haver d'incorporar els nivells de satisfacció, que en aquest cas s'haurien de definir des de la «situació d'elecció». Així, hi hauria dues possibles vies per a introduir elements de racionalitat limitada en la interpretació de

---

<sup>95</sup> Gairebé no existeixen intents de desenvolupar models per vincular els dos camps (vid. per exemple W. McDaniel i F. Sistrunk, "Management Dilemmas...", art. cit., 1991). No obstant això, si existeixen referències a la seva necessitat, com per exemple J. Andreoni, "Why Free Ride? Strategies and Learning in Public Goods Experiments", *Journal of Public Economies*, vol. 37, 1988.



l'acció col·lectiva. Per un costat, proposant mecanismes de càlcul que adoptin estratègies predeterminades a partir d'un cert nivell de complexitat, d'altre costat, introduint consideracions probabilístiques sobre les funcions d'utilitat (i no només sobre els mecanismes de decisió) que permetin eliminar la hipòtesi d'optimització<sup>96</sup>.

En general però, la fragilitat d'aquestes propostes mostra les dificultats per generar hipòtesis sobre mecanismes de decisió en situacions interactives quan manca la interpretació -no absolutament contextual- sobre els entorns físics, materials i institucionals dins dels quals es formulen les decisions. Així, l'intent de construir interpretacions teòriques a partir d'una teoria de la decisió, inclús quan aquesta es distancia dels pressupòsits de la racionalitat normativa o convencional, mostra també els seus límits. No obstant això, si es considera que la teoria de la decisió és simplement això -una teoria sobre com es formen i quins resultats tenen les decisions-. Aquesta pot tenir una gran utilitat, però només per interpretar una part -bàsicament, els resultats de les eleccions- dels processos socials i polítics. Addicionalment, es poden generar afirmacions més generals sobre els efectes de combinacions de «tipus ideals» per la producció d'acció col·lectiva; o bé poden ajudar a definir variables de les «situacions d'elecció». Però no poden substituir la conceptualització i construcció teòrica sobre aquestes situacions.

---

<sup>96</sup> Els millors exemples en aquest sentit són les propostes de Heiner i de Rosenthal, ja discutides al segon capítol (vid. especialment apt. 2.2.5.).

**CAPÍTOL CINQUÈ:**

**Conclusions.**

Als capítols precedents, s'ha presentat un marc per a l'aplicació de la teoria de jocs a les ciències socials. Aquesta aplicació, que pot resoldre algunes de les dificultats importants que planteja normalment la utilització de la teoria de jocs, requereix que l'adopció de la hipòtesi de racionalitat sigui justificada externament. Aquesta perspectiva ja implica una resposta a la pregunta inicial de la tesi, sobre el paper de la hipòtesi de la racionalitat com a fonament per a teories politològiques basades en l'acció. Tant a la part orientada a les qüestions metodològiques, com a la part en què s'ha explorat una de les línies de construcció teòrica d'aquest paradigma, s'han aportat diverses consideracions per defensar aquest status de la hipòtesi de racionalitat. No obstant això, és convenient insistir en una perspectiva de conjunt, que connectant les dues parts de la tesi, repregui les qüestions centrals plantejades a la introducció de la tesi, incloent la discussió sobre la definició d'una teoria de l'acció, en relació a les necessitats i possibilitats de la ciència política contemporània.

El propòsit inicial del treball realitzat era analitzar les possibilitats d'emprar en les explicacions dels models econòmics aplicats a la ciència política uns fonaments no basats exclusivament en el tradicional model de l'«homo oeconomicus», és a dir, en els pressupòsits convencionals sobre la racionalitat i l'interès propi. Però d'aquesta anàlisi ha sorgit una perspectiva distinta sobre l'aplicació de models d'origen econòmic a la ciència política. Així, la recerca desenvolupada s'ha centrat en establir els problemes fonamentals que apareixen en la formulació de models que consideren exclusivament la racionalitat com a hipòtesi de comportament. En primer lloc, s'ha mostrat que la definició de l'orientació d'interès dels individus no és un problema central a l'enfocament de l'«elecció racional». El següent pas de la recerca ha estat encaminat a l'anàlisi de l'estatus de la hipòtesi de racionalitat. De les dues grans possibilitats d'emprar aquesta hipòtesi, com a «tipus ideal» o com a principi teòric, l'enfocament de l'«elecció racional» es decanta clarament per la darrera, degut a la major capacitat explicativa que li

suposa. Tanmateix, la discussió que s'ha produït dins el propi enfocament als darrers anys sobre la plausibilitat d'aquesta hipòtesi, ha conduït a l'aparició de força propostes per flexibilitzar els principis que la componen, els quals generalment són presentats de forma axiomàtica. Un punt important d'aquesta discussió ha estat la revisió dels principis que calia abandonar o modificar per explicar més fidelment com efectuen decisions els individus. Però de l'anàlisi efectuada es pot inferir que no hi ha una hipòtesi de racionalitat única per a propostes descriptives, sinó que existeix un ventall de varietats, que s'adapten amb èxit desigual a diferents situacions d'elecció.

De les diferents formes instrumentals que faciliten la construcció teòrica a partir de principis de decisió a les ciències socials, una de les més atractives i prometedores és la teoria de jocs, en oferir un mitjà per l'anàlisi de les eleccions interdependents. Aquesta teoria s'ha vinculat generalment a una hipòtesi de decisió racional, però recents investigacions mostren que també pot integrar altres principis de decisió, com les normes o les convencions (en ocasions, augmentant el nombre d'equilibris possibles). Aquesta possibilitat, com s'exposa al segon capítol, pot derivar per evitar el problema de la teoria de jocs per justificar força situacions habituals de les interaccions humanes, les quals, tot i l'elaboració de complicats procediments analítics, resten sense solució si hom sols utilitza la hipòtesi de racionalitat (per exemple, el problema de la cooperació en un dilema del presoner que té un nombre finit de repeticions).

Aquestes transformacions de l'enfocament de l'«elecció racional», que estan prenent força darrerament, presenten com a principal problema la manca de determinisme, degut precisament a la diversitat de propostes -no competitives- que apareix<sup>1</sup>. Aquest problema difícilment es pot abordar discutint exclusivament sobre la hipòtesi de racionalitat. Per aquest motiu, la segona part de la tesi es deriva cap

---

<sup>1</sup> No és que les propostes siguin poc deterministes -aquest és un altre problema- sinó que l'enfocament pateix d'una manca de capacitat per establir criteris sobre quins principis aplicar a cada cas.

a l'anàlisi dels problemes existents en l'estudi dels fenòmens d'acció col·lectiva. Prenent com a punt de partida l'obra de Mancur Olson sobre la lògica de l'acció col·lectiva, que exposa els problemes de la cooperació en els grups grans, es contraposen les dues línies de solució que tradicionalment s'han emprat per analitzar aquests problemes. La solució «interna» proposa interpretar els problemes de l'acció col·lectiva aplicant exclusivament la hipòtesi de racionalitat, mentre que la solució «externa» reclama la necessitat d'introduir-hi altres supòsits de conducta (o bé que la racionalitat és afectada per altres variables). El treball d'Olson es pot considerar com la primera gran síntesi entre les dues línies de solució, en introduir el paper dels incentius selectius; però com a tal, deixa molts interrogants oberts, sobre els quals s'han generat intenses polèmiques. En tot cas, a la continuació de les discussions tampoc s'ha ofert una resposta definitiva, perquè malgrat que els defensors de la solució «interna» es mostren impossibilitats de resoldre problemes fonamentals, com la reducció de la propensió a la cooperació en augmentar la grandària dels grups, les solucions «externes» no aconsegueixen explicar els factors que reconeixen com significatius en la interpretació dels processos de decisió. Davant aquest dilema, en aquesta tesi es sosté que per facilitar la síntesi, no és convenient discutir sobre la preeminència d'una hipòtesi concreta -o cercar una alternativa global a la racionalitat-, sinó que la via més prometedora per apropar les dues línies de solució és proposar mecanismes que permetin ordenar i integrar variables importants pels resultats de l'acció col·lectiva en els processos de decisió.

Una via encaminada a resoldre aquest objectiu és estendre la teoria dels béns públics a qualsevol resultat produït per una acció col·lectiva, tot i separant conceptualment aquesta teoria de la seva tradicional vinculació amb la producció de béns materials. La proposta que integra el conjunt de fenòmens de l'acció col·lectiva es realitza amb models de la teoria de jocs que recullen les propietats dels béns públics. Per aquesta tasca, s'integren conceptualment diverses línies analítiques ja existents a la teoria de jocs, a les quals cal afegir-hi diferents supòsits simplificadors, especialment per tractar el problema de la manca d'informació completa en els jocs.

La modelització dels béns públics en termes de la teoria de jocs permet explicar una part important dels fenòmens d'acció col·lectiva, en concret, els resultats de les eleccions individuals -entesos com equilibris cooperatius en la interacció entre decisions individuals-, quan ja es coneixen les preferències i els mecanismes que guien aquestes eleccions. No obstant això, en desestimar la solució «interna», resta el problema d'explicar com es formen preferències i mecanismes d'elecció. És necessari referir-se a la «situació d'elecció», o sigui, el context institucional que envolta un procés de decisió. Des d'aquesta «situació d'elecció» s'ha de determinar quines són les orientacions d'interessos que segueixen les preferències, així com quins són els mecanismes d'elecció que utilitzen els individus. La forma com es produeixen aquestes «determinacions» és un camp que ja resta fora de la tesi, tot i que és fàcil deduir que poden ser-hi presents força perspectives distintes, que utilitzin diferents graus d'abstracció i que donin rellevància a distintes variables. Per a abastar tots els fenòmens d'acció col·lectiva, es presenta una instrumentalització de la teoria de jocs més àmplia que la convencional, que serveix com a mecanisme formal per integrar la diversitat d'orientacions d'interès i formes de decisió -racionals o no- en models que permeten interpretar els resultats com la producció de béns públics. Així, es fa necessari considerar que la hipòtesi de racionalitat -o les seves possibles variacions- és separable del nucli bàsic de la teoria de jocs i que sols és un mecanisme de decisió, entre altres de possibles (normes socials, aprenentatge evolutiu, etc.). En aquest sentit, es trenca el dilema entre solució «externa» i «interna», en introduir la teoria de jocs -protagonista habitual de les solucions «internes»- com un instrument formal que facilita el desenvolupament de les solucions «externes».

Als capítols anteriors s'ha examinat amb detall els diversos components d'aquestes afirmacions -que constitueixen una breu síntesi del treball elaborat-, tant els arguments destacats per la seva defensa, com els matisos i crítiques establertes en relació a cada aspecte. Al llarg del text no sempre s'articulen completament les diverses conclusions referides a aspectes concrets, degut a l'especificitat dels

diversos temes analitzats. És precisament el propòsit d'aquest darrer capítol emfasitzar el paper d'aquestes conclusions parcials dins la interpretació global que es proposa en la tesi. En aquest sentit, s'aporta més exactitud en la definició i situació d'aquests interrogants, per contribuir a una clarificació de les limitacions essencials -no resoltes- que presenta el paradigma científic analitzat. Per a tal fi, els problemes centrals de l'argumentació efectuada s'analitzen successivament en aquestes conclusions a través de cinc punts claus:

- i) el problema de l'estatus metodològic de la hipòtesi de racionalitat,
- ii) els supòsits inclosos en la noció de comportament racional i els problemes que genera la seva utilització a la teoria de jocs,
- iii) les distintes lògiques d'aplicació de la teoria de jocs a les ciències socials per constituir un enllaç entre els fenòmens agregats i els supòsits sobre individus,
- iv) la reconceptualització dels processos d'acció col.lectiva per facilitar la connexió entre les solucions «internes» i «externes»,
- v) la lògica de la construcció teòrica a partir de l'acció individual i la discussió sobre els seus límits.

## I.

La hipòtesi de racionalitat desenvolupada pels economistes sobre el comportament dels individus ha estat present extensament al conjunt dels capítols elaborats, tant per discutir els seus problemes interns, analitzant com pot servir per a fonamentar el desenvolupament de teoria social (mitjançant la teoria de jocs); com en un camp aplicat, examinant la seva instrumentalització en relació a les necessitats d'aplicació empíriques. Tot i reconeixent l'existència de diverses interpretacions, s'ha defensat una perspectiva sobre la hipòtesi de racionalitat: és convenient cercar fórmules per a millorar la plausibilitat de la seva aplicació. Aquesta havia d'aportar una certa capacitat de representació de les decisions dels individus, que no

contradigués la validesa de les proposicions sobre fenòmens agregats quan aquestes eren contrastades (vid. apt. 1.1.2.). Però la noció d'una racionalitat que gestiona mitjans i finalitats també pot ser entesa des d'altres perspectives. Per exemple, la concepció de Max Weber sobre l'acció racional es constitueix a partir de la formulació d'un concepte de decisió en forma de «tipus ideal». Weber argumentava que el seu punt de referència a contrastar era de natura metodològica, no empírica; així, encara que el «tipus ideal» no coincidís amb les observacions empíriques, aquest servia per proposar esquemes lògics que facilitessin la interpretació<sup>2</sup>.

No obstant això, es pot apuntar que la teoria econòmica ha desenvolupat una altra perspectiva de la utilització d'aquesta noció de racionalitat instrumental, que ha estat dominant en la seva expansió en altres ciències socials. Aquesta perspectiva intenta emprar la noció de racionalitat com un fonament per a la formulació de lleis i proposicions teòriques sobre diversos aspectes de la societat. En contrast, s'argumenta a la tesi que no es poden rebutjar completament els requisits de base empírica dels fonaments i, per tant, que la hipòtesi de racionalitat no es pot constituir en un principi d'aplicació universal (vid. secció 1.3.). És una expectativa legítima de totes les ciències intentar explicar la major quantitat possible de fenòmens amb els mínims requeriments possibles. La proposta d'un concepte unívoc de decisió racional intenta apropar-se a aquesta perspectiva, en no requerir gairebé cap hipòtesi addicional, i poder ésser aplicada indistintament a molts àmbits de les activitats humanes. Així, l'aparició d'aquest tipus d'explicacions en el camp de la ciència política i social s'ha convertit en un desafiament contemporani a la tradicional divisió de treball entre les disciplines econòmiques i socials, en què aquestes darreres gairebé no empraven models basats en la racionalitat. La possibilitat d'explicar «molt amb poc» ha atret recentment molts científics socials al

---

<sup>2</sup> Per analitzar la realitat empírica, Weber proposava la necessitat de combinar generalment amb altres elements interpretatius: "*Innerhalb des (subjektiv) sinnhaft auf das Handeln anderer bezogenen Handelns, des Gemeinschaftshandels also, wird nun das rationale Handeln von der Theorie deshalb als 'Idealtypus' gebraucht, weil es besonders eindeutig und ohne alle 'psychologische' Erörterungen konstruierbar ist*" a "Rationale Konstruktion und empirische Wirklichkeit", *Methodologische Schriften*, S. Fischer Verlag, Frankfurt, 1968, pàgs. 166-167.



camp de les teories sobre la racionalitat (vid. apt. 1.1.4.), decebut per les típiques aproximacions o bé bàsicament empíriques o bé excessivament globalistes. Ha afavorit també aquesta atracció el creixement de l'instrumental analític desenvolupat pels economistes, sofisticant les inicials assumpcions «naive» sobre el comportament racional que caracteritzava tradicionalment l'enfocament econòmic (vid. secció 1.2.). La major sensibilitat empírica dels científics socials continua fent difícil l'acceptació d'explicacions basades essencialment en supòsits sobre el comportament racional, la qual cosa planteja una contínua discussió sobre l'aplicació d'aquest enfocament en els seus àmbits.

Un aspecte destacat de la recerca desenvolupada en aquest treball ha estat l'anàlisi dels intents de fer més flexible la hipòtesi de racionalitat. En la recerca de la plausibilitat, han estat explorades dues direccions per reformular una noció empíricament fonamentada de la racionalitat: i) els intents de sofisticar alguns dels postulats sobre la noció convencional de racionalitat, especialment en relació a les actituds davant el risc i a les eleccions múltiples (vid. apt. 1.3.2.); i ii) la recerca d'un concepte distint de racionalitat instrumental, que no consideri el principi d'optimització, com la formulació de la racionalitat limitada, que utilitza la noció de satisfacció (vid. apt. 1.1.4.). Amb aquests intents, augmenta el nombre de situacions en que es més plausible l'aplicació de versions de la hipòtesi de la racionalitat. Però no s'aconsegueix una capacitat d'aplicació general i, a més, sorgeix el problema d'escollir entre les diverses variacions de la hipòtesi possibles. Així, el procés de reformulació de la hipòtesi de racionalitat, per seguir essent utilitzada com a principi únic de construcció teòrica, en desaparèixer els criteris normatius, constitueix un conjunt mancat de capacitat determinista. Sense desestimar les possibilitats de ser emprada com a «tipus ideal», l'exploració que es realitza recobra el seu sentit en aportar molts suggeriments sobre les possibilitats d'aplicar hipòtesis de racionalitat com a forma plausible de comportament en un context teòric en què aquesta no és una constant dels processos de decisió inherents a les activitats humanes. En la caracterització d'un determinat camp d'anàlisi, la introducció d'altres principis de

comportament, sense excloure la possibilitat d'aplicar els supòsits sobre racionalitat, pot facilitar l'explicació del conjunt de fenòmens analitzats.

## II.

Aquest punt es dedica a les formes de presentar acuradament una hipòtesi de comportament racional. Els economistes treballen amb una hipòtesi de comportament racional que inclou una gran varietat de supòsits, els quals no sempre són coincidents. Una vinculació freqüent associa la racionalitat al supòsit que els individus busquen l'interès propi en les seves accions. Per aquest motiu, es considera que els agents ordenen en funció d'aquest principi les preferències sobre els resultats possibles de les seves accions, però també es poden considerar formalment possibles altres orientacions d'interès amb una consistència similar (vid. apt. 1.1.3.).

Força fenòmens socials tan sols poden ser aclarits, quan hom desestima la invariabilitat del supòsit d'interès propi, i introdueix altres de complementaris, establint models sobre la motivació del comportament més adequats, sense descartar un predomini en els comportaments individuals de les motivacions egoistes. Les combinacions de supòsits que es suggereixen no han d'implicar la construcció de cap alternativa general, ja que ressorgirien els problemes de plausibilitat esmentats. Es tracta de formular models específics, o combinacions d'aquests, per a diferents contextos i categories de fenòmens socials (vid. apt. 1.2.1.). Hom pot mantenir la resta d'elements que constitueixen el model d'home presentat, acceptant la introducció d'altres possibles interessos o objectius en el comportament dels individus, per tal de contrastar la capacitat analítica del supòsit de racionalitat.

En establir l'existència d'una orientació d'interès individual -egoista o no-, cal associar aquesta, si es vol fer mesurable, amb un principi de comportament o un índex sobre la satisfacció personal. Aquesta associació és necessària per definir

preferències cardinals, ja que solament amb la hipòtesi d'elecció consistent és impossible justificar el procés de satisfacció individual (vid. apt. 1.2.2.). Aquests problemes de connexió també apareixen, essent inclús més evidents, en analitzar les propostes de fonamentar les decisions sota risc. L'axiomatització conjunta de les preferències i la probabilitat ha estat la via per mesurar la satisfacció en situacions d'informació incompleta<sup>3</sup>. Tot i que les propostes existents actualment presenten suggeriments interessants per establir principis d'ordenació de preferències integrats en un model d'home racional, que tenen en compte consideracions psicològiques, sembla difícil d'escollir una d'aquestes propostes com fórmula de connexió general entre un procés de decisió completament racional i un sol procediment de mesurar la satisfacció (vid. apt. 1.3.2.).

En el context esbossat, la hipòtesi de racionalitat es defineix formalment com un conjunt de relacions per escollir en situacions d'informació completa (consistència, equivalència) o incompleta (consistència, continuïtat, independència), que pretenen explicar com els individus estableixen les seves preferències i relacionen aquestes amb les seves eleccions<sup>4</sup>. Cal remarcar que tots els elements de la hipòtesi de racionalitat han estat molt discutits en el terreny empíric (vid. apt. 1.3.3.). Per exemple, la consistència apareix alterada perquè a la vida real les eleccions no solen presentar-se en forma aïllada. Les investigacions de caire experimental han mostrat que la formació de preferències i la realització d'eleccions no són en absolut processos completament dependents, i que pot aparèixer la

---

<sup>3</sup> Hi ha propostes des dels anys cinquantes (vid. apt. 1.3.1.): per un costat, sobre com es formen o es construeixen les preferències; d'altre costat, sobre com es gestionen aquestes amb els supòsits sobre racionalitat. A partir de les discussions sobre la plausibilitat de l'axioma d'independència, s'ha suggerit la seva substitució per un índex d'utilitat més complex vinculat a les actituds variables de comportament observades en relació al risc (vid. apt. 1.3.2.).

<sup>4</sup> Si les creences, eleccions i les accions, o només les eleccions s'han d'incloure en aquesta estructura interpretativa és una disputa interna dins l'enfocament, amb diversos punts de vista. Amb una perspectiva globalista, Elster apunta que *"rational action, then, involves three optimizing operations: finding the best action, for given beliefs and desires; forming the best grounded belief, for given evidence; and collecting the right amount of evidence, for given desires and prior beliefs. Here, desires are the unmoved movers.."* (*Nuts and Bolts...* op. cit., 1989, pàg. 4). Altres autors, especialment dins la teoria econòmica, tan sols estan interessats en la primera d'aquests tres operacions.

reversió de preferències en una elecció específica. La utilització de l'axioma d'independència (de les alternatives irrellevants) és un dels problemes més discutits a la literatura sobre eleccions en situació d'incertesa. En situacions de risc, aquest axioma estableix que les preferències probabilístiques són una funció lineal, normalment convexa, que reflecteixen una aversió al risc. Però moltes observacions experimentals mostren l'existència d'actituds favorables al risc quan els individus es troben en el terreny de les pèrdues. Fora del camp formal, els intents més interessants per explicar les eleccions sota incertesa apareixen quan s'analitzen empíricament les distorsions cognitives que afecten el comportament, tot i que aquestes anàlisis no construeixen una teoria de l'elecció aplicable per desenvolupar proposicions agregades. Pel moment, les discussions segueixen, amb una diversitat de propostes sobre variacions dels postulats de racionalitat, que indiquen la inexistència d'una solució única al problema.

D'aquestes controvèrsies es pot deduir que definir una noció única i plausible de comportament racional és cada cop un intent més quimèric. Hi ha eleccions clarament irracionals, però un cop descartades, el que apareix és un espai de possibilitats per a definir una conducta racional, que van des de l'eliminació d'algunes de les relacions que integren la hipòtesi de racionalitat, fins a la introducció de supòsits més sofisticats sobre les actituds dels individus davant el risc. Aquestes variacions formen un conjunt de mecanismes simples per expressar diverses formes de càlcul autònom de decisions (no afectat per distorsions o per normes socials, per exemple), que poden ser aplicats com a principi per explicar el comportament. L'elaboració d'una hipòtesi amb capacitat descriptiva de la racionalitat -davant la impossible obtenció d'una noció ideal amb capacitat descriptiva- en els models explicatius ha de prendre un sentit operacional des d'una nova perspectiva, segons la qual les necessitats de construcció de diverses interpretacions teòriques de fenòmens polítics o socials, aconduïxen a suposar distintes formes de càlcul (segons les eleccions simultànies a fer, la quantitat d'informació a processar, la significació del risc per l'elecció, etc.), moltes de les

quals mantenen elements que formen part de la hipòtesi de la racionalitat.

Del conjunt d'elements que formen un model d'home des de la concepció del paradigma analitzat, es distingeixen a la tesi dos grans blocs analítics, referents als dos distints nivells que es consideren presents en l'elecció. El primer, l'anàlisi de la formació de les preferències i de la definició dels interessos dels individus; el segon, referent als mecanismes i principis de decisió amb què s'adopta una elecció. Així, es defensa que l'articulació de les preferències generades per les orientacions d'interès, i els procediments de càlcul de decisions que estableixen les accions i estratègies que els individus adopten són dos processos de natura distinta, que es poden analitzar separadament, encara que habitualment es presentin de forma indissoluble (vid. secció 4.4.). El primer nivell pot ser més alterable segons els contextos, èpoques i camps interpretatius (per exemple, en els assumptes públics les motivacions poden ser menys orientades a l'interès personal, que en els assumptes privats), i un tipus ideal d'interès pot ser formulat fàcilment per a àmplies categories de fenòmens. D'altra banda, els procediments de càlcul estan influenciats per capacitats cognitives i físiques dels éssers humans, i tenen uns punts de referència normatius que poden ser comunament acceptats. En tot cas, sobre aquests temes resten moltes preguntes obertes sobre la seva interpretació i formulació, essent un camp obert pel desenvolupament de futures línies d'investigació a les ciències socials.

Existeixen dos elements conceptuals difosos en molts camps científics, que són formalment externs al model d'home, però que s'hi vinculen per acabar de caracteritzar el paradigma que ha desenvolupat la teoria econòmica. Aquests són: i) el principi d'optimització, considerat com una actitud permanent en el comportament dels individus, i que serveix per vincular els postulats sobre racionalitat amb l'interès individual, i la seva representació a través de la funció d'utilitat; i ii) la solució d'equilibri, que és l'estat ideal d'estabilitat, el qual es considera que pot ser assolit pels individus a través de les seves interaccions (vid.

apt. 2.1.1.). Cal cercar en el comportament dels individus i en les condicions dels contextos els factors que aconduïxen a aquest equilibri o bé que l'impedeixen. Com aquests processos constitueixen la pròpia explicació de les accions i els seus resultats, també és necessari definir la forma com s'assoleix l'equilibri, des de la perspectiva que ofereix aquest enfocament.

El principi d'optimització és una noció matemàtica que té un estatus especial en el paradigma que es discuteix. El seu rebuig, implica greus problemes per l'estructura interpretativa d'aquest paradigma, ja que qualsevol mecanisme de càlcul de decisions basat en una noció de racionalitat, es suposa que aplica aquest criteri matemàtic per determinar -igualant, maximitzant o minimitzant, entre altres possibilitats- qualsevol funció de satisfacció (vid. apt. 2.2.5.). Un exemple per qüestionar aquesta situació és la teoria de la racionalitat limitada. Quan s'han intentat explicar els principis d'aquesta a través de l'enfocament de l'«elecció racional», el concepte de satisfacció en la recerca d'objectius ha estat substituït pel concepte d'optimització en la recerca d'informació, o per l'aplicació de regles en situacions d'excessiva incertesa, per esmentar diversos exemples (vid. apt. 1.1.4.).

En tot cas, com s'ha observat a través de diversos models, el que també es pot proposar és substituir l'optimització per l'aplicació de la probabilitat, mitjançant distints mecanismes que «gestionin» racionalment les possibilitats de combinació, segons la intensitat de les motivacions. Però sovint cal aproximar cas per cas aquesta lògica de la substitució, degut a què si es rebutja el supòsit d'optimització, introduint alguns elements de decisió probabilístics junt a mecanismes de decisió racional, s'han d'introduir molts elements contextuals per explicar cada fenomen concret que s'observi.

**III.**

La teoria de jocs no és l'únic procediment per a constituir un enllaç formal entre l'anàlisi de fenòmens agregats i els supòsits sobre individus, però sí és un dels més sofisticats, útil per aplicar a algunes de les situacions més complexes (vid. apt. 4.2.1.). Aquest instrument matemàtic comparteix, junt amb altres constructes formals emprats a les ciències socials, la noció d'equilibri com l'element central a determinar en els seus models. Aquesta teoria cerca l'existència de punts d'equilibri en interaccions, punts en què ningú és capaç de millorar el seu resultat sense empitjorar el dels altres. Generalment s'assumeix que l'equilibri es deriva de les expectatives racionals dels agents que calculen interactivament (vid. apt. 2.1.1.). Però s'ha demostrat formalment que per trobar un equilibri sols és necessari assumir l'existència de decisions, així com d'algun mecanisme de canvi (o no canvi) de decisions en funció dels resultats obtinguts (amb una funció lineal i una distribució de probabilitat es pot provar l'existència d'equilibri (vid. apt. 2.2.4.).

Seguint un argument similar, la teoria de jocs ha estat utilitzada en el camp de la biologia evolutiva o en l'anàlisi de comportament humà guiat per normes, sense tenir necessitat dels supòsits de racionalitat (vid. apt. 2.2.3.). Però entre les normes i la racionalitat, o en altres termes, entre decisions a les quals es suposa un baix nivell i amb un alt nivell de complexitat en la seva formulació, sembla més plausible d'aplicar l'ampli camp de procediments de decisió que apareix entre els dos extrems (vid. apt. 2.2.5.). L'espai de variacions de la hipòtesi de racionalitat esbossat a l'apartat anterior, acceptant l'existència d'interessos ja prèviament definits (vid. apt. 2.2.2.), permet disposar d'un conjunt de formes de comportament racional amb major capacitat descriptiva, que de forma no competitiva, suggereixen una pluralitat de teories de jocs, o de desenvolupaments d'aquesta teoria, que també poden aportar solucions d'equilibri.

Desenvolupar aquesta aproximació és un propòsit que no està lliure de greus

problemes, ja que en augmentar la variabilitat, el nombre d'equilibris possibles tendeix a augmentar, cosa que dificulta la determinació d'una solució pels models d'interacció que es formulen. Existeixen diverses formes de mantenir la possibilitat de trobar solucions d'equilibri (vid. apt. 2.2.5.). Una possibilitat és limitar la capacitat dels jugadors de recollir informació, mantenint la visió d'aquests com maximitzadors d'una funció lineal d'utilitat. Altres possibilitats consisteixen en emprar models de formació de preferències més sofisticats, amb funcions d'utilitat quadràtiques, o que segons les condicions d'incertesa s'apliquin determinades regles. Existeixen propostes que proveeixen un concepte de solució d'equilibri en teoria de jocs sense necessitat de l'axioma de transitivitat. També és possible desestimar l'axioma d'independència i utilitzar funcions no lineals -mantenint els altres supòsits, la consistència i la continuïtat- per trobar solucions d'equilibri. Finalment, un tercer camí és desenvolupar models amb actors limitats cognitivament en la seva capacitat d'optimització. Una altra possibilitat és no acceptar les estratègies mixtes en la modelització dels jocs. Això provoca la impossibilitat d'una demostració general d'existència d'equilibri, però no impedeix explicar l'aparició d'equilibris en estratègies pures. La combinació de la teoria de jocs amb nocions de racionalitat limitada possiblement és l'enfocament més plausible de tots els ressenyats, però també és el que presenta més dificultats per a ser integrat en la teoria de jocs, degut a les dificultats per definir operativament la noció de satisfacció, aplicada en substitució de la d'optimització (vid. apt. 2.2.5.).

Aquestes vies per introduir distintes variacions de la hipòtesi de la racionalitat a la teoria de jocs, tenen diversos problemes formals, que són força discutits al capítol segon de la tesi, però la qüestió central és que presenten diverses possibilitats per tractar el conjunt de mecanismes de càlcul racional destacat a l'apartat anterior. Sense afegir elements conceptuals nous, l'estructura de la teoria de jocs aporta en cada cas propostes de solució, a través de la definició d'equilibris, entre decisions individuals en interacció (vid. apt. 2.2.1.). A més, cal afegir que fora dels mecanismes de càlcul racional, la teoria de jocs també pot ser aplicada a la recerca



d'equilibris en interaccions d'agents amb decisions estables (o sigui, fixades per convencions, o a través de normes socials). La diferència més significativa entre la diversitat de mecanismes de decisió proposats per aplicar a la teoria de jocs, i el mecanisme convencional basat en la utilitat esperada es concentra en el problema de la determinació dels equilibris, especialment en relació a la recerca de demostracions de validesa general, ja que la possibilitat de trobar solucions úniques pels jocs que es formulen és més reduïda. Apareix una major necessitat de combinar elements contextuals per precisar les solucions entre els equilibris possibles, encara que l'estructura analítica bàsica segueix essent suportada per la teoria de jocs. Deixant pel moment aquest darrer objectiu per les propostes normatives, es pot considerar que per establir proposicions sobre els processos socials no cal tan demostracions de validesa general, sinó especificacions ben determinades d'equilibris (vid. apt. 4.2.2.). En aquest sentit, tant la introducció de supòsits de simplificació sobre els jocs, com de supòsits sobre l'adquisició d'informació per part dels jugadors, han de contribuir a facilitar l'adopció a la ciència política d'explicacions plausibles sobre la interacció basades en la teoria de jocs, com un dels seus instruments formals més sofisticats.

#### IV.

Les possibilitats de l'estructura formal de la teoria de jocs per ésser aplicada al desenvolupament de la ciència política han sigut discutides a la segona part de la tesi. Per una banda, els problemes d'aplicació es deriven dels processos de simplificació que cal efectuar, bàsicament sobre els supòsits computacionals de la teoria de jocs; d'altra banda, la reinterpretació formulada sobre alguns elements del paradigma de l'«elecció racional», implica la necessitat d'explorar com en resta afectada l'anàlisi del camp de fèndems escollit, en el nostre cas, la lògica de l'acció col·lectiva. El conjunt d'estructures d'interacció que es poden interpretar mitjançant la teoria de jocs genera uns problemes teòrics específics per l'anàlisi de l'acció

col·lectiva. Aquests problemes, i no una construcció teòrica específica, és el que s'ha intentat definir i analitzar al llarg del treball presentat.

L'obra de M. Olson va suposar una important contribució a la interpretació dels fenòmens socials des d'una perspectiva basada en la hipòtesi de comportament racional. La seva anàlisi del funcionament de les associacions d'interessos a partir dels objectius individuals té com a punt central de l'argumentació la confrontació entre la racionalitat del comportament individual i la racionalitat dels objectius col·lectivament desitjats. Per Olson, un bé col·lectiu no es produeix només per la coincidència d'interessos d'un conjunt d'individus sobre la seva necessitat, en moltes circumstàncies també són necessaris estímuls individuals, per reforçar la contribució (vid. apt. 3.1.2.). En introduir elements de caràcter institucionalista per definir aquests estímuls -com els incentius-, Olson intentava integrar una nova perspectiva en les explicacions fonamentades en la hipòtesi de racionalitat dels individus (vid. apt. 3.2.1.). Olson analitza paral·lelament dos processos en la «lògica de l'acció col·lectiva», el primer, la relació dels individus amb la natura del bé públic en que estan interessats; el segon, les relacions interactives entre els individus degudes a la seva necessitat de coordinar-se per proveir el bé públic (vid. apt. 3.1.2.). En ambdós processos, que no acaba de relacionar teòricament de forma completa (vid. apt. 3.4.1.), Olson introdueix elements externs a la hipòtesi de racionalitat per desenvolupar les seves explicacions.

Posteriorment, la teoria ha estat molt analitzada, modificada i refinada analíticament, però s'ha mantingut la línia interpretativa definida per M. Olson. El seu desenvolupament ha aconduït a una integració més general de la teoria dels béns públics amb la lògica de l'acció col·lectiva, la qual cosa ha permès de relativitzar algunes de les afirmacions rotundes d'aquest autor (vid. apt. 3.4.3.). S'ha interpretat i definit un ampli ventall de «problemes» d'acció col·lectiva, segons la natura i característiques del bé públic en qüestió, integrant els dos processos que Olson analitzava per separat (vid. apt. 4.3.1.). El que ha quedat clar d'aquests refinaments,

és que l'obra d'Olson tot i ésser molt remarcable, era una síntesi teòrica que deixava més interrogants oberts dels que tancava (vid. apt. 3.2.2.). Per exemple, algunes intuïcions psicològiques claus sobre el comportament dels grups petits, són integrats, però no explicats, en la interpretació que proposa Olson sobre el funcionament dels grups d'interessos (vid. apt. 3.4.2.).

En resum, amb la seva síntesi, Olson deixà obert un camp de nous problemes, que han afrontat els qui han analitzat o refinat la seva proposta teòrica. Per un costat, proposant solucions «externes», s'ha comprovat el paper dels elements institucionalistes, rebutjant-se alguns i afegint-hi altres, formulant-se models de relacions més complexes i amb més variables, sense qüestionar ni integrar la hipòtesi bàsica d'Olson sobre el problema de l'acció col·lectiva (vid. secció 3.5.). D'altre costat, s'han extret de la interpretació d'Olson tots els elements que completaven el seu sentit, per a formular solucions «internes» força sofisticades en les que un supòsit normatiu de racionalitat sobre el comportament havia de justificar exclusivament el resultat de l'acció col·lectiva (vid. apt. 4.2.2.). Vint-i-cinc anys després de la seva publicació, el creixement teòric a partir de la síntesi era realment molt escàs, ja que l'argumentació d'Olson era sistemàticament rebutjada o implausiblement sofisticada. Davant aquesta situació, hom podria reflexionar que el propòsit d'Olson era força il·lusori, degut a què les tradicions teòriques que volia integrar estaven massa allunyades, amb coherències internes que perdien el seu sentit en connectar-se. De forma consistent, hom podria afegir que aquesta situació ha estat la causa de moltes confusions, apel·lacions i refutacions sense sentit en relació a l'obra d'Olson. Però també es pot argumentar que tot i tractar-se d'un problema teòric important, la font de confusió fonamental era deguda a lectures distintes des de tradicions científiques molt allunyades entre si -com es mostra en el tercer capítol-.

En tot cas, allò més important a destacar és que és possible construir una interpretació homogènia de l'acció col·lectiva, que englobi una àmplia diversitat de situacions d'interacció. Per definir aquestes situacions, el primer pas ha estat trencar

amb la hipòtesi determinista sobre l'existència d'un sol procés i un sol problema bàsic en l'acció col·lectiva (vid. secció 4.1.). Mantenint els supòsits sobre la racionalitat que constitueixen l'eix central de l'argumentació d'Olson, aquest trencament s'aconsegueix considerant només que els béns públics que es volen proveir col·lectivament poden presentar característiques de producció i de consum força diferents. D'aquesta primera proposta ja es pot observar que té poc sentit buscar una teoria general de l'acció col·lectiva, i que cal derivar la discussió cap a la incidència de les distintes variables sobre diverses modalitats de comportament interactiu, i no tant sobre un model exclusiu que integri tots els tipus possibles d'acció col·lectiva (vid. 4.3.1.). Si s'accepta l'existència de comportaments distints en els processos d'acció col·lectiva, la discussió sobre les solucions «internes» o «externes» pren una altra dimensió, ja que gràcies a la major variabilitat de l'acció col·lectiva presentada, es pot evitar sovint pressuposar la presència d'intervencions de segon ordre per explicar conductes de participació (vid. apt. 4.4.1.).

La representació dels fenòmens d'acció col·lectiva en models de teoria de jocs a través de la seva formulació prèvia en forma de dilemes de provisió de béns públics és l'estructura analítica proposada. Això permet integrar altres mecanismes de decisió, apart de la racionalitat (vid. apt. 4.4.3.). Els més freqüents entre aquests són les convencions, que es refereixen a mecanismes de decisió formats mitjançant la coincidència d'expectatives entre individus que interactuen. L'aparició de convencions facilita -com a clau única- la interacció entre els membres d'un grup, i afegeix seguretat a la interacció, inclús amb desconeguts, en estalviar recursos, en no haver de cercar informació i gastar temps per decidir (per exemple, es pot destacar l'estabilitat de les tradicions). Amb aquesta estratègia, hom pot mostrar que la modelització comparteix uns elements metodològics, degut al seu tractament a través de la teoria de jocs (per exemple, la definició d'estratègies, actors, etc.), i que també comparteix altres elements definitoris dels problemes analitzats (per exemple, la conceptualització dels models d'interacció a través dels caràcters distintius dels béns públics). Malgrat tot, els models divergeixen quan s'han d'aplicar a entorns

determinats, que presenten característiques específiques de relació entre els agents (vid. apt. 4.2.2.). Hom pot considerar que és precisament en aquests entorns sobre els que cal especificar les característiques de la «situació d'elecció» -com poden ser les intervencions de segon ordre-, per a disposar de models amb àmplia capacitat d'aplicació.

Finalment, cal destacar que al llarg d'aquest procés de discussió, han sorgit diversos problemes per aplicar la teoria de jocs. Primer, sobre la qüestió de com solventar els problemes derivats dels elevats requisits de computació que generalment es suposen en aplicar la teoria de jocs; i en segon lloc, sobre com es poden vincular i ampliar conceptualment les diverses formes de modelització en termes de teoria de jocs existents sobre l'acció col·lectiva, mantenint i afermant alhora una perspectiva interpretativa homogènia. Al llarg del treball es proposen distintes línies per avançar en la resolució d'aquests problemes, com per exemple la introducció d'alguns conceptes específics per interpretar les possibles formes de connectar amb el marc institucional, o l'elaboració d'una lògica de la simplificació dels models de la teoria de jocs; però encara resten força problemes pendents, solventats només provisionalment.

Acceptant una resolució provisional d'aquests darrers problemes, hom pot afirmar, que tot tipus de processos d'acció col·lectiva poden ser interpretats a través de models d'interacció basats en la teoria de jocs, sense necessitat de què integrin sempre el supòsits de racionalitat -aquesta és una discussió paral·lela, però no dependent-, ja que la teoria de jocs pot integrar tot tipus de mecanismes de decisió (vid. apt. 2.2.1.). Per il·lustrar-ho amb la definició: hom pot entendre els fenòmens d'acció col·lectiva com els processos interactius entre individus per a la consecució de béns públics.

En conjunt, cal afegir que la perspectiva establerta per l'anàlisi de l'acció col·lectiva -a través de l'individualisme metodològic i la teoria de jocs- aconsegueix un

major grau explicatiu dels processos socials i polítics en què es produeix la interacció entre individus, perquè aconseguix formalitzar els elements bàsics de les «situacions d'elecció» que generen relacions interactives entre individus, i a través d'aquestes, oferir una explicació del resultat de les eleccions. Això no obstant, cal remarcar -des d'aquesta perspectiva- que el desenvolupament no serveix per oferir una teoria general de l'acció col·lectiva. La reflexió efectuada permet concloure que almenys -deixant a banda la discussió sobre les modalitats de construcció teòrica a partir de la racionalitat- l'anàlisi dels problemes de l'acció col·lectiva proposada aconseguix oferir una definició, un tractament homogeni del fenomen, i la possibilitat d'aclarir un dels seus aspectes més destacats: el resultat col·lectiu de les accions individuals, coneixent o suposant les preferències i els mecanismes de decisió dels agents que intervenen en el procés, així com la natura del bé a produir.

## V.

Rescatant les consideracions realitzades fins al moment sobre la hipòtesi de racionalitat, sembla clar que no es pot defensar l'existència d'una definició unívoca de racionalitat dins el paradigma analitzat, sinó que el principi de comportament racional encobreix un conjunt de mecanismes de decisió, més o menys adaptats a distintes condicions d'elecció. Aquesta diversitat augmenta la plausibilitat de la seva aplicació com a hipòtesi de comportament, sense que es pugui arribar a defensar que aquests mecanismes es tractin de lleis de comportament empíricament contrastades, alhora que també perd sentit la confrontació sobre l'eficàcia normativa entre les diverses propostes existents de conceptualitzar la racionalitat.

Mantenir una presumpció de comportament basat en la racionalitat, té el gran avantatge que permet disposar d'una perspectiva interpretativa estable i compartida sobre tots els fenòmens socials. Però com s'ha il·lustrat a través de la discussió sobre la interpretació dels processos d'acció col·lectiva, apareixen fenòmens

difícilment explicables amb aquest principi (les intervencions de segon ordre, l'aparició de convencions), tot i considerar possible l'assumpció de diverses orientacions d'interès -formalment desvinculada del supòsits de racionalitat-. Al marge de la discussió sobre la modificació de la hipòtesi de racionalitat, es planteja el problema de com proposar explicacions on els principis de comportament puguin ser variables (en altres paraules: com explicar una «situació d'elecció» que determina la forma com es pren la decisió, però no determina el resultat). La resposta al problema plantejat no és resolta en aquesta tesi, però sí que, des d'una opció, s'estableixen algunes pautes encaminades a definir les possibles vies de resolució del problema plantejat (vid. secció 4.4.). Entre els elements interpretatius aportats prèviament, destaca la construcció d'un marc interpretatiu comú, basat en l'aparell analític de la teoria de jocs, que pugui integrar tots els fenòmens d'acció col.lectiva, gràcies a la flexibilització dels supòsits que guien la teoria de jocs convencional, junt a la introducció d'altres supòsits simplificadors sobre les condicions dels jocs (vid. secció 4.3.). En mostrar que existeix un nucli bàsic en els fenòmens d'acció col.lectiva que pot ser analitzat mitjançant els models de la teoria de jocs -els processos d'ajust interactiu entre individus en presència de béns públics-, es disposa d'un entorn comú per discutir el paper de la hipòtesi de racionalitat en l'anàlisi d'aquests fenòmens, sense que l'hipotètic rebuig d'aquest supòsit com a principi d'aplicació general impliqui la desaparició del marc analític. Disposant d'aquesta entorn comú, es poden discutir amb claredat els problemes de les distintes utilitzacions de la hipòtesi de racionalitat, que van d'ésser un principi inqüestionable per a la construcció teòrica (la línia principal que sorgeix de l'obra d'Olson), a ser considerada com un element addicional dels models d'acció individual basats en normes socials (els darrers treballs d'Elster), passant per ésser interpretada com un principi central, però no insubstituïble, en la formulació teòrica dels processos d'acció col.lectiva (la perspectiva dels models proposats per Schelling).

Si hom defensa una utilització no inqüestionable de la hipòtesi de racionalitat, apareixen implicacions importants per la capacitat de generalització de les propostes

teòriques basades en l'explicació intencional. Algunes d'aquestes implicacions es poden presentar en forma de preguntes, que suggereixen diverses línies de treball:

- Cal abandonar l'objectiu de cercar una teoria elemental del comportament humà, aplicable a un alt nivell d'abstracció per tots els àmbits socials possibles, i desenvolupar enfocaments teòrics sobre camps específics, en els quals es puguin aplicar mecanismes de decisió -no idèntics en tots els casos- adequats als contextos, no necessàriament d'arrel empírica?.

- Com s'ha d'interpretar i explicar la influència dels contextos sobre la determinació dels mecanismes de decisió?

- Implica aquesta situació que són els elements del context els que prenen més rellevància explicativa?.

- Podrien ésser aquests elements, des de caràcters tècnics i físics, fins a resultats de decisions anteriors, passant per constructes institucionals, integrats en una mateixa perspectiva interpretativa?

En tot cas, un cop destacades aquestes incògnites, sí es pot assenyalar que el que unifica l'enfocament de l'«elecció racional», marcant una diferència fonamental amb perspectives funcionals, estructurals i sistèmiques, és que l'aparell analític que s'utilitza, junt a la noció d'equilibri com a punt de referència interpretatiu i com a estructura inductiva comuna, comporta la perspectiva de que els resultats globals depenen de les accions individuals, i aquestes, de preferències i eleccions individuals. Davant la perspectiva tradicional que passava dels principis a l'aplicació sobre fenòmens concrets, degut al supòsit implícit que sols existeixen restriccions a la capacitat individual d'optimitzar quan s'introdueixen supòsits de comportament variables, es transforma la perspectiva de l'enfocament, per centrar-se en l'elaboració de propostes teòriques d'abast intermedi. Sinó, l'arbitrarietat en l'elecció de principis generals de decisió invalidaria qualsevol pretensió d'explicació teòrica (si no fos que hi hagués una determinació molt estricta d'altres teories).



Per acabar amb una paradoxa, es podria afirmar sorprenentment que la hipòtesi de racionalitat no té límits. Hi ha problemes, però és el context que hom adopta per aquesta hipòtesi, el que assenyala si són significatius o no els límits de la plausibilitat i la contrastació empírica de les aplicacions que es desenvolupen. Aparentment, pot semblar que hi ha una relació enfrontada entre la manca de plausibilitat i la dificultat per aconseguir capacitat de predicció; però aquest és tan sols un problema relatiu, ja que en totes les ciències es juga amb aquests dos factors, per aconseguir les prediccions més plausibles. El que és més significatiu per aquest dilema concret de les ciències socials és com integrar i vincular l'entorn o el context en la teoria social que es formula a partir de proposicions sobre l'acció individual: el problema és que incorporar l'entorn, com alguns autors han proposat, en forma de filtres, restriccions, o camps accessibles aparta la resolució de la qüestió, no ajuda a explicar com l'acció es forma en l'entorn.

La línia de recerca que s'ha seguit no es centrava en la definició d'una teoria de l'acció, sinó en la discussió d'un dels seus prolegòmens fonamentals, el d'un supòsit de comportament racional com a hipòtesi generalitzable. En aquest sentit, no cal confondre amb les discussions sobre la teoria de l'acció a la psicologia o a la filosofia. El sentit d'una teoria de l'acció per a les ciències socials pot ser disposar d'un fonament interpretatiu que permeti establir proposicions teòriques sobre fenòmens socials, polítics, econòmics. Però també pot ser disposar d'una teoria de l'acció suficientment flexible per connectar-se amb altres enfocaments que hi aportin continguts factics. La resposta a aquest dilema resta oberta en aquestes pàgines. Però es pot suggerir que la resposta no passa per la recerca d'un altre concepte de racionalitat, com s'ha proclamat força insistentment. Aquest itinerari ja s'ha seguit detalladament al llarg de la tesi, però infructuosament pel que fa a l'objectiu principal, tot i que han sorgit subproductes força interessants. L'alternativa és aconseguir models teòrics prou convincents, els quals parteixin de les condicions tècniques i físiques externes a l'acció (com per exemple, la natura d'un bé col·lectiu), així com d'interpretacions sobre les decisions prèvies, per determinar i

explicar els factors de les «situacions d'elecció» que incideixen en els mecanismes de decisió que els individus apliquen en les seves interaccions. Així, les eleccions dels individus es poden convertir en una variable important i decisiva, però no única i exclusiva.

## **Referències bibliogràfiques**

- Abreu, D. i A. Rubinstein, (1988): "The Structure of Nash Equilibrium in Repeated Games with Finite Automata", Econometrica, vol. 56.
- Ackermann, R., (1983): "Methodology and Economics", The Philosophical Forum, vol. 14.
- Aguar, F., (1990): "La lógica de la cooperación", Zona Abierta, vol. 54-55.
- Ainslie, G., (1985): "Rationality and the Emotions, a Picoeconomic Approach", Social Science Information, vol. 24.
- Akerlof, G. i J. Yellen, (1987): "Rational Models of Irrational Behavior", American Economic Review, vol. 77.
- Akerlof, G. i W. Dickens, (1982): "The Economic Consequences of Cognitive Dissonance", American Economic Review, vol. 72.
- Akerlof, G., (1970): "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism", Quarterly Journal of Economics, vol. 84.
- Albanese, R. i D. Van Fleet, (1986): "Rational Behavior in Groups: The Free-Riding Tendency", Academy of Management Review, vol. 10.
- Alchian, A.A. i W.R. Allen, (1973): "The Pure Economics of Giving", The Economics of Charity, The Institute of Economic Affairs, Londres.
- Allais, M., (1953): "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine", Econometrica, vol. 21.
- Allais, M., (1979): "Selected Findings of the 1952 Experiment", a: M. Allais i O. Hagen (Ed.), Expected Utility Hypotheses and the Allais Paradox, Reidel, Dordrecht.
- Allais, M., (1979): "The So-Called Allais Paradox and Rational Decisions Under Uncertainty", a: M. Allais i O. Hagen (Ed.), Expected Utility Hypotheses and the Allais Paradox, Reidel, Dordrecht.

- Allais, M., (1984): "The Foundations of the Theory of Utility and Risk. Some Central Points of the Discussions at the Oslo Conference ", a: O. Hagen i F. Wenstoep (Ed.), Progress in Utility and Risk Theory, Reidel, Dordrecht.
- Allison, G., (1971): Essence of Decision, Little, Brown and Co., Boston.
- Anand, P., (1987): "Are the Preference Axioms Rational?", Theory and Decision, vol. 23.
- Anbarci, N., (1989): "The Kalai/Smorodinsky Solution with Time Preferences", Economic Letters, vol. 31.
- Andreoni, J., (1988): "Why Free Ride? Strategies and Learning in Public Goods Experiments", Journal of Public Economics, vol. 37.
- Arminger, G., (1986): "Fallacies and Paradoxes Caused by Heterogeneity", a: A Diekmann i P. Mitter (Ed.), Paradoxical Effects of Social Behavior, Springer Verlag, Heidelberg-Viena.
- Aronson, E., (1979): El animal social, Alianza Editorial, Madrid.
- Arrow, K.J., (1951): "Mathematical Models in the Social Sciences", a: D. Lerner i H. Lasswell (Eds.), The Policy Sciences, Stanford University Press, Stanford, 1959.
- Aumann, R. i S. Sorin, (1989): "Cooperation and Bounded Recall", Games and Economic Behavior, vol. 1.
- Aumann, R., (1962): "Utility Theory without the Completeness Axiom", Econometrica, vol. 32.
- Aumann, R., (1976): "Agreeing to Disagree", Annals of Statistics, vol. 4.
- Aumann, R., (1981): "Survey of Repeated Games", Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Bibliographisches Institut, Manheim-Viena-Zurich.

- Aumann, R., (1985): "What is Game Theory Trying to Accomplish?", a: K.J. Arrow i S. Honkapohja (Ed.), Frontiers of Economics, Blackwell, Oxford.
- Aumann, R., (1987): "Game Theory", The New Palgrave. A Dictionary of Economics, vol. 2.
- Axelrod, R., (1981): "The Emergence of Cooperation among Egoists", American Political Science Review, vol. 75.
- Axelrod, R., (1984): La evolución de la cooperación, Alianza Editorial, Madrid, 1986.
- Axelrod, R., (1986): "An Evolutionary Approach to Norms", American Political Science Review, vol. 80.
- Balme, R., (1990): "L'action collective rationnelle dans le paradigme d'Olson", L'Année Sociologique, vol. 40.
- Banks, J. i R. Sundaram, (1990): "Repeated Games, Finite Automata and Complexity", Games and Economic Behavior, vol. 2.
- Barberà, S., (1991): "Alguns models del comportament racional en economia", a R. Marimon i X. Calsamiglia (Eds.), Invitació a la teoria econòmica, Aï Barcelona.
- Barnard, Ch., (1938): The Functions of the Executive, Harvard University Press, Cambridge.
- Barry, B., (1970): Los sociólogos, los economistas y la democracia, Amorrortu, Buenos Aires, 1974.
- Barry, B., (1982): "Methodology versus Ideology. The 'Economic Approach Revisited'", a: E. Ostrom (Ed.), Strategies of Political Inquiry, Sage, Beverly Hills.
- Basu, K., (1987): "Modeling Finitely-Repeated Games with Uncertain Termination", Economic Letters, vol. 23.

- Basu, K., (1990): "On the Non-Existence of a Rationality Definition for Extensive Games", International Journal of Game Theory, vol. 19.
- Baumol, W.J., (1952): Welfare Economics and the Theory of the State, Longmans, Londres.
- Becker, G., (1976): The Economic Approach to Human Behavior, Chicago University Press, Chicago.
- Behan McCullagh, C., (1991): "How Objective Interests Explain Action", Social Science Information, vol. 30.
- Bell, D., (1982): "Regret in Decision-Making under Uncertainty", Operations Research, vol. 20.
- Bell, D., H. Raiffa i A. Twersky, (1988): "Descriptive, Normative and Prescriptive Interactions in Decision-Making", a: D. Bell, H. Raiffa i A. Twersky (Ed.), Decision-Making: Descriptive, Normative and Prescriptive Interactions, Cambridge University Press, Cambridge.
- Bendor, J., (1987): "In Good Times and Bad: Reciprocity in an Uncertain World", American Journal of Political Science, vol. 31.
- Bentham, J., (1968): "An Introduction to the Principles of Morals and Legislation", a: Page, A. (Ed.), Utility Theory. A Book of Readings, Wiley, New York.
- Ben-Porath, E., (1990): "The Complexity of Computing a Best Response Automaton in Repeated Games with Mixed Strategies", Games and Economic Behavior, vol. 2.
- Bernheim, B.D., (1984): "Rationalizable Strategic Behavior", Econometrica, vol. 52.
- Bernoulli, D., (1728): "Exposition of a New Theory on the Measurement of Risk", Econometrica, vol. 22, 1954.

- Bianchi, M., (1990): "The Unsatisfactoriness of Satisficing from Bounded Rationality to Innovative Rationality", Review of Political Economy, vol. 2
- Bianco, W. i R. Bates, (1990): "Cooperation by Design: Leadership, Structure, and Collective Dilemmas", American Political Science Review, vol. 84.
- Bicchieri, C., (1988): "Strategic Behavior and Contrafactuals", Synthese, vol. 76.
- Bierman, H., (1989): "The Allais Paradox: a Framing Perspective", Behavioral Science, vol. 34.
- Binmore, K., (1987): "Modeling Rational Players I.", Economics and Philosophy, vol. 3.
- Binmore, K., (1988): "Modeling Rational Players II.", Economics and Philosophy, vol. 4.
- Binmore, K., (1990): Essays on the Foundations of Game Theory, Basil Blackwell, Cambridge.
- Bliss, C. i B. Nalebuff, (1984): "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good", Journal of Public Economics, vol. 25.
- Blom, R. i M. Kivinen, (1990): "Analytical Marxism and Class Theory", a: S. Clegg (Ed.), Organization Theory and Class Analysis. New Approaches and Issues, De Gruyter, Berlin-Nova York.
- Blumel, W. et al., (1986): "The Theory of Public Goods. A Survey", Journal of Institutional and Theoretical Economics, vol. 142.
- Bomze, I. i B. Pötscher, (1989): Game Theoretical Foundations of Ecological Stability, Springer Verlag, Berlin-Heidelberg.
- Bonacich, Ph., (1987): "Communication Networks and Collective Action", Social Networks, vol. 9.



- Boudon, R., (1970): L'analyse mathématique des faits sociaux, Plon, Paris.
- Boudon, R., (1982): La lógica de lo social, Rialp, Madrid.
- Boudon, R., (1986): "Individualisme et Hollisme dans les Sciences Sociales", a: P. Birnbaum i J. Leca (Ed.), Sur l'Individualisme, FNSP, Paris.
- Boudon, R., (1986): L'Ideologie, Fayard, Paris.
- Boudon, R., (1989): "Subjective Rationality and the Explanation of Social Behavior", Discussion Paper, 89/6, Max-Plank-Institut für Gesellschaftsforschung, Colonia.
- Bowman, J.R., (1982): "The Logic of Capitalist Collective Action", Social Science Information, vol. 21.
- Brandts, J., (1988): "Economia experimental y teoria de juegos: un panorama", Cuadernos Económicos de ICE, vol. 40.
- Brennan, J. i J. Buchanan, (1986): La razón de las normas, Norma Editorial, Madrid.
- Brewer, B. i R. Kramer, (1986): "Choice Behavior in Social Dilemmas: Effects of Social Identity, Group Size, and Decision Draming", Journal of Personality and Social Psychology, vol. 50.
- Brubaker, R., (1984): The Limits of Rationality. Essay on the Social and Moral Thought of Max Weber, Allen and Unwin, Londres.
- Buchanan, J., (1971): "Las bases de la acción colectiva", a: J. Casahuga (Ed.), Teoria económica y democracia, Instituto de Estudios Fiscales, Madrid.
- Bunge, M., (1989): "Game Theory is Not a Useful Tool for the Political Scientist", Epistemologia, vol. 12.
- Bunge, M., (1989): Mente y Sociedad, Alianza Editorial, Madrid.

- Burris, V., (1990): "Classes in Contemporary Capitalist Society: Recent Marxist and Weberian Perspectives", a: S. Clegg (Ed.), Organization Theory and Class Analysis. New Approaches and New Issues, De Gruyter, Berlin-Nova York.
- Caínzos, M.A., (1990): "Explotación, dominación y estructura de clase", Política y Sociedad, vol. 5.
- Camerer, C. i K. Weigelt, (1988): "Experimental Test of a Sequential Equilibrium Reputation Model", Econometrica, vol. 56.
- Carley, K., (1989): "The Value of Cognitive Foundations for Dynamic Social Theory", Journal of Mathematical Sociology, vol. 14.
- Carling, A.H., (1987): "The Schelling Diagram: On Binary Choice with Externalities", Behavioral Sciences, vol. 32.
- Cesco, J.C. i A. Neme, (1989): "A Nash Type Equilibrium in Games without Transitive Preferences", Mathematical Social Sciences, vol. 17.
- Chamberlin, J., (1974): "Provision of Collective Goods as a Function of Group Size", American Political Science Review, vol. 68.
- Chamberlin, J., (1978): "A Diagrammatic Exposition of the Logic of Collective Action", a: Hooker et al. (Ed.), Foundations and Applications of Decision Theory II, Reidel, Dordrecht.
- Clark, P. i J. Wilson, (1961): "Incentive Systems: A Theory of Organizations", Administrative Science Quarterly, vol. 6.
- Coats, A.W., (1983): "John R. Commons as a Historian of Economics: The Quest for the Antecedents of Collective Action", Research in the History of Economic Thought and Methodology, vol. 1.
- Cohen, M., J. March i J. Olsen, (1972): "A Garbage Can Model of Organizational Making", Administrative Science Quarterly, vol. 17.

- Coleman, J., (1987): "Competition and Cooperation", Ethics, vol. 98.
- Coleman, J., (1989): "Simulation Games and The Development of Social Theory", Simulation and Games, vol. 20.
- Coleman, J., (1990): Foundations of Social Theory, Harvard University Press, Cambridge.
- Colman, A. , (1982): Game Theory and Experimental Economics, Pergamon Press, Oxford.
- Colomer, J.M., (1990): El arte de la manipulación política, Anagrama, Barcelona.
- Commons, J.R., (1934): Institutional Economics: Its Place in Politics and Economy, University of Wisconsin Press, Madison, 1959.
- Commons, J.R., (1950): The Economics of Collective Action, University of Wisconsin Press, Madison, 1970.
- Condon, A., (1989): Computational Models of Games, MIT Press, Cambridge.
- Cooper, R., D. Dejong, R. Forsythe i T. Ross, (1990): "Selection Criteria in Coordination Games: Some Experimental Results", American Economic Review, vol. 80.
- Cornes, R. i T. Sandler, (1986): The Theory of Externalities, Public Goods and Club Goods, Cambridge University Press, Cambridge.
- Crawford, V. i H. Haller, (1990): "Learning How to Cooperate: Optimal Play in Repeated Coordination Games", Econometrica, vol. 58.
- Crawford, V., (1990): "Equilibrium without Independence", Journal of Economic Theory, vol. 50.
- Crenson, M., (1987): "The Private Stake in Public Goods: Overcoming the Illogic of Collective Action", Policy Sciences, vol. 20.

- Davidson, D., (1980): Essays on Actions and Events, Clarendon Press, Oxford.
- Domènech, A., (1989): De la Ética a la Política, Crítica, Barcelona.
- Donninger, Ch., (1986): "Is it Always Efficient to be Nice?. A Computer Simulation of Axelrod's Computer Tournament", a: A. Diekmann i P. Mitter (Ed.), Paradoxical Effects of Social Behavior, Springer Verlag, Heidelberg-Viena.
- Downs, A., (1957): Una teoria econòmica de la democràcia, Aguilar, Madrid, 1973.
- Dresher, M., S. Karlin i L.S. Shapley, (1950): "Polynomial Games", Annals of Mathematics Study (Contributions to Theory of Games, I), vol. 24.
- Dunleavy, P., (1988): "Group Identity and Individual Influence: Reconstructing the Theory of Interest Groups", British Journal of Political Science, vol. 8
- Dunleavy, P., (1991): Democracy, Bureaucracy and Public Choice, Harvester-Wheatsheaf, Nova York-Londres.
- Dupy, J.-P., (1989): "Common Knowledge, Common Sense", Theory and Decision, vol. 27.
- Earl, P.E., (1988): "Introduction", a: P.E. Earl (Ed.), Behavioral Economics I, Aldershot, Elgar.
- Einhorn, H. i R. Hogarth, (1988): "Behavioral Decision Theory: Processes of Judgement and Choice", a: D. Bell, H. Raiffa i A. Tversky (Ed.), Decision Making: Descriptive, Normative and Prescriptive Interactions, Cambridge University Press, Cambridge.
- Ellsberg, D., (1954): "Classic and Current Notions of 'Mesurable Utility'", Economic Journal, vol. 64.

- Elster, J. i A. Hylland, (1986): "Introduction", a: J. Elster i A. Hylland (Ed.), Foundations of Social Choice Theory, Cambridge University Press, Cambridge.
- Elster, J., (1976): "Some Conceptual Problems in Political Theory", a: B. Barry (Ed.), Power and Political Theory, Wiley, Londres.
- Elster, J., (1978): Logic and Society, Wiley, Londres.
- Elster, J., (1979): "Anomalies of Rationality: Some Unresolved Problems in the Theory of Rational Behavior", a: L. Lévi-Garboua (Ed.), Sociological Economics, Sage, Londres.
- Elster, J., (1982): "Marxism, Functionalism and Game Theory: The Case for Methodological Individualism", Theory and Society, vol. 11.
- Elster, J., (1983): Explaining Technical Change: A Case Study in the Philosophy of Science, Cambridge University Press, Cambridge.
- Elster, J., (1983): Sour Grapes: Studies in the Subversion of Rationality, Cambridge University Press, Cambridge.
- Elster, J., (1984): Ulysses and the Sirens. Studies in Rationality and Irrationality, Cambridge University Press, Cambridge.
- Elster, J., (1985): Making Sense of Marx, Cambridge University Press, Cambridge.
- Elster, J., (1985): "Rationality, Morality and Collective Action", Ethics, vol. 96.
- Elster, J., (1985): "Sadder but Wiser? Rationality and the Emotions", Social Science Information, vol. 24.
- Elster, J., (1985): "The Nature and Scope of Rational-Choice Explanation", a: E. LePore i B.P. McLaughlin (Ed.), Actions and Events, Basil Blackwell, Londres.

- Elster, J., (1985): "Three Challenges to Class", a: J. Roemer (Ed.), Analytical Marxism, Cambridge University Press, Cambridge.
- Elster, J., (1985): "Weakness of Will and the Free-Rider Problem", Economics and Philosophy, vol. 1.
- Elster, J., (1986): "Introduction", a: J. Elster (Ed.), Rational Choice, Oxford University Press, Oxford.
- Elster, J., (1986): "The Market and the Forum", a: J. Elster i A. Hylland (Ed.), Foundations of Social Choice Theory, Cambridge University Press, Cambridge.
- Elster, J., (1989): Nuts and Bolts for Social Sciences, Cambridge University Press, Cambridge.
- Elster, J., (1989): Solomonic Judgements, Cambridge University Press, Cambridge.
- Elster, J., (1989): The Cement of Society, Cambridge University Press, Cambridge.
- Elster, J., (1990): "Norms of Revenge", Ethics, vol. 100.
- Elster, J., (1990): "Rationality and Social Norms", , Paper prepared for the XII Congress of the ISA, Madrid.
- Elster, J., (1990): "Social Norms and Economic Theory", Journal of Economic Perspectives, vol. 4.
- Erev, I. i A. Rapoport, (1990): "Provision of Steep-Level Public Goods", Journal of Conflict Cooperation, vol. 34.
- Festinger, L., A. Pepitone i T.M. Newcomb, (1952): "Some Consequences of De-Individuation in a Group", Journal of Abnormal and Social Psychology, vol. 47.

- Field, A., (1984): "Microeconomics, Norms and Rationality", Economic Development and Cultural Change, vol. 32.
- Fishburn, P. i I. LaValle, (1988): "The Structure of SSB Utilities for Decision under Uncertainty", Mathematical Social Sciences, vol. 15.
- Fishburn, P., (1968): "Utility Theory", Management Science, vol. 14.
- Fishburn, P., (1972): "On the Foundations of Game Theory: The Case of Non-Archimedean Utilities", International Journal of Game Theory, vol. 1.
- Fishburn, P., (1982): "Nontransitive Measurable Utility", Journal of Mathematical Psychology, vol. 26.
- Fishburn, P., (1988): "Expected Utility: An Anniversary and a New Era", Journal of Risk and Uncertainty, vol. 1.
- Fishburn, P., (1989): "Retrospective on the Utility Theory of von Neumann and Morgenstern", Journal of Risk and Uncertainty, vol. 2.
- Fisher, I., (1918): "Is 'Utility' the Most Suitable Term for the Concept it is Used to Denote?", American Economic Review, vol. 8.
- Fleishman, J., (1988): "The Effects of Decision Framing and Others' Behavior on Cooperation in a Social Dilemma", Journal of Conflict Resolution, vol. 32.
- Flood, M., (1958): "Some Experimental Games", Management Science, vol. 5.
- Ford Runge, C., (1984): "Institutions and the Free Rider: The Assurance Problem in Collective Action", Journal of Politics, vol. 46.
- Fox, J. i M. Guyer, (1977): "Group Size and Others' Strategy in an N-Person Game", Journal of Conflict Resolution, vol. 21.
- Frey, B. i R. Eichenberger, (1989): "Anomalies and Institutions", Journal of Institutional and Theoretical Economics, vol. 145.

- Friedman, J. i R. Rosenthal, (1986): "A Positive Approach to Non-Cooperative Games", Journal of Economic Behavior and Organization, vol. 7.
- Friedman, J., (1990): Game Theory with Applications to Economics, Oxford University Press, Nova York.
- Friedman, M., (1953): Essays in Positive Economics, Chicago University Press, Chicago.
- Friedman, M., (1953): "La metodologia de la economia positiva", a: F. Hahn i M. Hollis (Eds.), Filosofia y teoria econòmica, FCE, Mèxic, 1986.
- Frisch, R., (1926): "A Problem of Pure Economics", a: J.S. Chipman et al. (Eds.), Preferences, Utility and Demand, Harcourt Brace Jovanovich, Nova York, 1971.
- Frohlich, N. i J. Oppenheimer, (1984): "Beyond Economic Man. Altruism, Egalitarianism and Difference Maximizing", Journal of Conflict Resolution, vol. 28.
- Fuentes Quintana, E., (1986): "James Buchanan. Premio Nobel de Economía 1986", Papeles de Economía Española, vol. 31-32.
- Fudenberg, D. i E. Maskin, (1986): "Folk Theorems for Repeated Games with Discounting and Incomplete Information", Econometrica, vol. 54.
- Fudenberg, D. i E. Maskin, (1990): "Evolution and Cooperation in Noisy Repeated Games", American Economic Review. Papers and Proceedings, vol. 80.
- Gellner, E., (1986): "L'animal qui évite les gaffes, ou un faisceau d'hypothèses", a: P. Birnbaum i J. Leca (Ed.), Sur l'Individualisme, FNSP, Paris.
- Georgescu-Roegen, N., (1968): "Utility", International Encyclopedia of the Social Sciences, vol. 16.



- George, D., (1984): "Meta-Preferences: Reconsidering Contemporary Notions of Free-Choice", International Journal of Social Economics, vol. 11.
- Gilbert, M., (1990): "Rationality, Coordination, and Convention", Synthese, vol. 84.
- Gilboa, I. i D. Samet, (1989): "Bounded versus Unbounded Rationality: The Tyranny of the Weak", Games and Economic Behavior, vol. 1.
- Gilboa, I., (1988): "The Complexity of Computing Best Response Automata in Repeated Games", Journal of Economic Theory, vol. 45.
- Giner, S., (1977): "Intenciones humanas y estructuras sociales: Aproximación crítica a la lógica situacional", Cuadernos Económicos del ICE, vol. 3-4.
- Goetze, D. i P. Galderisi, (1989): "Explaining Collective Action with Rational Models", Public Choice, vol. 62.
- Goodin, R., (1988): "Some New Sources of Social Conflict: Transformations of Mixed-Motive Games", British Journal of Sociology, vol. 39.
- Gottinger, H., (1982): "Computational Cost and Bounded Rationality", a: W. Balzer i W. Spohn (Ed.), Philosophy of Economics, Springer Verlag, Berlin-Heidelberg.
- Gómez Rodríguez, A., (1988): "Acción individual, lógica social", Zona Abierta, vol. 48-49.
- Granovetter, M., (1985): "Economic Action and Social Structure: The Problem of Embeddness", American Sociological Review, vol. 91.
- Granovetter, M., (1991): "Modelos de umbral de conducta colectiva", Zona Abierta, vol. 54-55.
- Greenberg, J., (1989): "An Application of the Theory of Social Situations to Repeated Games", Journal of Economic Theory, vol. 49.

- Grether, D., (1980): "Bayes Rules as a Descriptive Model: The Representativeness Heuristic", Quarterly Journal of Economics, vol. 95.
- Groser, M., (1989): "Die Logik organisierter Interessen", Jahrbuch für Neue Politische Ökonomie, vol. 8.
- Güth, W., W. Leiniger i G. Stephan, (1989): "On Supergames and Folk Theorems. A Conceptual Discussion", Diskussionsschriften, Heidelberg.
- Hadari, A., (1987): "What are Preference Explanations? The Interpretative Core of Economic Modeling", Social Science Quarterly, vol. 68.
- Hagen, O., (1979): "Towards a Positive Theory of Preferences under Risk", a: M. Allais i O. Hagen (Ed.), Expected Utility Hypotheses and the Allais Paradox, Reidel, Dordrecht.
- Hamburger, H., (1973): "N-Person Prisoner's Dilemma", Journal of Mathematical Sociology, vol. 3.
- Hamburger, H., (1979): Games as Models of Social Phenomena, Freeman, San Francisco.
- Hamburger, H., M. Guyer i J. Fox, (1975): "Group Size and Cooperation", Journal of Conflict Cooperation, vol. 19.
- Hammerstein, P., (1984): "The Biological Counterpart to Non-Cooperative Game Theory", Institute of Mathematical Economics, Working Paper, Bielefeld.
- Hampton, J., (1987): "Free-Rider Problems in the Production of Collective Goods", Economics and Philosophy, vol. 3.
- Hansen, J.-H., (1985): "The Political Economy of Group Membership", American Political Science Review, vol. 79.
- Hardin, R., (1971): "Collective Action as an Agreeable N-Prisoners' Dilemma", Behavioral Science, vol. 16.

- Hardin, R., (1982): Collective Action, Johns Hopkins University Press, Baltimore.
- Hardin, R., (1984): "Difficulties in the Notion of Economic Rationality", Social Science Information, vol. 23.
- Hardin, R., (1991): "Acting Together, Contributing Together", Rationality and Society, vol. 3.
- Hargreaves Heap, S., (1989): Rationality in Economics, Basil Blackwell, Oxford.
- Harsanyi, J. i R. Selten, (1988): A General Theory of Equilibrium Selection in Games, MIT Press, Cambridge.
- Harsanyi, J., (1967): "Games with Incomplete Information Played by Bayesian Players", Management Science, vol. 13.
- Harsanyi, J., (1977): Rational Behavior and Bargaining Equilibrium Selection in Games and Social Situations, Cambridge University Press, Cambridge.
- Harsanyi, J., (1987): "Bargaining", The New Palgrave. A Dictionary of Economics, vol. 1.
- Hartfiel, G., (1968): Wirtschaftliche und Soziale Rationalität, Ferdinand Enke, Stuttgart.
- Hart, S., (1985): "Nonzero-Sum Two-Person Repeated games with Incomplete Information", Mathematics of Operations Research, vol. 10.
- Hausner, M., (1954): "Multidimensional Utilities", a: R. Thrall et al. (Ed.), Decision Processes, Wiley, Nova York.
- Hayek, F.A., (1967): "Notes on the Evolution of Systems of Rules of Conduct", Studies in Philosophy, Politics and Economics, Routledge, Londres.

- Haywood, D.G., (1954): "Military Decision and Game Theory", Operations Research, vol. 2.
- Heackthorn, D., (1988): "Collective Sanctions and the Creation of Prisoner's Dilemma Norms", American Journal of Sociology, vol. 94.
- Heackthorn, D., (1989): "Cognitive Science, Sociology, and the Theoretic Analysis of Complex Systems", Journal of Mathematical Sociology, vol. 14.
- Head, J., (1962): "Public Goods and Public Police", Public Finance, vol. 17.
- Hechter, M., (1990): "On the Inadequacy of Game Theory for the Solution of Real World Collective Action Problems", a: K.S. Cook i M. Levy (Ed.), The Limits of Rationality, University of Chicago Press, Chicago.
- Heiner, R., (1983): "The Origin of Predictable Behavior", American Economic Review, vol. 73.
- Heiner, R., (1985): "Origin of Predictable Behavior: Further Modeling and Applications", American Economic Review. Papers and Proceedings, vol. 75.
- Heiner, R., (1989): "Imperfect Choice and Self-Stabilizing Rules", Economics and Philosophy, vol. 5.
- Herrero, M.J., (1989): "The Nash Program: Non-Convex Bargaining Problems", Journal of Economic Theory, vol. 49.
- Hesse, W. i U. Schulz, (1986): "Preference Structures by Payoffs in Experimental Games", a: R.W. Scholz (Ed.), Current Issues in West German Decision Research, Peter Lang, Frankfurt.
- Hey, J., (1982): "Search for Rules for Search", Journal of Economic Behavior and Organization, vol. 3.
- Hicks, J., (1956): A Revision of Demand Theory, Clarendon Press, Oxford.

- Hicks, J. i D. Allen, (1934): "A Reconsideration of the Theory of Value", Economica New Series (Econometrica), vol. 1.
- Hindess, B., (1984): "Rational Choice Theory and the Analysis of Political Action", Economics and Society, vol. 13.
- Hindess, B., (1988): Choice, Rationality and Social Theory, Uwin Hyman, Londres.
- Hirschman, A., (1977): Voz, salida y lealtad, FCE, Mèxic.
- Hirschman, A., (1978): Las pasiones y los intereses, FCE, Mèxic.
- Hirschman, A., (1983): Interés público y acción pública, FCE, Mèxic.
- Hirschman, A., (1989): "Doscientos años de retórica reaccionaria. El caso del efecto perverso", El Trimestre Económico, vol. 56.
- Hirschman, A., (1989): "Salida y voz: una esfera de influencia en expansión", Enfoques alternativos sobre la sociedad de mercado, FCE, Mèxic.
- Hirshleifer, J. i E. Rasmusen, (1989): "Cooperation in Repeated Prisoners' Dilemma with Ostracism", Journal of Economic Behavior and Organization, vol. 12.
- Hirshleifer, J. i J.C. Martinez Coll, (1988): "What Strategies Can Support the Evolutionary Emergence of Cooperation?", Journal of Conflict Resolution, vol. 32.
- Hirshleifer, J., (1983): "From Weakest-Link to Best Shot: The Voluntary Provision of Public Goods", Public Choice, vol. 41.
- Hirshleifer J. i J. Riley, (1979): "The Analytics of Uncertainty and Information -An Expository Survey", Journal of Economic Literature, vol. 17.

- Hollis, M. i E. Nell, (1975): Rational Economic Man, Cambridge University Press, Cambridge.
- Hollis, M., (1979): "Rational Man and Social Science", a: R. Harrison (Ed.), Rational Action. Studies in Philosophy and Social Science, Cambridge University Press, Cambridge.
- Holt, Ch., (1986): "Preference Reversals and the Independence Axioms", American Economic Review, vol. 76.
- Homans, G., (1961): Social Behavior: Its Elementary Forms, Harcourt, Brace and World, Nova York.
- Homans, G., (1986): "Fifty Years of Sociology", Annual Review of Sociology, vol. 12.
- Hong, C. i S. Waller, (1986): "Empirical Test of Weighted Utility Theory", Journal of Mathematical Psychology, vol. 29.
- Hovi, J., (1986): "Binary Games as Models of Public Goods Provision", Scandinavian Political Studies, vol. 9.
- Jacquemin, A., (1987): La nueva organización industrial, Vicens Vives, Barcelona, 1989.
- Jenkins, C., (1983): "Resource Mobilization Theory and the Study of Social Movements", Annual Review of Sociology, vol. 9.
- Jensen, N., (1967): "An Introduction to Bernoullian Utility Theory. I. Utility Functions", Swedish Journal of Economics, vol. 69.
- Jensen, N., (1967): "An Introduction to Bernoullian Utility Theory. II. Interpretation, Evaluation and Application: a Critical Survey", Swedish Journal of Economics, vol. 69.

- Johansen, L., (1982): "On the Status of the Nash Type of Noncooperative Equilibrium in Economic Theory", Scandinavian Journal of Economics, vol. 84.
- Johnson, R.L., (1988): "Symbolic Action and the Limits of Strategic Rationality", Political Power and Social Theory, vol. 7.
- Kacprzyk, J. i M. Roubens (Eds.), (1988): Non-Conventional Preference Relations in Decision-Making, Springer Verlag, Berlin-Heidelberg.
- Kadane, J. i P. Larkey, (1982): "Subjective Probability and the Theory of Games", Management Science, vol. 28.
- Kadane, J. i P. Larkey, (1983): "The Confusion of Is and Ought in Game Theoretic Contexts", Management Science, vol. 29.
- Kahan, J. i A. Rapoport, (1984): "Social Psychology and the Theory of Games: A Mixed-Motive Relationship", Representative Research in Social Psychology, vol. 14.
- Kahneman, D. i A. Tversky, (1979): "Prospect Theory: An Analysis of Decision under Risk", Econometrica, vol. 47.
- Kahneman, D., (1988): "Experimental Economics: A Psychological Perspective", a: R. Tietz, W. Albers i R. Selten (Eds.), Bounded Rational Behavior in Experimental Games and Markets, Springer Verlag, Berlin.
- Kalai, E. i W. Stanford, (1988): "Finite Rationality and Interpersonal Complexity in Finite Games", Econometrica, vol. 56.
- Kalay, E. i M. Smorodinsky, (1975): "Other Solutions to Nash's Bargaining Problem", Econometrica, vol. 43.
- Kant, M., (1785): Fundamentación de la metafísica de las costumbres, Espasa-Calpe, Madrid, 1981.

- Kauder, E., (1965): A History of Marginal Utility Theory, Princeton University Press, Princeton.
- Keller, B., (1988): "Olson's 'Logik des kollektiven Handelns'. Entwicklung, Kritik -und eine Alternative", Politische Vierteljahresschrift, vol. 29.
- Kelley, H. i J. Thibaut, (1978): Interpersonal Relations. A Theory of Interdependence, Wiley, Nova York.
- Kerr, N.L., (1989): "Illusions of Efficacy: The Effects of Group Size on Perceived Efficacy in Social Dilemmas", Journal of Experimental Social Psychology, vol. 25.
- Kimura, K., (1989): "Large Groups and the Tendance towards Faillure: A Critique of M. Olson's Model of Collective Action", Journal of Mathematical Sociology, vol. 40.
- Kliemt, H., (1983): "Der Homo Oeconomicus zwischen Ethik und Biologie -Aspekte seiner Geschichte und Wirkungsgeschichte", a: M.J. Holler (Ed.), Homo Oeconomicus I, N. Leudemann Verlag, Munich.
- Kline, M., (1985): Matemáticas. La pérdida de la certidumbre, Siglo XXI, Madrid.
- Knoke, D. i C. Wright-Isak, (1982): "Individual Motives and Organizational Incentives Systems", Research in the Sociology of Organizations, vol. 1.
- Knoke, D., (1988): "Incentives in Collective Action Organizations", American Sociological Review, vol. 53.
- Knoke, D., (1990): Organizing for Collective Action. The Political Economies of Associations, De Gruyer, Nova York.
- Kolm, S.K., (1986): "Le bouddhisme et les hommes économiques", Bulletin du MAUSS, vol. 6.



- Kondo, T., (1990): "Some Notes on Rational Behavior, Normative Behavior, Moral Behavior, and Cooperation", Journal of Conflict Resolution, vol. 34.
- Koslowski, P., (1985): "Economic Principle, Maximizing, and the Co-ordination of Individual in Economics and Philosophy", a: P. Koslowski (Ed.), Economics and Philosophy, J.C.B. Mohr, Tubinga.
- Kostreva, M.M., (1989): "Nonconvexity in Noncooperative Game Theory", International Journal of Game Theory, vol. 18.
- Kreps, D. i R. Wilson, (1982): "Sequential equilibria", Econometrica, vol. 50.
- Kreps, D., (1987): "Nash Equilibrium", The New Palgrave. A Dictionary of Economics, vol. 2.
- Kreps, D., P. Milgrom, P. Roberts i R. Wilson, (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma", Journal of Economic Theory, vol. 27.
- Laitin, D., (1988): "Language Games", Comparative Politics, vol. 22.
- Laitin, D., (1989): "Linguistic Revival: Politics and Culture in Catalonia", Comparative Studies in Society and History, vol. 31.
- Laitin, D., (1990): "Language and States", Instituto Juan March de Estudios e Investigaciones, Working Paper 90/3, Madrid.
- Lancaster, K., (1971): Consumer Demand: A New Approach, Columbia University Press, Nova York.
- Lange, P. von, W. Liebrand i D. Kuhlman, (1990): "Causal Attribution of Choice Behavior in Three N-Person Prisoner's Dilemma", Journal of Experimental Social Psychology, vol. 26.
- Lapeyronnie, D., (1988): "Mouvements sociaux et action politique. existe-t-il une théorie de la mobilisation des ressources?", Revue Française de Sociologie, vol. 29.

- Laver, M., (1980): "Political Solutions to the Collective Action Problem", Political Studies, vol. 28.
- Leibenstein, H., (1984): "On the Economics of Conventions and Institutions: An Exploratory Essay", Zeitschrift für die gesamte Staatswissenschaft, vol. 10
- Leist, A., (1989): "Kollektive Güter und individuelle Verantwortung", Analyse und Kritik, vol. 11.
- Leontief, W., (1989): "Input-Output Data Base for Analysis of Technological Change", Economic Systems Research, vol. 1.
- Levy, I., (1974): "On Indeterminate Probabilities", Journal of Philosophy, vol. 71.
- Lewis, C., (1969): Convention: A Philosophical Study, Cambridge University Press, Cambridge.
- Liebrand, W., (1984): "The Effect of Social Motives, Communication and Group Size on Behavior in an N-Person Multi-Stage Mixed-Motive Game", European Journal of Social Psychology, vol. 14.
- Lindenberg, S., (1985): "Rational Choice and Sociological Theory: New Pressures on Economics as a Social Science", Journal of Institutional and Theoretical Economics, vol. 141.
- Lipman, B.L., (1986): "Cooperation among Egoists in Prisoner's Dilemma and Chicken Games", Public Choice, vol. 51.
- Lipnowski, I. i S. Maital, (1983): "Voluntary Provision of a Pure Public Good as the Game of 'Chicken'", Journal of Public Economics, vol. 20.
- Luce, D. i H. Raiffa, (1957): Games and Decisions, Wiley, New York.
- MacCarthy, J. i M. Zald, (1977): "Resource Mobilization and Social Movements: A Partial Theory", American Journal of Sociology, vol. 82.

- MacCrimmon, K. i D. Messick, (1976): "A Framework for Social Motives", Behavioral Science, vol. 21.
- Machina, M., (1981): "'Rational' Decision Making versus 'Rational' Decision Modelling?", Journal of Mathematical Psychology, vol. 24.
- Machina, M., (1982): "'Expected Utility' Analysis without the Independence Axiom", Econometrica, vol. 50.
- Machina, M., (1983): "Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom", a: B. Stigum i F. Wenstoeop (Ed.), Foundations of Utility and Risk Theory with Applications, Reidel, Dordrecht.
- Machina, M., (1987): "Choice Under Uncertainty: Problems Solved and Unsolved", Economic Perspectives, vol. 1.
- Machina, M., (1989): "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty", Journal of Economic Literature, vol. 27.
- Machlup, F., (1967): "L'Homo Oeconomicus et ses collegues", a: E.M. Claassen (Ed.), Les Fondements Philosophiques des Systèmes Economiques, Payot, Paris.
- Macy, M., (1989): "Walking Out of Social Traps", Rationality and Society, vol. 1.
- Manin, B., (1987): "On Legitimacy and Political Deliberation", Political Theory, vol. 15.
- March, J., (1978): "Bounded Rationality, Ambiguity and the Engineering of Choice", Bell Journal of Economics, vol. 9.
- Margolis, H., (1981): "A New Model of Rational Choice", Ethics, vol. 91.
- Margolis, H., (1982): Selfishness, Altruism and Rationality. A Theory of Social Choice, Cambridge University Press, Cambridge.

- Marimon, R., (1988): "Introducción a los juegos dinámicos", Cuadernos Económicos de ICE, vol. 40.
- Marimon, R. i X. Calsamiglia (Eds.) (1991): Invitació a la teoria econòmica, Ariel Barcelona.
- Martinez Coll J.C. i J. Hirshleifer, (1991): "Selección y mutación en juegos evolutivos", Ms..
- Martinez Coll, J.C. i J. Hirshleifer, (1991): "The Limits of Reciprocity: Solution Concepts and Reactive Strategies in Evolutionary Equilibrium Models", Rationality and Society, vol. 3.
- Martin, J., G. Kleindorfer i W. Brashers, (1987): "The Theory of Bounded Rationality and the Problem of Legitimation", Journal for the Theory of Social Behavior, vol. 17.
- Marwell, G. i P. Oliver, (1984): "Collective Action and Social Movements", Research in Social Movements, conflict and Change, vol. 7.
- Marwell, G. i R. Ames, (1979): "Experiments on the Provision of Public Goods I. Resources, Interest, Group Size, and the Free-Rider Problem", American Journal of Sociology, vol. 84.
- Marwell, G. i R. Ames, (1981): "Economists Free Ride, does Anyone Else?: Experiments on the Provision of Public Goods IV.", Journal of Public Economics, vol. 15.
- Marwell, G., P. Oliver i P. Prahl, (1988): "Social Networks and Collective Action. A Theory of the Critical Mass III.", American Journal of Sociology, vol. 94.
- Mas Colell, A., (1974): "An equilibrium Existence Theorem for a General Model without Complete or Transitive Preferences", Journal of Mathematical Economics, vol. 1.

- Mas Colell, A., (1977): "On Revealed Preference Analysis", Review of Economic Studies, vol. 44.
- Mas Colell, A., (1988): "Algunos comentarios sobre la teoria cooperativa de los juegos", Cuadernos Económicos de ICE, vol. 40.
- Maynard Smith, J., (1982): Evolution and the Theory of Games, Cambridge University Press, Cambridge.
- Maynard Smith, J., (1986): "Evolutionary Game Theory", a: D. Farner et al. (Ed.), Evolution, Games and Learning, North-Holland, Amsterdam.
- McCaleb, T. i R. Wagner, (1985): "The Experimental Search for Free Riders: Some Reflections and Observations", Public Choice, vol. 47.
- McClendon, J.F., (1986): "Existence of Solutions of Games with Some Non-Convexity", International Journal of Game Theory, vol. 15.
- McClennen, E., (1983): "Sure-Thing Doubts", a: B. Stigum i F. Wenstoep (Ed.), Foundations of Utility and Risk Theory with Applications, Reidel, Dordrecht.
- McCormick, K., (1989): "An Essay on the Assumption of Rational Behavior in Economics", Review of Social Economy, vol. 47.
- McDaniel, W. i F. Sistrunk, (1991): "Management Dilemmas and Decisions", Journal of Conflict Resolution, vol. 35.
- Meiksins, E., (1989): "Rational Choice Marxism: Is the Game Worth the Candle?", New Left Review, vol. 177.
- Messick, D. i C. McClintock, (1968): "Motivational Bases of Choice in Experimental Games", Journal of Experimental Social Psychology, vol. 4.
- Messick, D. i W. Thorngate, (1967): "Relative Maximization in Experimental Games", Journal of Experimental Social Psychology, vol. 3.

- Miller, G.A., (1956): "The Magical Number Seven, Plus or Minus Two. Some Limits on Our Capacity of Processing Information", Psychological Review, vol. 63.
- Mill, J., (1828): "Essay on Government", a: J. Lively i J. Rees (Ed.), Utilitarian Logic and Politics, Clarendon Press, Oxford, 1978.
- Mironescu, C., (1978): "Individual and Collective Interest in Game Theory", a: P. Birnbaum (Ed.), Democracy, Consensus and Social Contract, Sage, Londres.
- Mirowski, Ph., (1986): "Institutions as Solutions Concepts in a Game Theory Context", a: Ph. Mirowski (Ed.), The Reconstruction of Economic Theory, Kluwer, Boston.
- Mjoset, L., (1985): "Review Essay. The Limits of Neoclassical Institutionalism", Journal of Peace Studies, vol. 22.
- Moe, T., (1979): "On the Scientific Status of Rational Models", American Journal of Political Science, vol. 23.
- Moe, T., (1980): The Organization of Interests, University of Chicago Press, Chicago.
- Moguin, Ph., (1986): "Simon, Stigler et les théories de la rationalité limitée", Social Science Information, vol. 25.
- Moore, M. i M. Chapman Moore, (1989): "Adaptative Learning, Adaptative Utility, and Rational Behavior in a Repeated Prisoner's Dilemma", Journal of Risk and Uncertainty, vol. 2.
- Morgan, J., (1978): "Multiple Motives, Group Decisions, Uncertainty, Ignorance, and Confusion: A Realistic Economics of the Consumer requires Some Psychology", American Economic Review. Papers and Proceedings, vol. 68.

- Morgenstern, O., (1968): "Game Theory. Theoretical Aspects", International Encyclopaedia of Social Sciences, vol. 6.
- Morgenstern, O., (1976): "The Collaboration Between O. Morgenstern and J. von Neumann on the Theory of Games", Journal of Economic Literature, vol. 14.
- Moulin, E., (1986): Game Theory in the Social Sciences, Nova York University Press, Nova York.
- Mueller, D., (1979): Elección pública, Alianza Editorial, Madrid, 1984.
- Muhsan, H.V., (1977): "An Algebraic Theory of the Commons", a: G. Hardin i J. Baden (Ed.), Managing the Commons, Freeman and Co., San Francisco.
- Myerson, R., (1978): "Refinements of the Nash Equilibrium Concept", International Journal of Game Theory, vol. 7.
- Myerson, R., (1984): "Two-Person Bargaining Problems with Incomplete Information", Econometrica, vol. 52.
- Myerson, R., (1985): "Negotiation in Games: A Theoretical Overview", a: L. Hurwicz, D. Schmeidler i H. Sonnenschein (Ed.), Social Goals and Social Organization, Cambridge University Press, Cambridge.
- Nagel, E., (1963): "Assumptions in Economic Theory", American Economic Review. Papers and Proceedings, vol. 53.
- Nakamura, Y., (1990): "Bilinear Utility and a Threshold Structure for Nontransitive Preferences", Mathematical Social Sciences, vol. 19.
- Nash, J., (1951): "Non-Cooperative Games", Annals of Mathematics, vol. 54.
- Nash, J., (1953): "Two person Cooperative Games", Econometrica, vol. 21.
- Neumann, J. von, (1928): "Zur Theorie der Gessellschaftspiele", Mathematische Annalen, vol. 100.

- Neumann, J. von, i O. Morgenstern, (1944): Theory of Games and Economic Behaviour, Princeton University Press, Princeton, 1953.
- Neyman, A., (1985): "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma", Economic Letters, vol. 19.
- Offe, C. i H. Wiesensthal, (1980): "Two Logics of Collective Action", Political Power and Social Theory, vol. 1.
- Offe, C. i U.K. Preuss, (1990): "Democratic Institutions and Moral Resources", , Arbeitspapier 5/90, Zentrum für Sozialpolitik, Bremen.
- Offe, C., (1988): "Endogeneous Preferences and Institutional Choice", Ms., Bielefeld.
- Oliver, P. i G. Marwell, (1988): "The Paradox of Group Size in Collective Action: A Theory of the Critical Mass II.", American Sociological Review, vol. 53.
- Oliver, P., (1980): "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations", American Journal of Sociology, vol. 85.
- Oliver, P., (1984): "Rewards and Punishments as Selective Incentives. An Apex Game", Journal of Conflict Resolution, vol. 28.
- Oliver, P., (1989): "Bringing the Crowd Back In: The Non-Organizational Elements of Social Movements", Research in Social Movements, Conflict and Change, vol. 11.
- Oliver, P., G. Marwell i R. Teixeira, (1985): "A Theory of the Critical Mass I. Interdependence, Group Heterogeneity, and the Production of Collective Action", American Journal of Sociology, vol. 91.
- Olson, M. i R. Zeckhauser, (1966): "An Economic Theory of Alliances", Review of Economic Studies, vol. 48.



- Olson, M. i R. Zeckhauser, (1970): "The Efficient Production of External Economies", American Economic Review, vol. 60.
- Olson, M., (1965): The Logic of Collective Action, Harvard, Cambridge, 1971.
- Olson, M., (1982): The Rise and Decline of Nations, Yale University Press, New Haven.
- Olson, M., (1984): "Microeconomic Incentives and Macroeconomic Decline", Weltwirtschaftliches Archiv, vol. 120.
- Olson, M., (1985): "Space, Organization and Agriculture", American Journal of Agricultural Economics, vol. 67.
- Olson, M., (1987): "Collective Action", The New Palgrave. A Dictionary of Economics, vol. 1.
- Olzak, S., (1989): "Analysis of Events in the Study of Collective Action", Annual Review of Sociology, vol. 15.
- Opp, K.-D., (1985): "Sociology and Economic Man", Journal of Institutional and Theoretical Economics, vol. 141.
- Opp, K.-D., (1986): "Soft-Incentives and Collective Action: Participation in the Anti-Nuclear Movement", British Journal of Political Science, vol. 16.
- Opp, K.-D., (1986): "The Evolution of a Prisoner's Dilemma in the Market", a: A. Diekmann i P. Mitter (Ed.), Paradoxical Effects of Social Behavior, Springer Verlag, Heidelberg-Viena.
- Opp, K.-D., (1989): "The Evolutionary Emergence of Norms", British Journal of Social Psychology, vol. 21.
- Orbell, J., P. Schwartz-Shea i R. Simmons, (1984): "Do Cooperators Exit More Readily than Defectors?", American Political Science Review, vol. 78.

- Ordeshook, P., (1986): Game Theory and Political Theory. An Introduction, Cambridge University Press, Cambridge.
- Ostrom, E., (1986): "An Agenda for the Study of Institutions", Public Choice, vol. 48.
- Ostrom, E., (1990): Governing the Commons: The Evolution of Institutions for Collective Action, Cambridge University Press, Nova York.
- Ostrom, V. i E. Ostrom, (1971): "Public Choice: A Diferent Approach to the Study of Public Administration", Public Administration Review, vol. 31.
- Ouchi, W., (1980): Teoria Z, Orbis, Barcelona, 1985.
- Ovejero, F., (1989): La quimera fértil. La posibilidad de una teoria de la historia, Ms., Barcelona.
- Ovejero, F., (1989): "Diez años de marxismo analítico", Arbor , núm. 526.
- Owen, G., (1982): Game Theory, Academic Press, Nova York.
- Page, A. (ed.), (1968): Utility Theory. A Book of Readings, Wiley, Nova York.
- Parfit, D., (1984): Reasons and Persons, Clarenton Press, Oxford.
- Parker, G.A. i P. Hammerstein, (1985): "Game Theory and Animal Behavior", a: P.J. Greenwood et al. (Ed.), Evolution. Essays in Honour of John Maynard Smith, Cambridge University Press, Cambridge.
- Pasinetti, L., (1985): Cambio estructural y crecimiento económico, Pirámide, Madrid.
- Patzak, M., (1984): "Rationalmodell und analitische Handlungstheorie", a: M. Holler (Ed.), Homo Oeconomicus III, N. Leudemann Verlag, Munich.
- Payne, J., (1982): "Contingent Decision Behavior", Psychological Butletin, vol. 92.

- Peleg, B., (1966): "The Independence of Game Theory of Utility Theory", Bulletin of the American Mathematical Society, vol. 72.
- Petit, Ph., (1990): "Virtus Normativa: Rational Choice Perspectives", Ethics, vol. 100.
- Pizzorno, A., (1978): "Political Exchange and Collective Identity in Industrial Conflict", a: C. Crouch i A. Pizzorno (Ed.), The Resurgence of Industrial Conflict in Western Europe since 1968, MacMillan, Londres.
- Pizzorno, A., (1989): "Algun otro tipo de alteridad: Una crítica a las teorías de la decisión racional", Sistema, vol. 88.
- Plon, M., (1974): "On the Meaning of the Notion of Conflict and its Study in Social Psychology", European Journal of Social Psychology, vol. 4.
- Pommerehne, W. i F. Scheider, (1980): "Wie steth's mit dem Trittbrettfahren? -Eine experimentelle Untersuchung", Zeitschrift für die gesamte Staatswissenschaft, vol. 136.
- Ponsati, C., (1988): "Juegos de negociación", Cuadernos Económicos de ICE, vol. 40.
- Popkin, S., (1979): The Rational Peasant, University of California Press, Berkeley.
- Popper, K., (1967): "La rationalité et le statut du principe de rationalité", a: E.M. Claassen (Ed.), Les Fondements philosophiques des systèmes économiques, Payot, Paris.
- Przeworski, A., (1985): Capitalism and Social Democracy, Cambridge University Press, Cambridge.
- Przeworski, A., (1985): "Marxism and Rational Choice", Politics and Society, vol. 14.
- Putnam, H., (1988): Representation and Reality, MIT Press, Cambridge.

- Quartone, G. i A. Tversky, (1988): "Contrasting Rational and Psychological Analyses of Political Choice", American Political Science Review, vol. 82.
- Rabbie, J., J. Schot i L. Visser, (1989): "Social Identity Theory: A Conceptual and Empirical Critique from the Perspective of a Behavioral Interaction Model", European Journal of Social Psychology, vol. 19.
- Radner, R., (1975): "Satisficing", Journal of Mathematical Economics, vol. 2.
- Raiffa, H., (1982): The Art and Science of Negotiation, Belknap Press, Cambridge.
- Ramsey, F.P., (1931): "Truth and Probability", a: R.B. Braithwaite (Ed.), The Foundations of Mathematics and Other Essays, Humanities Press, Nova York, 1950.
- Rapoport, A. i C. Orwant, (1962): "Experimental Games: A Review", Behavioral Science, vol. 7.
- Rapoport, A., (1985): "Applications of Game-Theoretic Concepts in Biology", Butlletin of Mathematical Biology, vol. 47.
- Rapoport, A., (1987): "Prisioner's Dilemma", The New Palgrave. A Dictionary of Economics, vol. 3.
- Rapoport, A., (1988): "Experiments with N-Person Social Traps", Journal of Conflict Resolution, vol. 32.
- Rasmusen, E., (1989): Games and Information, Basil Blackwell, Oxford.
- Raub, W. i Th. Voss, (1986): "Conditions for Cooperation in Problematic Social Situations", a: A. Diekmann i P. Mitter (Ed.), Paradoxical Effects of Social Behavior, Springer Verlag, Heidelberg-Viena.
- Raub, W., (1988): "An Analysis of the Finitely Repeated Prisioner's Dilemma", European Journal of Political Economy, vol. 4.

- Rawls, A., (1972): A Theory of Justice, Oxford University Press, Oxford.
- Reisman, D., (1990): Theories of Collective Action, MacMillan, Londres.
- Ricardo, D., (1818): Els principis d'economia política i tributació, Edicions 62, Barcelona, 1984.
- Ricard, J., (1988): "Juegos con información incompleta", Cuadernos Económicos de ICE, vol. 40.
- Ricard, J., (1988): "Una introducción a la teoría de juegos", Cuadernos Económicos de ICE, vol. 40.
- Roemer, J. (Ed.), (1985): Analytical Marxism, Cambridge University Press, Cambridge.
- Rosenberg, A., (1980): "Obstacles to the Nomological Connection of Reasons and Actions", Philosophy of Social Sciences, vol. 10.
- Rosenberg, A., (1983): "If Economics isn't a Science, What is It?", The Philosophical Forum, vol. 14.
- Rosenthal, R.W., (1989): "A Bounded-Rationality Approach to the Study of Noncooperative Games", International Journal of Game Theory, vol. 18.
- Rothenberg, L., (1988): "Organizational Maintenance and the Retention Decision in Groups", American Political Science Review, vol. 82.
- Roth, A. i M. Malouf, (1979): "Game-Theoretic Models and the Role of Information in Bargaining", Psychological Review, vol. 86.
- Roth, A. (Ed.), (1985): Game-Theoretic Models of Bargaining, Cambridge University Press, Cambridge.
- Rouquette, M.L., (1975): Les Rumeurs, PUF, Paris.

- Rubinstein, A., (1986): "Finite Automata Play the Repeated Prisoner's Dilemma", Journal of Economic Theory, vol. 39.
- Ryll, A., (1989): "Die Spieltheorie als instrument der Gesellschaftsforschung", Max-Plank-Institut für Gesellschaftsforschung, Discussion Paper 89/10, Colonia.
- Samuelson, P., (1948): "Consumption Theory in Terms of Revealed Preference", Economica, vol. 15.
- Samuelson, P., (1955): "The Pure Theory of Public Expenditure", Review of Economics and Statistics, vol. 37.
- Samuleson, P., (1955): "Diagrammatic Exposition of a Theory of Public Expenditure", Review of Economics and Statistics, vol. 37.
- Savage, L., (1954): Foundations of Statistics, Wiley, New York.
- Scharpf, F., (1989): "Decision Rules, Decision Styles and Policy Choices", Journal of Theoretical Politics, vol. 1.
- Scharpf, F., (1989): "Games Real Actors Could Play: The Problem of Complete Information", Max-Plank-Institut für Gesellschaftsforschung, Discussion Paper 89/9, Colonia.
- Scharpf, F., (1990): "Games Real Actors Could Play. The Problem of Mutual Predictability", Rationality and Society, vol. 2.
- Scharpf, F., (1990): "Games real Actors Could Play: The Problem of Connectedness",,, Max-Plank-Institut für Gesellschaftsforschung, Discussion Paper 90/8, Colonia.
- Scharpf, F., (1991): "Political Institutions, Decision Styles and Policy Choices", a: R. Czada i A. Windhoff-Héritier (Ed.), Political Choice. Institutions, Rules and the Limits of Rationality, Campus Verlag, Frankfurt.

- Schelling, T., (1960): The Strategy of Conflict, Harvard University Press, Cambridge.
- Schelling, T., (1961): "Experimental Games and Bargaining Theory", World Politics, vol. 14.
- Schelling, T., (1968): "Some Thoughts on the Relevance of Game Theory to the Analysis of Ethical Systems", Journal of Conflict Resolution, vol. 12.
- Schelling, T., (1973): "Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities", Journal of Conflict Resolution, vol. 17.
- Schelling, T., (1978): Micromotives and Macrobehavior, Norton and Co., Nova York.
- Schelling, T., (1984): "Self-Command in Practice, in Policy and in a Theory of Rational Choice", American Economic Review, vol. 74.
- Schelling, T., (1967): "What is Game Theory?", a: T. Schelling (Ed.), Choice and Consequence, Harvard University Press, Cambridge, 1984.
- Schick, F., (1984): Having Reasons: An Essay on Rationality and Sociability, Princeton University Press, Princeton.
- Schoemaker, P., (1982): "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations", Journal of Economic Literature, vol. 20.
- Schotter, A., (1981): The Economic Theory of Social Institutions, Cambridge University Press, Cambridge.
- Schotter, A., (1986): "The Evolution of Rules", a: R. Langlois (Ed.), Economics as a Process, Cambridge University Press, Cambridge.
- Schulz, U. i T. May, (1989): "The Recording of Social Orientations with Ranking and Pair Comparison Procedures", European Journal of Social Psychology, vol. 19.

- Scott, R., (1972): "Avarice, Altruism and Second Party Preferences", Quarterly Journal of Economics, vol. 81.
- Seidendeld, T., (1988): "Decision Theory without 'Independence' or without 'Ordering'", Economics and Philosophy, vol. 4.
- Selten, R. i R. Stoecker, (1986): "End Behavior in Sequence of Finite Prisoner's Dilemma Supergames. A Learning Theory Approach", Journal of Economic Behavior and Organization, vol. 7.
- Selten, R., (1985): "Comment", a: K.J. Arrow i S. Honkapohja (Ed.), Frontiers of Economics, Blackwell, Oxford.
- Sen, A., (1969): "A Game-Theoretic Analysis of Theories of Collectivism in Allocation", a: T. Majumdar (Ed.), Growth and Choice, Oxford University Press, Londres.
- Sen, A., (1973): "Behaviour and the Concept of Preference", Economica, vol. 40.
- Sen, A., (1977): "Rational Fools. A Critique of the Behavioural Foundations of Economic Theory", Phylosophy and Public Affairs, vol. 6.
- Sen, A., (1982): Choice, Welfare and Measurement, Basil Blackwell, Oxford.
- Sen, A., (1985): Commodities and Capabilities, North-Holland, Amsterdam.
- Sen, A., (1987): "Rational Behavior", The New Palgrave. A Dicctionary of Economics, vol. 4.
- Sen, A., (1989): "Economic Methodology: Heterogeneity and Relevance", Social Research, vol. 56.
- Shapley, L.S. i M. Shubik, (1954): "A Method for Evaluating the Distribution of Power in a Committee System", American Political Science Review, vol. 48.
- Shubik, M., (1982): Game Theory in the Social Sciences, MIT Press, Cambridge.



- Shubik, M., (1984): A Game Theoretical Approach to Political Economy, MIT Press, Cambridge.
- Shubik, M., (1985): "Comment", a: K.J. Arrow i S. Honkapohja (Eds.), Frontiers of Economics, Blackwell, Oxford.
- Shubik, M., (1986): "Games with Perceptive Commanders but with Indoctrinated or Less Perceptive Subordinates", a: A. Diekmann i P. Mitter (Eds.), Paradoxical Effects of Social Behavior, Physica Verlag, Heidelberg.
- Silver, Ch., (1984): "Utilitarian Participation", Social Science Information, vol. 23.
- Simon, H., (1957): Models of Man, Wiley, Nova York.
- Simon, H., (1976): "From Substantive to Procedural Rationality", a: J. Latsis (Ed.), Method and Appraisal in Economics, Cambridge University Press, Cambridge.
- Simon, H., (1978): "Rationality as Process and as Product of Thought", American Economic Review, vol. 68.
- Simon, H., (1983): Reason in Human Affairs, Stanford University Press, Stanford.
- Simon, H., (1987): "Bounded Rationality", The New Palgrave. A Dictionary of Economics, vol. 1.
- Simon, H., (1987): "Satisficing", The New Palgrave. A Dictionary of Economics, vol. 3.
- Simon, H., (1990): "Invariants of Human Behavior", Annual Review of Psychology, vol. 41.
- Simon, L. i W. Zame, (1990): "Discontinuous Games and Endogenous Sharing Rules", Econometrica, vol. 58.

- Skala, H.J., (1974): "Nonstandart Utilities and the Foundation of Game Theory", International Journal of Game Theory, vol. 3.
- Skala, H.J., (1975): Non-Archimedean Utility Theory, Reidel, Dordrecht.
- Slovic, P., B. Fischhoff i S. Lichtenstein, (1983): "Preference Reversals: A Broader Perspective", American Economic Review, vol. 73.
- Snidal, D., (1985): "Coordination versus Prisoner's Dilemma: Implications for International Cooperation and Regimes", American Political Science Review, vol. 79.
- Sobel, J., (1985): "Utility Maximizers in Iterated Prisoner's Dilemmas", a: R. Campbell i L. Sowden (Eds.), Paradoxes of Rationality and Cooperation, University of British Columbia Press, Vancouver.
- Sonnenschein, H., (1971): "Demand Theory without Transitive Preferences, with Applications to the Theory of Competitive Equilibrium", a: J. Chipman et al. (Ed.), Preferences, utility and Demand, Harcourt Brace Jovanovich, Nova York.
- Sorin, S., (1986): "On Repeated Games of Complete Information", Mathematics of Operations Research, vol. 11.
- Sraffa, P., (1960): Producció de mercaderies per mitjà de mercaderies, Edicions 62, Barcelona, 1985.
- Srivastava, S., (1990): "Informational Requirements and Strategic Complexity in Repeated Games", Games and Economic Behavior, vol. 2.
- Stigler, G. i G. Becker, (1977): "De Gustibus non est Disputandum", American Economic Review, vol. 67.
- Stigler, G., (1950): "The Development of Utility Theory", Journal of Political Economy, vol. 58.

- Stigler, G., (1974): "Free Riders and Collective Action: An Appendix to Theories of Economic Regulation", Bell Journal of Economics, vol. 5.
- Stigler, G., (1961): "The Economics of Information", Journal of Political Economy, vol. 69.
- Streeck, W., (1991): "Interest Diversity and Organizing Capacity: Two Class Logics of Collective Action?", a: R. Czada i A. Windhoff-Héritier (Ed.), Political Choice. Institutions, Rules and the Limits of Rationality, Campus Verlag, Frankfurt.
- Stroebe, W. i B. Frey, (1982): "Self-Interest and Collective Action: The Economics and Psychology of Public Goods", British Journal of Social Psychology, vol. 21.
- Strotz, R., (1953): "Cardinal Utility", American Economic Review, vol. 43.
- Sugden, R., (1982): "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty", Economic Journal, vol. 92.
- Sugden, R., (1985): "Why be Consistent? A Critical Analysis of Consistency requirements in Choice Theory", Economica, vol. 52.
- Sutton, J., (1986): "Non-Cooperative Bargaining Theory: An Introduction", Review of Economic Studies, vol. 53.
- Sweeney, J., (1974): "Altruism, the Free Rider Problem and Group Size", Theory and Decision, vol. 4.
- Taylor, M. i H. Ward, (1982): "Chickens, Whales, and Lumpy Goods: Alternative Models of Public-Goods Provision", Political Studies, vol. 30.
- Taylor, M., (1987): The Possibility of Cooperation, Cambridge University Press, Cambridge.
- Taylor, M., (1989): "Structure, Culture and Action in the Explanation of Social Change", Politics and Society, vol. 17.

- Taylor, M., (1990): "Cooperation and Rationality: Notes on the Collective Action Problem and Its Solutions", a: K.S. Cook i M. Levi (Eds.), The Limits of Rationality, University of Chicago Press, Chicago.
- Thaler, R., (1987): "The Psychology of Choice and the Assumptions of Economics", a: R. Hogarth i M. Reder (Ed.), Rational Choice, University of Chicago Press, Chicago.
- Thompson, G., (1987): "John von Neumann", The New Palgrave. A Dictionary of Economics, vol. 3.
- Tietzel, M., (1990): "Virtue, Vice and Dr. Pangloss", ORDO, vol. 41.
- Tietz, R., (1988): "Experimental Economics: Ways to Model Bounded Rational Bargaining Behavior", a: R. Tietz, W. Albers i R. Selten (Ed.), Bounded Rational Behavior in Experimental Games and Markets, Springer Verlag, Berlin-Heidelberg.
- Tilly, Ch., (1978): Fram Mobilization to Revolution, Addison-Wesley, Reading.
- Tsebelis, G., (1990): "Are Sanctions Effective? A Game-Theoretic Analysis", Journal of Conflict Resolution, vol. 34.
- Turner, R., (1981): "Collective Behavior and Resource Mobilization as Approaches to Social Movements: Issues and Continuities", Research in Social Movements, Conflict and Change, vol. 4.
- Tversky, A. i D. Kahneman, (1987): "Rational Choice and the Framing of Decisions", a: R. Hogarth i M. Reder (Ed.), Rational Choice, University of Chicago Press, Chicago.
- Tversky, A. i D. Kahneman, (1988): "Rational Choice and the Framing of Decisions", a: D. Bell, H. Raiffa i A. Tversky (Ed.), Decision-Making, Reidel, Dordrecht.
- Tversky, A., (1969): "Intransitivity of Preferences", Psychological Review, vol. 76.

- Tversky, A., (1975): "A Critique of Expected Utility Theory: Descriptive and Normative Considerations", Erkenntnis, vol. 9.
- Tversky, A., P. Slovic i D. Kahneman, (1990): "The Causes of Preference Reversal", American Economic Review, vol. 80.
- Uhlander, C., (1989): "Relational Goods' and Participation: Incorporating Sociability into a Theory of Rational Action", Public Choice, vol. 62.
- Ullmann-Marglit, E., (1977): The Emergence of Norms, Clarenton Press, Oxford.
- Van de Kragt, A., J. Orbell i R. Dawes, (1983): "The Minimal Contribution Set as a Solution to Public Good Problem", American Political Science Review, vol. 77.
- Vanberg, V., (1988): "Rules and Choice in Economics and Sociology", Jahrbuch für Neue Politische Ökonomie, vol. 7.
- Varian, H., (1986): Análisis microeconómico, Editorial Bosch, Barcelona.
- Varian, H., (1987): Microeconomia intermedia, Editorial Bosch, Barcelona.
- Vega, F., (1988): "Amenazas increíbles, percepciones insostenibles: refinamientos del equilibrio de Nash en juegos dinámicos", Cuadernos Económicos de ICE, vol. 40.
- Voss, T., (1983): "Rationale Akteure in der soziologischen Analyse", a: M.J. Holler (Ed.), Homo Oeconomicus I, N. Leudemann Verlag, Munich.
- Wallmeier, H.-M., (1986): "Games and Incomplete Information -A Survey", Institute of Mathematical Economics, Working Paper , Bielefeld.
- Wang, M., K. Hipel i N. Fraser, (1988): "Modeling Misperceptions in Games", Behavioral Science, vol. 33.

- Ward, H., (1987): "The Risks of a Reputation for Toughness: Strategy in Public Goods Provision Problems Modeled by Chicken Supergames", British Journal of Political Science, vol. 17.
- Weber, M., (1927): Gesammelte Aufsätze zur Wissenschaftslehre, J.C.B. Mohr, Tubinga, 1967.
- Weber, M., (1968): Methodologische Schriften, S. Fischer Verlag, Frankfurt.
- Weise, P., (1989): "Homo Sociologicus und Homo Oeconomicus. Die Schreckenmänner der Sozialwissenschaften", Zeitschrift für Soziologie, vol. 18.
- Weizsäcker, C. von, (1971): "Notes on Endogenous Change of Tastes", Journal of Economic Theory, vol. 3.
- Weldes, J., (1989): "Marxism and Methodological Individualism", Theory and Society, vol. 18.
- Wiesenthal, H., (1987): "Rational Choice, Ein Überblick über Grundlinien, Theorienfelder und neuere Themenadquisition eines sozialwissenschaftlichen Paradigmas", Zeitschrift für Soziologie, vol. 16.
- Wiesenthal, H., (1990): "Unsicherheit und Multiple-Self-Identität. Eine Spekulation über die Voraussetzungen strategischen Handelns", Max-Planck-Institut für Gesellschaftsforschung, Discussion Paper 90/2, Colonia.
- Wiesenthal, H., (1990): "'Two Logics of Collective Action' Revisited. Oder: Ein zweiter Blick auf die 'strukturelle Machtposition des Kapitals'", Ms., Bremen.
- Wilde, K. et al., (1985): "Knowledge, Uncertainty and Behavior", American Economic Review. Papers and Proceedings, vol. 75.
- Williamson, O., (1975): Markets and Hierarchies. Analysis and Antitrust Implications, Free Press, Nova York.

- Wilson, J., (1986): "Subjective Probability and the Prisoner's Dilemma", Management Science, vol. 32.
- Wiston, G., (1989): "Imperfectly Rational Choice", Journal of Economic Behavior and Organization, vol. 12.
- Witt, U., (1986): "Evolution and Stability of Cooperation without Enforceable Contracts", KYKLOS, vol. 39.
- Witt, U., (1989): "The Evolution of Economic Institutions as a Propagation Process", Public Choice, vol. 62.
- Wright, E.O., (1985): Classes, Verso, Londres.
- Wrong, D., (1961): "The Oversocialized Conception of Man in Modern Sociology", American Sociological Review, vol. 26.
- Yamagishi, T., (1986): "The Provision of a Sanctioning System as a Public Good", Journal of Personality and Social Psychology, vol. 51.
- Yillock, H. i D. Morrison, (1979): "Group Size and Contributions to Collective Action: An Examination of Olson's Theory Using Data from Zero Population Growth, Inc.", Research in Social Movements, Conflict and Change, vol.2.
- Zemel, E., (1988): "Small Talk and Cooperation: A Note on Bounded Rationality", Journal of Economic Theory, vol. 49.
- Zinlt, R., (1986): "Ökonomisches Rationalitätskonzept und normorientertes Verhalten", Jahrbuch für Neue Politische Ökonomie, vol. 5.
- Zinlt, R., (1989): "Der Homo Oeconomicus: Ausnahmeerscheinung in jeder Situation oder Jedermann in Ausnahmesituation?", Analyse und Kritik, vol. 11.