

0108
3A
2g

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE FILOSOFÍA Y LETRAS
DIVISIÓN DE ESTUDIOS DE POSGRADO

LA RAZÓN DESENCANTADA

Jon Elster y la teoría de elección racional

TESIS
QUE PARA OPTAR POR EL GRADO DE
DOCTORA EN FILOSOFÍA
PRESENTA
ELISABETTA DI CASTRO STRINGHER

MÉXICO, D. F.

22943

1999



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradezco al CONACyT y a la Facultad de Filosofía y Letras de la UNAM el apoyo que me otorgaron para realizar mis estudios de doctorado. También a Mariflor Aguilar, Gerardo de la Fuente, Bolívar Echeverría, Carlos Pereda, Carmen Silva y Francisco Valdés la lectura y los comentarios que hicieron a la tesis. De manera especial quiero agradecer a Paulette Dieterlen la asesoría en la elaboración de este trabajo además de la amistad y el apoyo que siempre me ha brindado. A todos ellos mi más sincero reconocimiento

ÍNDICE

INTRODUCCIÓN.....	3
I. ALGUNAS OBSERVACIONES PREVIAS.....	12
1. Más allá de los <i>postmodernos</i>	13
2. ¿Razón vs pasiones?.....	24
3. El interés, una pasión peculiar.....	28
II. LA TEORÍA DE ELECCIÓN RACIONAL.....	42
1. Conceptos básicos.....	46
2. Estructura de la explicación.....	64
3. Límites y fracasos de la teoría.....	74
III. ENTRE SIRENAS.....	84
1. Ulises y las sirenas.....	85
2. Las sirenas de la razón.....	98
IV. LA RACIONALIDAD DE LOS DESEOS.....	115
1. Racionalidad amplia.....	116
i) Creencias.....	117
ii) Deseos.....	119
2. La incorporación de las emociones.....	129
i) Emociones racionales.....	138
ii) Deseos racionales.....	144
A MODO DE CONCLUSIONES.....	150
REFERENCIAS BIBLIOGRÁFICAS.....	161

INTRODUCCIÓN

La *razón* es uno de los temas y problemas fundamentales que ha acompañado a la filosofía a lo largo de su historia. Remontándonos al mundo griego, podemos recordar la insistencia de Aristóteles en que "lo peculiar del hombre es la actividad del alma, de acuerdo parcial o totalmente con la razón".¹ En el Renacimiento, entre las posibilidades que tiene el hombre, como un maravilloso camaleón, se encuentra el poder cultivar las semillas racionales que lo transformarían en un ser celestial.² Por su parte, los iusnaturalistas explican y justifican el estado social como un ente artificial construido a partir de un contrato que voluntariamente realizan los hombres.³ Finalmente, el pensamiento ilustrado centra sus esfuerzos y confianza en la capacidad de la razón para transformar no sólo al mundo sino también a los propios hombres: "¡Ten el valor de servirte de tu *propia* razón!: he aquí el lema de la ilustración."⁴

Dentro de este desarrollo de la filosofía, podemos ubicar dos momentos culminantes. La crítica de Kant a la metafísica, -con la que también se deja atrás la polémica entre racionalistas y empiristas-, parte de una concepción de la razón que es inmanente al sujeto de conocimiento y que es capaz de producir (tanto teórica como prácticamente) el "mundo", como ya lo habían descubierto los investigadores de la naturaleza al seguir el método experimental:

Entendieron que la razón sólo reconoce lo que ella misma produce según su bosquejo, que la razón tiene que anticiparse con los principios de sus juicios de acuerdo con leyes constantes y que tiene que obligar a la naturaleza a responder sus preguntas, pero sin dejarse conducir con andaderas, por así decirlo.⁵

¹ Aristóteles, *Ética Nicomaquea*, 1098a, p.1178.

² Cf. Giovanni Pico della Mirandola, *Discurso sobre la dignidad del hombre*, p. 49.

³ Cf. Thomas Hobbes, *Leviatán o la materia, forma y poder de una república eclesiástica y civil*, especialmente el capítulo XVII "De las causas, generación y definición de un estado", pp. 137-141.

⁴ Immanuel Kant, "¿Qué es la Ilustración?" en *Filosofía de la historia*, p. 25.

⁵ Immanuel Kant, *Crítica de la razón pura*, p. 18.

Con esta *revolución copernicana*, que permite establecer que “sólo conocemos *a priori* de las cosas lo que nosotros mismos ponemos en ellas”,⁶ el problema del conocimiento ya no es un problema de coincidencia con un supuesto orden ya establecido, sea natural o divino, sino que radica más bien en la capacidad que tiene la propia razón. De ahí la necesidad de su autoconocimiento y la búsqueda de garantías sobre sus pretensiones legítimas.

Para Hegel, este temor kantiano a equivocarse es ya el error mismo.⁷ La lograda unidad sujeto-objeto no puede limitarse al ámbito exclusivamente epistemológico en la medida en que el conocimiento no puede estar fuera de lo absoluto. Recordemos el célebre aforismo en el que se sintetiza la propuesta hegeliana: “Lo que es racional es real; y lo que es real es racional”.⁸ Aquí, la razón ya no es sólo inmanente al sujeto en tanto que individuo sino también es una cualidad de la realidad en donde éstos interactúan con su entorno y entre sí mismos:

Comprender lo que es, es tarea de la filosofía, porque lo que es, es la razón. Por lo que concierne al individuo, cada uno es, sin más, hijo de su tiempo; y, también, la Filosofía es el propio tiempo apprehendido con el pensamiento.⁹

Frente a la concepción abstracta universal, Hegel se preocupa por recuperar las determinaciones históricas, incorporar la historia en el concepto formal, elevando así la historia particular y lo real al rango de concepto racional.

Sin duda Kant y Hegel representan dos momentos culminantes de la historia de la filosofía, en donde la modernidad consolida uno de sus principios característicos: reconducir la “realidad” a la fuerza normativa de la razón. Ya sea desde la Ilustración o desde el Historicismo, esta “revolución racional”

⁶ *Ibid.*, p. 21.

⁷ Cf. G. W. F. Hegel, *Fenomenología del espíritu*, p. 52.

⁸ G. W. F. Hegel, *Filosofía del derecho*, p. 14.

⁹ *Ibid.*, p. 15.

marcó los principales problemas a los que tuvo que enfrentarse el siglo XIX y XX.¹⁰ Su piedra angular es el proceso de secularización por el que se deja atrás el tiempo medieval y la Ciudad de Dios, para pasar en cambio al tiempo del progreso y la Ciudad del Hombre; es el paso de la época de la comunidad a la de la sociedad, de un vínculo fundado en la obligación a otro basado en el contrato, de la voluntad sustancial a la voluntad electiva.¹¹

Si bien el término *secularización* surgió originalmente en el ámbito jurídico -durante la Reforma designaba la expropiación de los bienes eclesiásticos ya sea a favor del príncipe o de las iglesias nacionales reformadas-, en el siglo XIX el término se extendió hasta abarcar el sentido del desenvolvimiento histórico de la sociedad occidental moderna: racionalización y desencanto del mundo. Recordemos la formulación de Max Weber

[...] se sabe o se cree que en cualquier momento en que se *quiera* se *puede* llegar a saber que, por tanto, no existen en torno a nuestra vida poderes ocultos e imprevisibles, sino que, por el contrario, todo puede ser *dominado mediante el cálculo y la previsión*. Esto quiere decir simplemente que se ha excluido lo mágico del mundo.¹²

El nexo entre secularización, racionalización y desencanto del mundo se presenta como un destino histórico irreversible, aunque ello obviamente no supone un mundo sin ideologías ni una desacralización *tout court*, sino más bien, como ha señalado Giacomo Marramao, que la realidad en tanto racionalización del mundo sigue siendo la condición de la elección y del obrar

¹⁰ Cf. Luis F. Aguilar Villanueva, *Weber: la idea de ciencia social*, especialmente capítulo I "La apertura de la cuestión sobre el conocimiento de la historia: Kant y Hegel", pp. 21-71.

¹¹ Si bien hoy presenciamos alarmantes procesos de *desecularización* con la vuelta de diversos fundamentalismos, la modernidad le apostó a un despliegue continuo de racionalización (sobre la cual, sin embargo, ya Weber había advertido, con su célebre metáfora de la jaula de hierro, que también podía conducir a la irracionalidad). Cf. Max Weber, *La ética protestante y el espíritu del capitalismo*.

¹² Max Weber, "La ciencia como vocación" en *El político y el científico*, p. 200.

del individuo moderno.¹³ De esta manera, la íntima relación entre racionalidad y formas de actuar permite pensar al mundo moderno, a la realidad social, como un conjunto de acciones guiadas (o que pueden ser guiadas) por la racionalidad (ya sea de los propios individuos o de un espíritu absoluto).

Sin embargo, a finales del siglo XX, nos seguimos preguntando por la razón. Si bien podríamos decir que en algún sentido la razón siempre ha sido problemática, -incluso para los autores ubicados bajo la etiqueta de los racionalista, como era ya el caso de Descartes-, nuestra pregunta se caracteriza por cuestionar al mismo tiempo las esperanzas y expectativas puestas en ella. Si tuviéramos que dar alguna referencia, habría que señalar que al menos desde la Primera y Segunda Guerras Mundiales hemos descubierto que hay cosas peores que la propia muerte y que la historia no necesariamente implica progreso ni se mueve en esa dirección.¹⁴

Parece que así como en su momento se desencanto el mundo, también hay que desencantar a la razón, porque ésta se fue encantando dentro del proceso de secularización hasta convertirse en un mito e incluso llevar a la hiperracionalización. La *razón desencantada* sería la respuesta a la necesaria redefinición del concepto de razón que exige la caída de los grandes proyectos y supuestos del siglo XVIII.¹⁵

¹³ Básicamente, la noción de secularización remite a tres principios fundamentales: el principio de acción electiva o de autodecisión individual, el principio de diferenciación y especialización progresivas (aumento de complejidad) y el principio de legitimación o de cambio. Aunque para Hans Blumenberg la auténtica metáfora de la modernidad no es la secularización (que sería un término derivado) sino la legitimidad; aquí el énfasis de la revolución copernicana se pone en la capacidad del individuo para apoderarse de nuevo de su propio destino. (Cf. Giacomo Marramao, *Poder y secularización*, pp. 23-27)

¹⁴ Esto tampoco tiene que significar el fin de las ideologías o de la propia historia, como se insistió a raíz de la caída del Muro de Berlín (Cf. Francis Fukuyama, *El fin de la historia*). Precisamente la pérdida de fundamentos y garantías últimas, puede llevar a que las diversas posiciones, sean teóricas o políticas, adquieran una mayor y pesada relevancia en tanto son una toma de posición que hay que construir y poder defender.

¹⁵ Como veremos en el primer capítulo, con ello no pretendemos una redefinición "postmoderna", más bien, simpatizando con el planteamiento del sugerente libro de Giacomo Marramao, *L'ordine disincantato*, se trataría más bien de una redefinición "postclásica": no se trata de liquidar o disolver los núcleos de la reflexión clásica, sino de traducirlos o transferirlos a un código categorial y epistemológico distinto.

La exigencia de esta necesaria redefinición surge desde diversos ámbitos, no exclusivamente desde la filosofía. Por ejemplo, en las matemáticas, frente a las pretensiones de una formalización exhaustiva, destaca el famoso trabajo de Kurt Gödel sobre la incompletud de la aritmética;¹⁶ en la física, frente a las visiones deterministas, se levantan los sistemas dinámicos inestables cuyas leyes sólo pueden expresar posibilidades;¹⁷ en la economía, frente a las soluciones óptimas del análisis económico clásico, Herbert A. Simon incorpora la noción de elecciones satisfactorias;¹⁸ en las ciencias sociales, frente a la elección social colectiva, Kenneth J. Arrow planteó su ya célebre teorema en el que se cuestiona la compatibilidad de la democracia y las elecciones sociales racionales;¹⁹ finalmente, en la filosofía, frente a la racionalidad cognitivo-instrumental, Jürgen Habermas desarrolla el concepto de racionalidad comunicativa.²⁰ Con estos ejemplos sólo mencionamos algunos de los casos más significativos.

El problema de esta investigación es el concepto de *razón*. Frente a un excesivo optimismo se plantea la pertinencia de una redefinición del concepto de razón en la que, si bien se cuestionan las esperanzas y expectativas depositadas en ella, no se presupone tampoco una reivindicación del *reino de lo irracional*. Parecería que a lo largo de la historia, las diversas definiciones se han elaborado a partir de ciertas contraposiciones básicas, como son racional-irracional, universal-particular y necesario-contingente. La hipótesis es que ya no es tan fácil, si es que alguna vez lo fue, contraponer lo racional (vinculado

¹⁶ Cf. La divulgación de la propuesta gödeliana que hace Alfred Tarski en el artículo "Truth and Proof" publicado en *Scientific American*, junio 1969, pp. 63-73.

¹⁷ Cf. Ilya Prigogine, *El fin de las certidumbres*. El Premio Nobel de Química 1977 llega incluso a afirmar que con los últimos desarrollos de la física nos encontramos "en el punto de partida de una nueva racionalidad que ya no identifica ciencia y certidumbre, probabilidad e ignorancia." (*Ibid.*, p. 15)

¹⁸ Cf. Herbert A. Simon, "De la racionalidad sustantiva a la procesal" en *Filosofía y teoría económica*, comps. Frank Hahn y Martin Hollis, pp. 130-171. Para el Premio Nobel de Economía 1981, "está condenado al fracaso todo esfuerzo tendiente a pronosticar y prescribir el comportamiento económico humano a partir de la inferencia deductiva surgida de un pequeño conjunto de premisas irrefutables." (*Ibid.*, p. 167)

¹⁹ Cf. Kenneth J. Arrow, *Social Choice and Individual Values*.

²⁰ Cf. Jürgen Habermas, *Teoría de la acción comunicativa*.

con lo universal y necesario) a lo irracional (vinculado con lo particular y contingente). Partimos de la idea de que ya no se puede sostener, e incluso puede llegar a ser "peligrosa", la defensa de un concepto de razón elaborado a partir de dichas contraposiciones. Más que enfrentamiento entre dos ámbitos contrapuestos, lo racional, lo universal y lo necesario también pueden presuponer de manera interna, y sin subsumirlos o asimilarlos, sus respectivos espacios irracionales, particulares y contingentes. Parecería que, perdidos los fundamentos y las garantías, las definiciones dicotómicas se vuelven demasiado pobres y ahora el problema es más bien pensar precisamente los juegos y mutuas implicaciones. Con esto lo racional deja de presentarse como una esfera compacta, se vuelve porosa, con fisuras; asimismo, el concepto de razón deja de ser uno y universal, pasa a ser plural y local. Finalmente, nos encontramos con una multiplicidad de imágenes que ya no pueden ser englobadas por una supuesta universalidad porque no son necesariamente compatibles entre sí.

Aquí es pertinente hacer dos observaciones. Esta pluralidad de imágenes de la razón, lejos de desembocar en el "todo es igual" o en el nihilismo, subraya más bien la relevancia de las especificidades problemáticas, así como la necesaria toma de posición en toda construcción teórica. De esta manera, la contraposición entre lo racional y lo irracional, así como entre lo universal y lo particular, y lo necesario y lo contingente, serían una de las posibles imágenes de la razón, cuyo sentido y relevancia obedece a los problemas específicos a los que se pretende dar respuesta.

La toma de posición que anima este trabajo no es novedosa. Recupera el espíritu de un sector importante de las discusiones de las últimas décadas, que bien puede quedar resumido en la siguiente expresión de Javier Muguerza:

[...] por más que las lanzas de la moderna racionalidad no pasen muchas veces de ser frágiles cañas, creo sinceramente que aún cabe romper alguna que otra en pro de la razón, moderna o no.²¹

Al menos desde los años setenta, entre las publicaciones filosóficas podemos observar una preocupación creciente, con distintos tonos y perspectivas, por la revisión del concepto de razón. Mencionemos como ejemplos, además de los estudios sobre la naturaleza de la razón,²² las reflexiones sobre la crisis de la razón,²³ la razón sin esperanza,²⁴ la razón erótica,²⁵ y la razón enfática.²⁶

Más allá de las propuestas concretas de los diversos autores, aquí lo que nos interesa destacar es el cuestionamiento de algunos conceptos básicos, como son los de totalidad, sujeto, sustancia, necesidad y progreso. Asimismo, que la caída del supuesto optimismo en la razón, heredado del siglo XVIII, no implica entrar al *reino de lo irracional*, sino exige más bien reformular y redefinir lo que podemos entender por racional. Esta reformulación se puede rastrear a partir no sólo de los usos y límites de la razón sino también de otras alternativas de racionalidad. De hecho, la redefinición del concepto de razón que, como ya señalamos, se está llevando a cabo en diversas disciplinas, no sólo en la filosofía, cuestiona algunos de los principales supuestos ontológicos y epistemológicos con los que se ha caracterizado a la modernidad, al tiempo que elabora y recupera nuevos conceptos, entre ellos los de incertidumbre, aleatoriedad, incompletud, complejidad y riesgo.

Nosotros nos centraremos en el análisis de una de las teorías contemporáneas que consideramos especialmente significativa: la teoría de elección racional. Decimos que es especialmente significativa no sólo porque se

²¹ Javier Muguerza, *La razón sin esperanza (Siete trabajos y un problema de ética)*, p. 20.

²² Cf. Nicholas Rescher, *La racionalidad. Una indagación filosófica sobre la naturaleza y la justificación de la razón*; y Robert Nozick, *La naturaleza de la racionalidad*.

²³ Cf. Aldo Gargani (ed.), *Crisi della ragione. Nuovi modelli nel rapporto tra sapere e attività umane*.

²⁴ Cf. Javier Muguerza, *op. cit.*

²⁵ Cf. Antoni Domènech, *De la ética a la política. De la razón erótica a la razón inerte*.

²⁶ Cf. Carlos Pereda, *Razón e incertidumbre*.

presenta como una teoría que defiende un concepto fuerte de razón, sino también por el impacto que ha tenido a nivel práctico en la toma de decisiones individuales y colectivas. Por otra parte, además de ser una propuesta que, desde la perspectiva filosófica, ha sido poco trabajada en México, a sus críticos internos les ha tocado ocupar un *no-lugar* peculiar: no existen para los defensores ortodoxos de la teoría, pero tampoco para sus acérrimos críticos externos.²⁷

Desde la teoría de elección racional rastrearemos el deslinde de los supuestos que ahora podemos llamar tradicionales así como la construcción de nuevos conceptos y problemas. Este análisis nos proporcionará los elementos con los que, a modo de conclusiones, trataremos de esbozar los líneas de una posible redefinición del concepto de razón. Para ello, nos apoyaremos en la obra de uno de sus más apasionados defensores que es también, al mismo tiempo, uno de sus más incisivos críticos: Jon Elster.

El trabajo está estructurado en cuatro partes:

1. En la primera, presentamos una breve ubicación histórica de algunos conceptos claves de la investigación. En primera instancia cuestionamos las visiones simplistas que vinculan esquemáticamente la modernidad con la idea de un razón fuerte. En especial ubicamos la relación razón-pasiones a partir de dos de sus principales propuestas y concluimos con algunos pasajes de la historia moderna de una pasión peculiar: *el interés*.
2. En la segunda parte, presentamos un panorama general de la teoría de elección racional en el que ubicamos sus principales conceptos y problemas. Veremos sus conceptos básicos, como son la racionalidad optimizadora, la

²⁷ Parecería que con la asignación del Premio Nobel de Economía 1998 a Amartya Sen se empieza a reconocer la necesidad de reintroducir la reflexión filosófica en los estudios y discusiones económicas. Aunque también es probable que para los economistas ortodoxos ese año no hubo propiamente Premio Nobel de Economía.

paramétrica y la estratégica; además de la estructura de su explicación, que descansa en una triple relación entre acción, creencias y deseos del agente; para concluir con un punteo de sus principales fallas, las cuales obedecen básicamente a la negación de las condiciones de optimalidad o a la negación de lo que normativamente prescribe la teoría.

3. En la tercera parte, nos centramos en dos fallas principales de la teoría de elección racional: cuando el agente actúa contra su mejor juicio y cuando el agente debe decidir bajo incertidumbre. El primer caso lleva a postular la necesidad de reconocer una racionalidad imperfecta y el segundo a advertir sobre una posible adicción a la razón. Frente a ambas fallas se presenta como una solución razonable la abdicación racional de la razón.

4. En la cuarta y última parte, nos detenemos en uno de los elementos básicos de la explicación de elección racional: los deseos del agente. Frente a la *teoría estricta* de la racionalidad se postula la pertinencia de una *teoría amplia* que se pregunte también por la racionalidad de los deseos. Finalmente, con este planteamiento, la teoría de elección racional termina por incorporar tanto la autonomía como las emociones del agente. De esta manera, se va conformando un modelo explicativo mucho más complejo, tal vez más acorde con la vida de los hombres que también es mucho más compleja que una simple jerarquización de preferencias guiadas por el egoísmo.

I. ALGUNAS OBSERVACIONES PREVIAS

En esta primera parte, presentamos algunas observaciones que consideramos pertinentes antes de analizar la teoría de elección racional. Básicamente se trata de una breve ubicación histórica en relación a algunos términos que son claves para nuestra investigación. Estas observaciones se agrupan en tres secciones:

1. En la primera, intentamos plantear que, más allá de los posibles aciertos del pensamiento llamado *postmoderno*, su difusión llevó a una visión esquemática de la modernidad frente a la cual cualquier cuestionamiento a una razón fuerte podría etiquetarse de “postmoderno”. La intención de esta sección es rescatar una visión más compleja de la modernidad y subrayar que en sus *clásicos* también podemos encontrar posiciones no tan optimistas con respecto a la capacidad o alcances de la razón.
2. En la segunda sección, nos centramos en la relación razón-pasiones y punteamos algunos de los aspectos más relevantes de dos propuestas altamente significativas: en primer lugar, veremos la relación de enfrentamiento entre razón y pasiones, que aun hoy es bastante generalizada, y que está ejemplificada por Descartes; después esbozamos la posición de Hume en donde la razón, lejos de enfrentarse a las pasiones, se une a ellas para alcanzar de la mejor manera los fines del hombre cuando actúa.
3. En la tercera y última sección, tratamos de rescatar algunos pasajes de la historia moderna del término *interés*, en tanto fruto del intento por

contrarrestar las pasiones destructivas de los hombres y darle cierta estabilidad a la convivencia humana.

1. Más allá de los *postmodernos*

Entre los pasajes más citados de las dos últimas décadas, sin duda se encuentra el siguiente:

En origen, la ciencia está en conflicto con los relatos. Medidos por sus propios criterios, la mayor parte de los relatos se revelan fábulas. Pero, en tanto que la ciencia no se reduce a enunciar regularidades útiles y busca lo verdadero, debe legitimar sus reglas de juego. Es entonces cuando mantiene sobre su propio estatuto un discurso de legitimación, y se la llama filosofía. Cuando ese metadiscurso recurre explícitamente a tal o tal otro gran relato, como la dialéctica del Espíritu, la hermenéutica del sentido, la emancipación del sujeto razonante o trabajador, se decide llamar "moderna" a la ciencia que se refiere a ellos para legitimarse.¹

Inmediatamente después se aclara que en las sociedades más desarrolladas (postindustriales) el estatuto del saber ha cambiado y se propone caracterizarlo como "postmoderno":

Simplificando al máximo, se tiene por "postmoderna" la incredulidad con respecto a los metarrelatos. Esta es, sin duda, un efecto del progreso de las ciencias; pero ese progreso, a su vez, la presupone. Al desuso del dispositivo metanarrativo de legitimación corresponde especialmente la crisis de la filosofía metafísica, y la de la institución universitaria que dependía de ella. [...] Así, la sociedad que viene parte menos de una antropología newtoniana (como el estructuralismo o la teoría de sistemas) y más de una pragmática de las partículas lingüísticas. Hay muchos juegos de lenguaje diferentes, es la heterogeneidad de los elementos. Sólo dan lugar a una institución por capas, es el determinismo local.²

¹ Jean-François Lyotard, *La condición postmoderna. Informe sobre el saber*, p. 9.

² Jean-François Lyotard, *Ibid.*, p. 10.

Entre los principales *topoi* de la postmodernidad podemos mencionar el adiós al pensamiento del un progreso unilineal, la comprensión de las diferencias y los conflictos a partir de la ruptura con las exigencias de unidad y totalidad, así como la propuesta de que, sin fundamentos primeros (o últimos), las diversas relaciones remiten permanentemente a otras.³ A partir de este adiós a las "ilusiones fundamentalistas", podríamos ciertamente simpatizar con los planteamientos postmodernos, sin embargo, éstos también han alentado perspectivas catastrofistas, la arbitrariedad, el *todo se vale*, el irracionalismo y el cinismo.⁴ Obviamente, se podría tratar de distinguir diversas versiones de la "postmodernidad" e incluso, dentro de ellas, ubicar los elementos que se podrían considerar adecuados o rescatables. Nosotros no nos ocuparemos de ello, sólo pretendemos señalar que muchas de los planteamientos teóricos de la cultura contemporánea cumplirían con algunas características que acertadamente han descrito los postmodernos, lo cual no lleva necesariamente a asumir dicho planteamiento en su conjunto y mucho menos sus consecuencias.

Mas allá de la pretendida contraposición entre modernidad y postmodernidad, -en donde el inicio de la primera se ubica normalmente en Bacon y Descartes, así como su término y el inicio de la segunda en Nietzsche, Weber y Heidegger-, esta dicotomía ha llevado a difundir una visión esquemática de la modernidad que de manera paradójica muchos intelectuales han asumido para manifestar precisamente su rechazo a las pretensiones de

³ Cf. Wolfgang Iser, "Topoi de la posmodernidad" en *El final de los grandes proyectos*, comps. Hans Rudi Fischer, Arnold Retzer y Jochen Schweitzer, pp. 36-56.

⁴ Aquí no nos ocuparemos propiamente del pensamiento *postmoderno*, sin embargo podemos mencionar cómo el *espíritu postmoderno* puede terminar justificando (incluso alentando) el cinismo y la (re)caída en los grandes prejuicios que alguna vez se creyó definitivamente desterrados por el pensamiento crítico. (Cf. Bolívar Echeverría, *Las ilusiones de la modernidad*, especialmente los ensayos "Postmodernidad y cinismo" y "La identidad evanescente", pp. 39-74)

los autollamados o no *postmodernos*.⁵ Aunque por otro lado, debemos reconocer que este enfrentamiento también ha estimulado una fértil relectura de los *clásicos* de la modernidad ya que, a partir de sus propios textos, se ha tratado de hacerle frente a esa pobre etiqueta que busca encasillarlos. Empecemos por establecer, aunque sea a grandes rasgos, cuál sería esa visión.⁶

En la contraposición entre modernidad y postmodernidad, se ha llegado a caracterizar a la primera como la época:

- a) de una razón fuerte que pretende construir explicaciones totalizantes del mundo y que está ligada a la idea de que el desarrollo histórico del pensamiento es interminable e implica una progresiva iluminación;
- b) del orden epistemológico de la razón con una estructura monolítica y unificante;
- c) de la apropiación y reapropiación de los fundamentos (el pensamiento como acceso al fundamento);
- d) de la autolegitimación del saber científico y de la total coincidencia entre verdad y emancipación;
- e) del tiempo lineal que se caracteriza por la superación (una novedad envejece y es sustituida por otra); y,

⁵ Nos referimos a la difusión de los planteamientos postmodernos. Sin embargo, haciéndole un poco de justicia, habría que señalar que propiamente la "postmodernidad" no se presenta como una determinación temporal, sino más bien como un nuevo enfoque y actitud que actualmente es el predominante pero que está presente también en algunas propuestas de la época moderna. El propio Lyotard señala: "A medida que la discusión se desarrolla en el plano internacional, la complejidad de la 'cuestión posmoderna' se agrava. Cuando la enfoqué, en 1979, en torno de la cuestión de los 'grandes relatos', mi intención era simplificarla, pero me temo que fui más allá de lo necesario." (Jean-François Lyotard, *La posmodernidad (explicada a los niños)*, p. 29)

⁶ Para ello seguiré el sugerente artículo de Paolo Rossi, "Idola della modernità" (en *Moderno-postmoderno. Soggetto, tempo, sapere nella società attuale*, ed. de G. Mari, pp. 14-31), en donde el autor destaca los principales elementos que, de acuerdo con las publicaciones de la década 1975-1985, configurarían lo moderno en contraposición a lo postmoderno.

- f) de la positividad del desarrollo tecnológico cuyos proyectos pueden prever y dominar de manera total.

Frente a ello, la postmodernidad se ha llegado a caracterizar como la época:

- a) de un debilitamiento de las pretensiones de la razón que cuestiona el sentido de la historia y del humanismo;
- b) de la emergencia de una pluralidad de modelos y paradigmas de racionalidad no homogéneos ni reducibles uno al otro sino sólo vinculados a la especificidad de su respectivo campo de aplicación;
- c) de un pensamiento sin fundamentos (una crítica de la razón instrumental que reconoce el carácter enigmático);
- d) del reconocimiento del carácter discontinuo y paradójico del desarrollo de la ciencia;
- e) de la disolución de la categoría de lo nuevo y de la experiencia del fin de la historia; y,
- f) de la aparición de la ciencia y la técnica como riesgosas más que liberadoras del esfuerzo y de la necesidad.

De acuerdo con esta visión, en donde la modernidad queda linealmente identificada con la *racionalidad* y la *ciencia*, las preocupaciones por una *razón débil*, por la *pluralidad de racionalidades* y por la *carencia de fundamentos*, estarían asociadas con las propuestas *postmodernas*.⁷ Sin embargo, al menos por lo que se refiere en especial a la tradición filosófica, aunque ciertamente no sólo en ella, se puede encontrar una diversidad de alternativas que, a lo largo

de su desarrollo, se componen y recomponen también de diversas maneras, con diferentes ritmos y niveles. De hecho, se trata de una tradición en la que, finalmente, entran en juego temas que todavía hoy son *nuestros* temas.⁸ En este sentido, las visiones simplistas y las esquemáticas contraposiciones en la historia de la filosofía, pueden llevar a desconocer la rica diversidad de posiciones que se da y se ha dado en toda polémica. Ello, no sólo por lo que se refiere en especial a la oposición modernidad-postmodernidad, sino también por ejemplo, dentro del ámbito de la primera, en el debate racionalistas-empiristas. Estos enfrentamientos, lejos de dejarnos tranquilamente en las contraposiciones, nos llevan justo a preguntarnos por la diversidad dentro de cada una de las posiciones enfrentadas.

En general, las propuestas que nos invitan a pensar por parejas antitéticas no sólo lo hacen con fines clasificatorios sino también como una estrategia de análisis.⁹ Y, si bien las dicotomías pueden volverse muy atractivas en tanto permiten sistematizar la historia del pensamiento -y de hecho también se pueden presentar como una estrategia de análisis conceptual muy efectiva por lo que se refiere a un primer acercamiento o esbozo general-, normalmente las visiones dicotómicas se vuelven demasiado rígidas y esquemáticas frente a la complejidad y riqueza del objeto a analizar.

⁷ Como mencionamos en la introducción, frente a esta asociación simple se pueden recuperar diversas propuestas que se han centrado en destacar precisamente las diferencias entre distintos tipos de racionalidades.

⁸ Aunque, como bien ha señalado Isaiah Berlin, "Uno de los signos de los escritores de genio es que lo que dicen puede, a veces, tocar un nervio central en las mentes o los sentimientos de los hombres que pertenecen a otras épocas, culturas o perspectivas, y establecen trenes de pensamiento y arrastran consecuencias que no se les ocurrieron o pudieron no ocurrírseles a tales escritores, y todavía menos preocupar sus mentes." (Isaiah Berlin, "Vico y el ideal de la Ilustración" en *Contra la corriente. Ensayos sobre historia de las ideas*, p. 188)

⁹ Por ejemplo, Norberto Bobbio nos ofrece un análisis de la historia del pensamiento político occidental a partir de parejas conceptuales antitéticas que dividen el universo de estudio en dos subclases recíprocamente exclusivas y conjuntamente exhaustivas. Entre las ventajas de esta estrategia, el autor ha destacado tres *usos privilegiados*: definir negativamente un término (el débil) a partir del otro (el fuerte), delinear una filosofía de la historia y evidenciar el posible juicio de valor (Cf. Norberto Bobbio, "La grande dicotomia" en *Dalla Struttura alla funzione. Nuovi studi di teoria del diritto*, especialmente pp. 145-148).

De esta manera, y más allá de las posibles críticas y divergencias con lo que se ha llamado el pensamiento *postmoderno*, debemos reconocer que éste ha planteado como uno de los grandes problemas contemporáneos el rescate de las diferencias y de lo inconmensurable. Obviamente, el aceptar esta problemática no tiene porque implicar también la aceptación de que el *saber postmoderno* es el que "Hace más útil nuestra sensibilidad ante las diferencias, y fortalece nuestra capacidad para soportar lo inconmensurable."¹⁰

Si partimos de que nuestro acceso a la *realidad* es mediado a través de la construcción de *mapas* y de la elección de *hilos conductores*, y, que un conjunto finito de datos es compatible con una diversidad de hipótesis que pueden ser a su vez incompatibles entre sí, entonces no contamos, ni podremos contar, con *el mapa* ni *el hilo conductor*. En otras palabras, no hay una única reducción de la *realidad* sino una multiplicidad cuyas propuestas pueden ser igualmente legítimas aunque no por ello igualmente fecundas. Así, en el caso específico de la reflexión filosófica, se debe reconocer la diversidad de alternativas e ideas que se han producido, acompañado y recompuesto de múltiples maneras a lo largo de la historia. Como bien ha señalado Paolo Rossi,

No es verdad que, como cada hombre tiene una sola cara, entonces cada edad ha tenido y debe tener un propio e inconfundible rostro, una específica *episteme*. El diálogo crítico entre teorías, tradiciones, metafísicas, ideologías, imágenes de la ciencia, métodos de investigación ha existido siempre y todavía -al contrario- continua, insistente, real. Incluso en las épocas que fueron polémicamente ocultadas en una sola dimensión -aquella de la oscuridad y de la supersticiosa barbarie- como acontece para el Medioevo. No hay ni épocas, ni estadios mono-paradigmáticos.¹¹

¹⁰ Jean-François Lyotard, *La condición postmoderna*, p. 11.

¹¹ Paolo Rossi, *op. cit.*, pp. 23-24.

Obviamente, entre la diversidad de propuestas también se encuentra la estrategia dicotómica. Estrategia que, a pesar de ser poco fértil,¹² ha tenido un gran éxito y una amplia difusión. Esta situación se ha llegado a explicar en los siguiente términos:

Pero hay una explicación también para la difundida adopción de esquemas extraordinariamente rígidos y al pensar por dicotomías. La transformación del pasado en un desierto o en el lugar de la infinita repetición de la misma cosa no está privada de elementos fuertemente autogratificantes. No es difícil sentirse originales portadores de mensajes radicalmente nuevos si, sobre la base de alguna fórmula, alcanzamos primero a reducir, después a eliminar la variedad, la multiplicidad, la contradictoriedad de las voces del pasado y de las del presente.¹³

Como dijimos al inicio, la contraposición moderno-postmoderno ha llevado a difundir una visión esquemática de la modernidad que incluso asumen algunos que rechazan las pretensiones *postmodernas*. Frente a ello, creemos pertinente rechazar la propia contraposición junto con sus supuestos; en especial, el enfrentamiento entre visiones simplistas de *la racionalidad y la ciencia*, por un lado, y, por otro, *lo irracional y lo acientífico*. Finalmente, al menos desde el siglo XVII sabemos que son “Dos excesos: excluir la razón, admitir sólo la razón.”¹⁴

Es precisamente en el siglo XVII en donde se ubica la gran revolución *espiritual* que tiene como *raíz y fruto a la ciencia moderna*. Revolución que se ha caracterizado, como ya señalamos en la introducción, por la secularización

¹² “Es de verdad difícil (y en general escasamente útil) hacer converger todo lo que acaece en una única, gran ruptura bien delimitada y representable en la forma de una oposición simple. Las contraposiciones, las ‘parejas oposicionales’, el pensar por dicotomías y por alternativas radicales son instrumentos insuficientes y parciales. No sirven de hecho, por ejemplo, para entender las relaciones físico-mente, teórico-observacional, humano-animal, naturaleza-cultura, interno-externo, magia-ciencia, relativismo-realismo. Dicotomías como éstas se revelan, conforme la investigación procede, capaces sólo de bloquear las indagaciones.” (Paolo Rossi, *Ibid.* p. 23)

¹³ Paolo Rossi, *Ibid.*, p. 24.

¹⁴ Blaise Pascal, *Pensieri*, parágrafo 128 (253 Brunschvicg), p. 148.

de la conciencia, el descubrimiento de la subjetividad y la aspiración de dominar a la naturaleza. Estas características se presentan como los aspectos concomitantes y expresión de la destrucción del viejo cosmos (finito, cerrado y jerárquicamente ordenado) y su necesaria sustitución por un nuevo universo (indefinido, infinito y con componentes idénticos).¹⁵ Imagen del mundo que se fue esbozando en los dos siglos anteriores, por un lado, y, por otro, que se consolidará definitivamente en el siglo posterior.¹⁶

Entre los principales supuestos de la llamada *Ilustración* se pueden destacar los siguientes: las cosas tienen una naturaleza permanente, los cambios en el mundo obedecen a leyes universales y estas leyes pueden ser descubiertas gracias al uso correcto de la razón y una pertinente observación. Por lo que se refiere en especial a la naturaleza humana, se pueden aplicar métodos similares a los de la física en la ética y la política, lo que posibilitaría la construcción de un gobierno de la razón que finalmente encauzaría al hombre hacia la felicidad alejándolo de las injusticias políticas y miserias morales.

Frente a este gran optimismo, hay que destacar que sus críticos son también *tan viejos como el movimiento mismo*:

Esta fe en los poderes de la razón y la ciencia de ninguna manera fue universalmente apoyada en la Europa occidental, ni aun a mitad del siglo XVIII; cuando menos no con igual confianza y fervor: fue regularmente asaltada por las dudas insidiosas de los escépticos, por la hostilidad de los defensores ortodoxos de la autoridad de la iglesia y del estado, por los defensores de la variedad, individual y local, y de los valores locales y tradicionales, así como por los campeones de la imaginación artística, libre de reglas y regulaciones universales[...]¹⁷

¹⁵ Cf. Alexandre Koyré, *Del mundo cerrado al universo infinito*.

¹⁶ Cf. Luis Villoro, *El pensamiento moderno. Filosofía del Renacimiento* y Norberto Bobbio, "El modelo iusnaturalista" en *Sociedad y Estado en la filosofía política moderna*.

En los inicios de la modernidad, aun sin desafiar abiertamente las creencias cristianas, se hace frente a la autoridad de la tradición con el tribunal de la razón; para ello se ubican las principales fuentes del error y se sistematiza el método crítico. Tenemos por ejemplo, por un lado, a Francis Bacon para quien "la verdad, que no tiene más juez que ella misma, nos enseña que su investigación, conocimiento y sentimiento [...] son el mayor bien que puede concederse a los mortales."¹⁸ Por otro, también se encuentra René Descartes para quien

El buen sentido es la cosa mejor repartida del mundo [...] la facultad de juzgar y distinguir lo verdadero de lo falso, que es propiamente lo que llamamos buen sentido o razón, es naturalmente igual en todos los hombres [...aunque] No basta, en efecto, tener el ingenio bueno; lo principal es aplicarlo bien.¹⁹

Sin embargo, el rechazo al principio de autoridad (ya sea de la tradición o incluso de las sagradas escrituras) así como la defensa de la autonomía de la razón y del método de las ciencias naturales, tuvo desde el principio opositores, no sólo por parte de las iglesias sino también por parte de una tradición relativista y escéptica que se remonta hasta los antiguos griegos.

Así, también se llega a postular que los valores morales y políticos descansan sobre convenciones humanas, las cuales son obviamente cambiantes. Incluso, se llega a cuestionar la existencia de eslabones lógicos entre las verdades de *hecho* y las verdades *a priori* de la lógica y la matemática. De esta manera, la pretensión de establecer un único sistema de conocimiento a partir de cadenas lógicas de axiomas válidos universalmente (no sujetos a modificación o refutación por experiencia de tipo empírico) se empieza a tambalear.

¹⁷ Isaiah Berlin, "Hume y las fuentes del antirracionalismo alemán" en *Contra la corriente*, p. 235.

¹⁸ Francis Bacon, "De la verdad" en *Ensayos sobre moral y política*, pp. 2-3.

¹⁹ Rene Descartes, *Discurso del método*, p. 29.

El ejemplo por excelencia es Hume. Aquí sólo mencionaremos algunos aspectos especialmente significativos frente a esa rígida caracterización de la modernidad que presentamos al inicio del capítulo.²⁰ La propuesta humeana parte de una crítica a la razón no sólo como facultad primordial del conocimiento sino también como esencia distintiva del hombre: el conocimiento y la identidad personal descansan más bien en la imaginación. Asimismo, la propuesta humeana parte de una crítica al sustancialismo como fundamento del conocimiento y de la identidad personal: el conocimiento y la identidad personal son tratados como problemas de creencias, sobre las cuales ni siquiera se pregunta por su fundamento o verdad, sino más bien por su origen. Recordemos cómo formula su toma de posición:

Ahora bien: dado que nada hay presente a la mente sino las percepciones, y que todas las ideas se derivan de algo que con anterioridad se hallaba ya ante la mente, se sigue que nos es imposible concebir o formar una idea de algo que sea específicamente distinto a las ideas e impresiones. Dirijamos nuestra atención fuera de nosotros cuanto nos sea posible; llevemos nuestra imaginación a los cielos, o a los más extremos límites del universo: nunca daremos realmente un paso fuera de nosotros mismos, ni podremos concebir otra clase de existencia que la de las percepciones manifiestas dentro de esos estrechos límites. Este es el universo de la imaginación, y no tenemos más ideas que las allí presentes.²¹

La idea de causalidad, de mundo externo e incluso la del propio yo, son creencias que, al descansar en la experiencia, se ubican en el ámbito de lo probable, no en el del conocimiento necesario, aunque la constancia y las costumbres nos lleven a concebirlas como algo inevitable. Por ello, si bien la sensación o experiencia es necesaria, ésta no es suficiente para explicar el

²⁰ Para ello recuperaremos algunas de las sugerentes observaciones del trabajo de Carmen Silva, "El naturalismo humeano y el problema mente-cuerpo" publicado en *El problema de la relación mente-cuerpo*, comps. Laura Benítez y José A. Robles, pp. 163-177.

²¹ David Hume, *Tratado de la naturaleza humana*, Libro I, Parte II, 2 vol., p. 169.

origen de las creencias, son los principios de asociación los que finalmente proporcionan la explicación de las creencias. En este sentido, tenemos que para Hume, la explicación de las cuestiones humanas no descansa en la razón, sino más bien en la imaginación.

Con este ejemplo y los señalamientos anteriores, podemos reconocer una diversidad dentro de las propuestas modernas. Con ello, sin embargo, tampoco se pretende desconocer que efectivamente hay algunos elementos compartidos por los diversos autores modernos, y que de hecho son los que permitirían seguir hablando de la modernidad como una época del desarrollo de la humanidad y en especial de la filosofía. Por ejemplo, incluso los críticos de los defensores de la autonomía de la razón y del método de las ciencias naturales, compartirían un núcleo común de ideas: todo hombre busca satisfacer sus necesidades básicas, no sólo por lo que se refiere al alimento y a la seguridad sino también la paz, la felicidad, la justicia, la verdad, la virtud e incluso, de alguna u otra manera, la perfección moral. Si bien los medios pueden diferir, los fines últimos son fundamentalmente semejantes. Como destaca Isaiah Berlin,

Un grupo de principios universales e inalterables gobernaba el mundo para deístas y ateos, para optimistas y pesimistas, puritanos, primitivos y creyentes en el progreso y en los más ricos frutos de la ciencia y la cultura; estas leyes gobernaban la naturaleza animada e inanimada, los hechos y los acontecimientos, medios y fines, la vida privada y pública, todas las sociedades, épocas y civilizaciones; sólo por apartarse de ellos los hombres caían en el delito, el vicio, la miseria. Los pensadores podían diferir acerca de qué leyes eran éstas, o como descubrirlas, o quienes estaban calificados para exponerlas; que estas leyes eran reales y podían ser conocidas, ya sea con certeza o sólo con probabilidad, seguía siendo el dogma central de la Ilustración.²²

²² Isaiah Berlin, "La contra-ilustración" en *Contra la corriente*, p. 62.

De esta manera, si bien no se pretende negar que la modernidad está marcada efectivamente como una época que rompe con una vieja concepción del mundo y su correspondiente pensamiento tradicional,²³ por ahora nos interesa destacar que se va dibujando como un mosaico mucho más complejo y rico de la caracterización que se esbozamos al inicio. Por lo que se refiere en especial al concepto de razón, tratamos de subrayar la presencia de propuestas que precisamente ponen en cuestión el primer inciso de dicha caracterización: una razón fuerte que pretende construir explicaciones totalizantes del mundo y que está ligada a la idea de que el desarrollo histórico del pensamiento es interminable e implica una progresiva iluminación.

2. ¿Razón vs pasiones?

Dentro de la diversidad de propuestas que podemos encontrar en la modernidad, para los fines del presente trabajo, nos interesa destacar especialmente una característica: la contraposición entre razón y pasiones. Parecería, como ha planteado Remo Bodei, que la razón es como una espejo de agua, que cuando es agitado por las pasiones deja de reflejar la realidad e impide que la voluntad encuentre alternativas a las inclinaciones del momento.²⁴ En este sentido, las pasiones terminan por ser condenadas en tanto perturban o anulan temporalmente a la razón. Pensemos primero en las propuestas que, de manera muy optimista, parten de una confianza en la capacidad de la razón, que por cierto está ligada al desarrollo de las matemáticas modernas. Recordemos a Descartes:

²³ Aunque también aquí no habría que suponer una época *monoparadigmática*. Así como entre las consecuencias positivas del discurso *postmoderno* destacamos una fértil relectura de los *clásicos* de la modernidad, mencionemos que también estimuló una nueva mirada al Medioevo. Si bien rebasa los límites de nuestro trabajo, podemos plantear la posibilidad y pertinencia de rastrear la diversidad del mundo que se presupone deja atrás la modernidad.

Más de todo esto se ha de concluir, no que sólo se debe aprender aritmética y geometría, sino únicamente que los que buscan el recto camino de la verdad, no deben ocuparse de ningún objeto acerca del cual no puedan tener una certeza igual a la de las demostraciones aritméticas y geométricas.²⁵

Partiendo de la una problemática unidad de las dos sustancias, para Descartes,²⁶ las pasiones son afecciones del alma que así como alteran a ésta comúnmente son en el cuerpo una acción. La fuerza o debilidad del alma se puede ver en la fortaleza que tiene la voluntad para vencer a las pasiones y detener los movimientos del cuerpo que las acompañan. Así como considera que la razón es lo mejor repartido, también plantea que son pocos los hombres tan débiles de no querer más que lo que les dicta la pasión. Recurriendo al ejemplo del adiestramiento de un perro de caza que logra reprimir el impulso de correr atrás de una perdiz o el de huir cuando oye el disparo de un fusil, señala:

Porque ya que podemos, con un poco de maña, cambiar los movimientos del cerebro en los animales desprovistos de razón, es evidente que podemos hacerlo aún mejor en los hombres; y que incluso quienes tienen el alma más débil podrían adquirir un dominio completamente absoluto sobre todas sus pasiones, si emplearan suficiente habilidad al formarlas y guiarlas.²⁷

²⁴ Cf. Remo Bodei, *Una geometría de las pasiones. Miedo, esperanza y felicidad: filosofía y uso político*, p. 11.

²⁵ René Descartes, *Reglas para la dirección del espíritu* en *Dos opúsculos*, Regla II, p. 97. En Pascal podemos ver cómo se declara la victoria de la razón incluso sobre el azar y la contingencia. Retomemos un pasaje en donde se refiere a la repartición del azar en los juegos sujetos a ella: "[...] los resultados de la suerte ambigua son atribuidos a la contingencia fortuita más que a la necesidad natural. Es por ello que la cuestión ha errado inciertamente hasta nuestros días: pero ahora, permaneciendo rebelde a la experiencia, no ha podido escapar al imperio de la razón. Y, gracias a la geometría, la hemos reducido con tal seguridad a un arte exacto que participa de su certeza y progresa audazmente. [...] se puede, extrayendo su nombre de ambas, arrogarse, en pleno derecho, el título estupefaciente: *La geometría del azar.*" (Blaise Pascal, "A la muy ilustre Academia de Ciencias Parisina 1654" en *Obras matemáticas. Selección de textos*, pp. 108-109)

²⁶ Cf. René Descartes, *Las pasiones del alma*, especialmente artículos II, XLVIII, XLIX, L.

²⁷ René Descartes, *Ibid.*, artículo L, p. 130.

De hecho, Descartes llega a plantear que la diferencia entre las almas *grandes* y las *vulgares* radica fundamentalmente en que estas últimas se dejan llevar por las pasiones, siendo dichosas o desgraciadas dependiendo de lo que les acaece; en cambio, las almas *grandes* hacen razonamientos tan fuertes que, aunque tengan también pasiones, la razón permanece siendo la dueña, logrando de esta manera que incluso las aflicciones contribuyan a su perfecta felicidad.²⁸

La antípoda nuevamente es Hume: "La razón es, y sólo debe ser, esclava de las pasiones, y no puede pretender otro oficio que el de servirías y obedecerlas."²⁹ En este conocido pasaje se pueden ver los antecedentes de uno de los principales planteamientos contemporáneos sobre la acción humana: la acción es causada por los deseos, creencias y otras inclinaciones del agente, lo cual no es incompatible con el hecho de que sea una acción libre.³⁰

Frente a la posición que considera que las pasiones deben estar bajo el control de la razón y que un hombre racional actúa con base en sus conocimientos bien fundados, Hume advierte que la razón por sí misma no es suficiente para actuar e incluso ni siquiera es el factor determinante, sólo puede dirigir los impulsos o acciones, pero no propiamente producirlos:

[...] el impulso no surge de la razón, sino que es únicamente dirigido por ella. De donde surge la aversión o inclinación hacia un objeto es de la perspectiva de dolor o placer. Y estas emociones se extienden a las causas y efectos de ese objeto, tal como nos son señaladas por la razón y la experiencia.³¹

²⁸ Cf. René Descartes, *Ibid.*, nota pp. 220-221, en donde se reproduce parte de una carta del 18 de mayo de 1645 dirigida a la princesa Isabel de Bohemia.

²⁹ David Hume, *op. cit.*, Libro II, Parte III, p. 617.

³⁰ El planteamiento contemporáneo lo desarrollaremos ampliamente en el siguiente capítulo a partir de la propuesta de Donald Davidson.

³¹ David Hume, *op. cit.*, Libro II, Parte III, p. 616.

Es a partir de un fin o propósito, establecido por una propensión o aversión, que tiene sentido el uso del razonamiento para poder llevar a cabo la mejor acción. Finalmente, las propensiones o aversiones, en tanto pasiones, son las causas de toda acción: actuamos sólo porque el acto no nos es indiferente, tenemos un deseo o una aversión que nos impulsa a actuar. Sin embargo, con ello, Hume también pone en cuestión el enfrentamiento entre pasiones y razón: se trata de dos ámbitos distintos, uno correspondiente a la existencia y a la modificación de la misma (placer-dolor), el otro al descubrimiento de la verdad o falsedad. En este sentido, las pasiones no pueden ser irracionales, simplemente son y la razón no puede oponerse a ellas. De hecho, sería impropio hablar del combate entre pasiones y razón.

Aquí es pertinente hacer una observación. Hume se plantea la posibilidad de que las pasiones pudieran oponerse a la razón, y concluye que esta oposición sólo podrían darse si las pasiones estuvieran acompañadas de un juicio. Esto puede ocurrir a su vez de dos maneras: cuando la afección se basa en la suposición de un objeto que en realidad no existe, o cuando al poner en acto la pasión no elegimos adecuadamente los medios y nos engañamos en nuestro juicio sobre las causas y efectos. En estos casos, lo irracional no recae propiamente en la pasión sino más bien en el juicio que la acompaña.³²

De acuerdo con esto, tenemos que lo único que puede oponerse a una pasión es otra pasión: en ocasiones los hombres actúan conscientemente en contra de sus propios intereses, así como también a veces reprimen una pasión violenta al perseguir sus deseos. Prevalecen las pasiones violentas, dirigidas hacia objetos particulares en algún sentido presentes, o las pasiones apacibles, como el apetito general hacia el bien y la aversión contra el mal. Por ello, plantea Hume:

³² Cf. David Hume, *Ibid*, p. 618.

Podemos observar, en general, que ambos principios actúan sobre la voluntad y que, allí donde se oponen directamente, uno de ellos prevalece, según el carácter *general* o la disposición *actual* de la persona. Lo que llamamos fuerza de espíritu implica el predominio de las pasiones apacibles sobre las violentas, aunque podemos observar fácilmente que no existe hombre alguno que posea esta virtud con tal constancia que, en alguna ocasión, no se someta a las incitaciones de la pasión y el deseo.³³

De esta manera, lejos de la visión cartesiana que parte de una oposición razón-pasiones, en la que finalmente la razón puede salir vencedora, la propuesta humeana sostiene que la razón colabora inevitablemente con las pasiones, y son más bien estas últimas las que entran en conflicto entre sí. Mas allá de las posibles objeciones que se pueden formular a la propuesta de Hume, sin duda algunos aspectos de su manera de explicar la acción de los seres humanos siguen presentes en las teorías contemporáneas. Como hemos dicho, nosotros nos ocuparemos especialmente de la explicación de la teoría de elección racional en la que se retoma, aunque en otro contexto, una de sus tesis principales: en la producción de toda acción interviene un deseo o propensión.³⁴

3. El interés, una pasión peculiar

Intereses y *deseos* son todavía hoy conceptos claves dentro de la explicación de la acción. Pero antes de entrar a las propuestas contemporáneas, quedándonos en el mismo periodo del desarrollo de la modernidad, pasaremos a un nivel menos abstracto en el que no nos preocuparemos tanto por destacar las diferencias y rupturas entre los diversos autores, sino más bien trataremos de recuperar las semejanzas y paulatinas transformaciones de algunos aspectos de sus principales propuestas

³³ David Hume, *Ibid*, pp. 621-622.

³⁴ Cf. Barry Stroud, *Hume*, especialmente capítulo VII "Acción, razón y pasión", pp. 201-239.

vinculados al problema de las pasiones. Siguiendo el atractivo trabajo de Albert O. Hirschman, podremos descubrir que "lo nuevo surgió de lo antiguo en mayor medida de lo que generalmente se cree."³⁵

El problema a tratar es el siguiente. A principios de la era cristiana, San Agustín proporcionó al pensamiento medieval una de sus directrices básicas, la cual condenaba tres impulsos o pasiones del hombre: el deseo de dinero y posesiones, el deseo de poder y el deseo sexual.³⁶ Aunque el propio San Agustín, al mismo tiempo, también aceptó la posibilidad de recurrir a un *vicio* para frenar a los otros, y por ello llegó a concederle cierto apoyo, aunque sea limitado, a la *búsqueda de la gloria*. Este apoyo se acrecentó con el ideal caballeresco (eminentemente aristocrático), al grado de que la búsqueda del honor y la gloria se convirtió en la piedra de toque de la virtud y la grandeza del hombre. Ya en el Renacimiento esta visión era dominante e incluso alcanzó a llegar hasta el siglo XVII.³⁷ Sin embargo, menos de un siglo después, se alabó el impulso adquisitivo y las actividades con él relacionadas. Generalmente, este giro de ciento ochenta grados se explica como el resultado de la victoria de una ideología (burguesa) sobre otra (aristocrática), pero se trata de un acontecimiento más complejo. Aquí sólo nos acercaremos a algunos de sus elementos básicos a partir del sugerente trabajo de reconstrucción conceptual realizado por Hirschman.

³⁵ Albert O. Hirschman, *Las pasiones y los intereses. Argumentos políticos en favor del capitalismo antes de su triunfo*, p. 12.

³⁶ Como destaca Alasdair MacIntyre, para San Agustín a través de la disciplina ascética se podía ascender en la escala de la razón y recibir la iluminación de Dios; la mente iluminada podía escoger correctamente entre los objetos del deseo, distinguiendo la *cupiditas* (el deseo de las cosas sensibles) de la *caritas* (el deseo de lo celestial), en donde finalmente esta última saldría victoriosa. (Cf. Alasdair MacIntyre, *Historia de la ética*, p. 118)

³⁷ Obviamente, también hubo posiciones en contra que de hecho llevaron a la *demolición del héroe*. Hirschman destaca: "Se demostró que todas las virtudes heroicas eran formas de la mera autoconservación (Hobbes), del amor a sí mismo (La Rochefoucauld), de la vanidad y el escape frenético del verdadero conocimiento de sí mismo (Pascal)." (Albert O. Hirschman, *op. cit.*, pp. 19-20). Sin embargo, esta crítica a los valores tradicionales no era para proponer un nuevo código moral, que correspondería a los intereses o las necesidades de una nueva clase social, sino para restaurar finalmente la posición agustiniana.

El inicio de este giro se ubica en el Renacimiento, pero no por que haya surgido una *ética* nueva para el *individuo*, sino más bien debido a un viraje en la teoría del Estado que buscaba mejorar su funcionamiento dentro del orden existente.³⁸ Podemos recordar que para Maquiavelo los filósofos morales y políticos se habían ocupado sólo de lo *imaginario* y no habían proporcionado las orientaciones para el mundo *real* en donde el príncipe debe desarrollarse. Esta exigencia de tener un enfoque *realista* pasó después del príncipe al individuo, es decir, de la naturaleza del Estado a la naturaleza humana. De hecho, en el siguiente siglo se podía tener la pretensión de descubrir las leyes de la conducta humana; en el *Leviatán*, por ejemplo, Hobbes se ocupa de la naturaleza del hombre antes de proceder a examinar la comunidad. Incluso, en el siglo XVIII se siguió sosteniendo que el tema de la ciencia política es el hombre tal y como es *realmente*.

Por otra parte, fue también en el Renacimiento cuando surgió la convicción de que ya no se podía confiar a la moral y a la religión el freno de las pasiones destructivas de los hombres.³⁹ Se necesitaban nuevas formas que fueran más eficaces para modelar el patrón de las acciones humanas, las cuales se empezaron a buscar a partir de la disección de la propia naturaleza del hombre. Frente a la dependencia del preceptor religioso, se propusieron al menos tres líneas de argumentación:

a) la más obvia, e incluso ya tradicional para ese momento, es apelar a la coerción y la represión, en este sentido, al Estado se encomienda la tarea

³⁸ Al respecto podemos recordar también las observaciones hechas por Perry Anderson en relación al Estado Absolutista: "nunca fue un árbitro entre aristocracia y burguesía ni, mucho menos, un instrumento de la naciente burguesía contra la aristocracia: fue el nuevo caparazón político de una nobleza amenazada." (Perry Anderson, *El Estado Absolutista*, p. 12)

³⁹ Obviamente, con ello no se pretende que sea una idea original del Renacimiento, sus antecedentes pueden remontarse al mundo antiguo. Sin embargo, en relación con el mundo medieval respresenta una ruptura fundamental que sienta las bases de la teorías modernas del Estado.

de contener, aun por la fuerza si es necesario, las peores manifestaciones y las consecuencias más peligrosas de las pasiones humanas;

- b) controlar las pasiones en lugar de simplemente reprimirlas, es decir, poner a trabajar las pasiones en favor del bienestar general, el encargado de esta hazaña, como medio transformador y civilizador, es el Estado o la propia sociedad; y
- c) discriminar entre las diversas pasiones y combatir el fuego con el fuego, es decir, utilizar una pasión para compensar a las otras que se consideran más destructivas.

De acuerdo con nuestro problema inicial, nos centraremos en especial en el desarrollo de la última solución, de la que Hirschman sintetiza:

la idea de conducir el progreso social enfrentando con astucia una pasión a otra se convirtió en un pasatiempo intelectual bastante común durante el siglo XVIII [...] El principio de la pasión compensadora había surgido en el siglo XVII a causa de su visión sombría de la naturaleza humana y de la creencia general de que las pasiones son peligrosas y destructivas. En el curso del siglo siguiente, la naturaleza humana y las pasiones habrían de ser rehabilitadas ampliamente [...] el remedio de la pasión compensadora continuaba invocándose aunque ahora se afirmara que las pasiones son vigorizantes y no perniciosas.⁴⁰

Esta compensación de las pasiones a partir de ellas mismas, se puede concebir como el fundamento mismo del nuevo Estado. Recordemos, por ejemplo, cuando Hamilton justifica el principio de reelección presidencial: si por buena conducta se puede esperar la continuación de honores, se vacilará en sacrificar éstos por la ganancia.

⁴⁰ Albert O. Hirschman, *Ibid.*, pp. 34-35.

Para denotar a las pasiones a las que se les asigna la función de compensar a otras, se utilizó la palabra *interés*. Sin embargo, por un lado, para que la propuesta de las pasiones compensadoras fuese operativa, se necesitaba saber, al menos en términos generales, cuáles eran las pasiones que podían ser las domadoras y cuáles eran las salvajes que debían ser domadas. Por otra parte, si bien en Hobbes la doctrina del contrato social es resultado de la estrategia de compensación, recurre a ella sólo una vez para fundar el Estado, como si se pudieran resolver de una vez por todas los problemas creados por los hombres pasionales. Pero, incluso para algunos de sus contemporáneos, esto no era posible y propusieron que la estrategia de compensación era necesaria de forma continua. De esta manera, enfatiza Hirschman,

Para este propósito resultaba claramente conveniente una formulación más general y permanente de la asignación de papeles. Tal formulación surgió en efecto y asumió la forma de una oposición de los *intereses* de los hombres a sus *pasiones*, y de un contraste entre los efectos favorables que se obtienen cuando los hombres se guían por sus intereses y la situación desastrosa que prevalece cuando los hombres dan rienda suelta a sus pasiones.⁴¹

Finalmente, entonces, la pasión peculiar que tenía la función de compensar a las otras pasiones, terminó contraponiéndose a ellas. Los *intereses* de personas y grupos acabaron por oponerse a las pasiones, al tiempo que fue adquiriendo un significado fundamental de ventaja económica. Sin embargo, hay que aclarar, tampoco fue siempre así. A finales del siglo XVI el término *interés* en tanto preocupaciones, aspiraciones y ventajas, era corriente en Europa Occidental y su significado no se limitaba sólo a los aspectos materiales del bienestar personal, por el contrario, abarcaba la *totalidad* de las aspiraciones humanas y remitía a un aspecto de reflexión y

⁴¹ Albert O. Hirschman, *Ibid.*, p. 39.

cálculo sobre la forma en que estas aspiraciones podían alcanzarse. De hecho, el análisis de la noción *interés* surgió primero en un ámbito distinto al de los individuos y su bienestar material. Como ya mencionamos, a la base de la búsqueda de un mayor realismo en los estudios del comportamiento humano, se encontraba la preocupación por mejorar la calidad de la actuación del Estado. Preocupación que llevó a la primera definición de *interés* así como a una investigación más detallada sobre el mismo.

La fuente se encuentra nuevamente en Maquiavelo. Aunque él no le asignó un nombre específico, sí describió un comportamiento característico de los gobernantes que llevó después a la consolidación de los términos inicialmente sinónimos de *interés* y *razón* de Estado y cuyo uso se generalizó en la segunda mitad del siglo XVI. Por una parte, no se trataba sólo de una declaración de independencia con respecto a los preceptos y reglas moralizantes, sino también de identificar una voluntad racional, libre de pasiones e impulsos momentáneos, que pudiese orientar de manera clara y sensata al príncipe.⁴² Aunque esto pronto se vió poco útil: si las normas tradicionales del comportamiento virtuoso eran difíciles de *alcanzar*, el *interés* resultó difícil de *definir*. Con relativa facilidad se puede afirmar que en general el interés de un rey es mantener e incrementar el poder y la riqueza de su reino, pero se trata de un principio que no proporciona ninguna *regla de decisión* precisa para las situaciones concretas. Si bien surgieron algunas propuestas, parece que la historia de los intentos de crear dichas reglas no es nada exitosa.

Sin embargo, frente al confuso horizonte que tuvo el concepto *interés* en el caso del príncipe o del Estado, fue en cambio muy fructífero cuando se pasó al ámbito de los grupos o individuos dentro del Estado. Aquí, la conjunción del

⁴² Curiosamente, la nueva doctrina del interés del príncipe llegó a criticar y prevenir contra la rendición ante las pasiones poco tiempo después de que habían sido ridiculizados, por poco realistas e inútiles, los antiguos preceptos moralizantes y religiosos. Muchos de los defensores

interés propio y la racionalidad, piedra de toque del arte de gobernar, resultó ser una categoría muy útil y enriquecedora.

El paso del interés del gobernante al de los diversos grupos gobernados siguió caminos distintos en Inglaterra y Francia. A principios del siglo XVII, Inglaterra importó de Francia e Italia el concepto *interés*. El interés que debía guiar a los príncipes se convirtió más tarde en interés nacional ya que, con la Revolución y la Guerra Civil, el interés de Inglaterra no se debatía en relación con otros países sino con los protagonistas principales de sus luchas internas. A finales de siglo, ya recuperada la estabilidad política, se discute cada vez más los intereses de grupos e individuos en términos de aspiraciones económicas.⁴³

El desarrollo del término *interés* en Francia se asemejó al inglés. A principios del siglo XVII era de uso común el término en sentido maquiavélico y se utilizaba en sentido amplio, es decir, no sólo remitía a la riqueza, sino también al poder y a la influencia del soberano. Sin embargo, poco a poco se va limitando su significado a la búsqueda de la ventaja material, lo que puede explicarse por la ya en ese entonces vieja relación del interés con el préstamo de dinero así como por la afinidad del cálculo racional (implícito en el concepto de *interés*) con la naturaleza de las actividades económicas. Asimismo, no olvidemos de que en la Francia del siglo XVII el poder se encontraba muy concentrado y el único espacio en donde una persona ordinaria podía aspirar a ascender era precisamente en el ámbito económico.

Obviamente, cuando se contrapusieron los intereses a las pasiones, se llegó a una oposición que pudo tener diversos significados dependiendo de cómo se entendieran los primeros, es decir, si se manejaban en un sentido amplio o en un sentido estrecho. Por ejemplo,

de estos últimos aprovecharon este nuevo e inesperado aliado, y se trató de demostrar, por ejemplo, cómo *el razonable amor de sí mismo* (el interés) se alía a la moral *contra* las pasiones.

⁴³ Probablemente este desplazamiento haya sido motivado por un cambio en el significado de *interés público* en donde la *abundancia* se volvió un componente fundamental.

Una máxima tal como la de "El interés no mentirá" fue originalmente una exhortación a perseguir *todas* nuestras aspiraciones de una manera ordenada y razonable; aconsejaba la introducción de un elemento de eficiencia calculadora, así como de prudencia, en el comportamiento humano, cualquiera que fuese la pasión que lo motive básicamente. Pero debido a la desviación semántica antes mencionada del término "intereses", la oposición entre intereses y pasiones pudo significar o transmitir también una idea diferente, mucho más notable en vista de los valores tradicionales, a saber: que *un conjunto de pasiones, conocidas hasta ahora como codicia, avaricia, o amor por el lucro, podía utilizarse convenientemente para enfrentar y frenar a otras pasiones tales como la ambición, el ansia de poder, o el deseo sexual.*⁴⁴

De esta manera, se llegó a conjuntar la anterior propuesta sobre las pasiones compensadoras y la nueva doctrina del interés. En el siglo XVII, al igual que en la Baja Edad Media, la avaricia era una de las peores pasiones, pero cuando el dinero tuvo la etiqueta de intereses, no sólo pasó a ser aclamada sino que incluso se le asignó el papel de frenar a las otras pasiones que por mucho tiempo se consideraron menos reprobables. Ello no fue simplemente porque se usó un nuevo término relativamente más neutro y débil, sino porque el nuevo término tenía una connotación *positiva y curativa* producto de su asociación con la idea de conducir más lúcidamente los asuntos humanos, sean éstos públicos o privados.

La oposición entre intereses y pasiones apareció por primera vez en la obra de Henri Rohan *Sobre el interés de los príncipes y estados de la cristiandad*. Esta dicotomía fue discutida ampliamente en las décadas siguientes tanto por ingleses como por franceses quienes se encargaron de aplicarla al estudio de la conducta humana en general. De hecho, la idea del interés se volvió un paradigma y la mayor parte de la acción humana se explicó de pronto por el interés propio. Si bien pasó así al siglo XVIII, se volvió una

⁴⁴ Albert O. Hirschman, *Ibid.*, p. 47.

noción tan evidente que no hubo quien se preocupara por definirla con precisión. Tampoco se explicaba su relación con las dos grandes categorías que, desde Platón, habían marcado la dirección de los estudios sobre la motivación humana, es decir, su relación con las pasiones y con la razón. Frente a este panorama, plantea Hirschman:

Pero es precisamente en el marco de esta dicotomía tradicional que puede entenderse el surgimiento de una tercera categoría a fines del siglo XVI y principios del XVII. Una vez considerada destructiva la pasión e ineficaz la razón, la concepción de que la acción humana podría describirse completamente por su atribución a la una o la otra significaba una perspectiva muy sombría para la humanidad. En consecuencia, la introducción del interés entre las dos categorías tradicionales de la motivación humana llevaba un mensaje de esperanza. En efecto, se veía al interés participando de la mejor naturaleza de cada una de aquellas categorías, como la pasión del amor a sí mismo elevada y contenida por la razón, y como la razón dotada de dirección y fuerza por esa pasión. La forma híbrida de la acción humana resultante se consideraba libre de la naturaleza destructiva de la pasión y de la ineficacia de la razón.⁴⁵

Obviamente, tampoco hubo unanimidad al respecto, algunos incluso rechazaron esta nueva concepción y, siguiendo con la visión de San Agustín, no veían mucha diferencia entre las pasiones y los intereses. Pero, en general, los críticos más bien dudaban de que el interés, entendido como amor de sí mismo razonable, pudiese tener el mismo o mayor peso que las pasiones; otros refutaban la preeminencia del interés, ya no tanto por la interferencia de las pasiones, sino por la incapacidad de los hombres para percibir sus propios intereses; otros más prevenían contra el peligro de eliminar a las pasiones y, en este sentido, concebían a los intereses y a las pasiones como elementos igualmente determinantes del comportamiento humano y no debía ser motivo de desaliento la ocasional victoria de las últimas sobre las primeras.

⁴⁵ *Ibid.*, p. 50.

Fue hasta el siglo XVIII cuando se criticó más fuertemente la idea de que el interés es predominante en la acción humana. Pero en ese momento la concepción sobre las pasiones ya no era la misma, se había modificado: si a principios del siglo XVII se consideraban totalmente viciosas y destructivas, a finales de ese siglo y aun más en el siguiente, las pasiones fueron rehabilitadas como la esencia de la vida y una fuerza potencialmente creadora. Por ello, la crítica a la preeminencia de los intereses en la conducta del hombre en tanto se relega el papel que pueden jugar también las pasiones, significa una cosa en el siglo XVII y otra muy distinta en el siglo XVIII. En el primer caso, se estaba señalando que el mundo era aun *peor* que lo propuesto, en cambio, en el siguiente siglo, se defendía que el mundo era *mejor* cuando las pasiones participan activamente.

La nueva interpretación se finca en el rechazo, dado por la Ilustración, de la concepción trágica y pesimista del hombre y la sociedad tan característica del siglo XVII. La nueva concepción, según la cual las pasiones *mejoran* un mundo gobernado sólo por el interés, aparece totalmente articulada en Hume [...] Naturalmente, una vez limitado el significado de los intereses a la ventaja material, la idea de que “el interés gobierna al mundo” debía perder inevitablemente mucho de su atractivo anterior. En realidad, la frase se convierte en un lamento [...]⁴⁶

Como dijimos, parecía que con la idea de que el interés era una motivación dominante del comportamiento humano se había descubierto una base realista para un orden social viable. Por un lado, la propuesta tenía varias ventajas específicas: la más general era la posibilidad de previsión, la cual llevaba a suponer que cada hombre al perseguir su interés beneficiaría también a los otros, ya que su curso de acción sería tan transparente como el de una persona virtuosa. De esta manera, del funcionamiento esperado *en la política* podía surgir la posibilidad de una ganancia mutua. Por otra parte, esta

⁴⁶ *Ibid.*, p. 54.

concepción también implicó graves dificultades: por ejemplo, se objetaba que justo la incapacidad de predicción es poder o que en política internacional frecuentemente los intereses de las partes son contrapuestos. Sin embargo, más allá de la política, en el campo de las actividades económicas, los beneficios que se podían obtener de la posibilidad de previsión de la conducta humana basada en el interés, eran enormes. Por ello, en este ámbito la propuesta fue finalmente mucho más exitosa.

Así pues, el subproducto de los individuos que actuaban en forma previsible de acuerdo con sus intereses económicos no era un *equilibrio* inestable, sino una *red* vigorosa de relaciones interdependientes. En consecuencia, se esperaba que la expansión del comercio interno creara comunidades más cohesivas, mientras que el comercio exterior ayudaría a evitar las guerras entre ellas.⁴⁷

De esta manera, destaca Hirschman, si para el pensamiento económico anterior a Hume y Adam Smith, el comercio era un juego de suma cero, en el siglo XVII y XVIII se esperaba de la expansión del comercio efectos (no sólo económicos sino también políticos, sociales y morales) benéficos para todos. Frente al fluctuante comportamiento apasionado y la inconstancia como una de las principales dificultades a vencer para la creación de un orden social viable, la capacidad de previsión se presentaba como una constancia y, por lo tanto, como una muy buena razón para aceptar con agrado la concepción de un mundo gobernado por el interés.

El supuesto básico es que, en contraposición al comportamiento guiado por las pasiones, los hombres son firmes, constantes y metódicos en la búsqueda de sus propios intereses. Esto explicaría también por que finalmente el término *interés* se fue restringiendo de un sentido amplio a una pasión particular, el amor al dinero, ya que se trata de una pasión que se caracteriza precisamente por la constancia y la igualdad en el tiempo y entre personas.

⁴⁷ *Ibid.*, p. 58.

Curiosamente, la insaciabilidad del deseo de ganancia, en tanto implicaba constancia, se convertía en virtud; concepción que terminó por imponerse al adjudicarle una cualidad adicional: la inocuidad.

El comportamiento que obedece al interés y a la ganancia se consideraba superior al orientado por las pasiones, en tanto estas últimas eran salvajes y peligrosas, en cambio, la búsqueda del propio interés se presentaba como algo inocente. Probablemente, frente a los ejércitos saqueadores y los piratas de la época, cobró fuerza la imagen del comerciante como un hombre pacífico e inofensivo. En el caso específico de Francia, incluso más que en Inglaterra, influyó también la forma en que se veía a los grupos sociales: quien no perteneciera a la nobleza no podía compartir las virtudes heroicas ni las pasiones violentas, sólo podía buscar sus propios intereses (actividad que se consideraba más inocua que los pasatiempos y hazañas de la aristocracia).

Obviamente, siempre hay críticas y las concepciones sobre los intereses y las pasiones no son la excepción. Se refutó que el hombre se gobierne solamente por el interés o el amor a sí mismo; se hicieron nuevas distinciones entre las pasiones para presentar algunas como menos dañinas que otras y algunas incluso como benéficas. En el siglo XVIII, principalmente en Inglaterra, la oposición entre pasiones benignas y malignas fue el equivalente de la oposición del siglo anterior entre los intereses y las pasiones. El deseo que actúa con cálculo y racionalidad equivale a lo que se entendía por interés en el siglo XVII. Retomemos nuevamente a Hume, como vimos, para él una pasión tranquila puede vencer a una pasión violenta, de hecho, este autor llega a plantear que una de las consecuencias de toda profesión industriosa es hacer que el amor por la ganancia prevalezca sobre el amor por el placer.

Con esta propuesta culmina el movimiento de ideas que aquí sólo hemos punteado. Finalmente, se acaba alabando al sistema capitalista en la medida en que activa ciertas inclinaciones humanas "benignas" en contra de otras inclinaciones que son "malignas". De esta manera, el capitalismo se presenta

como una forma de reprimir y atrofiar algunos aspectos negativos de la naturaleza humana en aras de un beneficio tanto personal como social. Sin duda, uno de los representantes más significativos de esta concepción es un querido amigo de Hume. Para concluir esta sección, recordemos el conocido pasaje en donde Adam Smith enfatiza que la necesaria ayuda que requerimos de los demás tiene como base el interés personal:

No es la benevolencia del carnicero, del cervecero o del panadero la que nos procura el alimento, sino la consideración de su propio interés. No invocamos sus sentimientos humanitarios sino su egoísmo; ni les hablamos de nuestras necesidades, sino de sus ventajas.⁴⁸

Asimismo, no olvidemos su optimista convicción de que al buscar su propio interés cada individuo, sin saberlo ni proponérselo, trabaja también para la sociedad en su conjunto:

[...] sólo piensa en su ganancia propia; pero en éste como en otros muchos casos, es conducido por una mano invisible a promover un fin que no entraba en sus intenciones. Mas no implica mal alguno para la sociedad que tal fin no entre a formar parte de sus propósitos, pues al perseguir su propio interés, promueve el de la sociedad de una manera más efectiva que si esto entrara en sus designios.⁴⁹

En esta última sección, recuperamos el brillante trabajo de reconstrucción histórica que realizó Hirschman, en donde se puede observar que una de las principales contraposiciones que han acompañado a la historia de la filosofía, es mucho más compleja y rica de lo esperado. Las pasiones y la razón no necesariamente conforman una dicotomía; aunque se puedan enfrentar también se pueden acompañar, yuxtaponer y recomponer. En el caso

⁴⁸ Adam Smith, *Investigación sobre la naturaleza y causas de la riqueza de las naciones*, p. 17.

⁴⁹ Adam Smith, *Ibid.*, p. 402

de la historia moderna del término *interés*, destacan los esfuerzos teóricos por construir un enlace que al vincular estos distintos ámbitos tampoco disuelva sus diferencias. Sin pretender ser exhaustivos, ésta ha sido sólo una breve introducción para adentrarnos a la problemática contemporánea en donde, como dijimos, el rechazo a las formas rígidas de racionalidad tampoco conlleve la disolución de todo principio o compromiso. En especial, para el caso de la teoría de la elección racional, habrá que reconocer no sólo sus propios límites y los límites de la razón, sino también sus vínculos con los deseos y sobre todo con las emociones. Tal vez, como ha sugerido Remo Bodei, nos encontraremos con que

las pasiones se revelarían como formas simbólicas y expresivas transmitidas y -a su modo- refinadas por tradiciones específicas y no, por el contrario, como simples pulsiones naturales, primitivas e inmutables. [y] No habría, por consiguiente, ninguna necesidad de que fueran solamente reprimidas o elevadas al sublime reino de la una razón majestuosa o misericordiosa [...]⁵⁰

⁵⁰ Remo Bodei, *Una geometría de las pasiones*, 1995, p. 63.

II. LA TEORÍA DE ELECCIÓN RACIONAL

El individualismo metodológico, como su nombre lo indica, explica algún problema o fenómeno social a partir de los individuos, en especial a partir de ciertas hipótesis sobre su conducta. Se pueden distinguir diversos individualismos metodológicos dependiendo de cuáles sean las unidades identificadas (un individuo físico, un individuo jurídico, una institución colegiada, etc.) y cuáles sean las hipótesis sobre sus respectivos comportamientos (egoísta, altruista, cooperativo, etc.).¹ Thomas Hobbes fue el primero en articular con claridad una explicación de este tipo.² Como sabemos, su propuesta metodológica fue retomada, si bien con grandes divergencias de contenido, por la mayoría de los pensadores ilustrados (incluso se ha llegado a hablar de un *modelo hobbesiano*)³ y, al menos desde principios del siglo XIX, fue también fuertemente combatida. Aunque el debate entre sus defensores y sus críticos reaparece bajo diversas formas a lo largo de los dos últimos siglos, sin duda el individualismo sigue siendo uno de los principales métodos de las ciencias sociales.⁴

Por lo que se refiere en especial a *las hipótesis sobre el comportamiento* de los individuos, hay que destacar el consenso que, más o menos generalizado, existe en la teoría económica moderna y que conforma el

¹ Cf. Steven Lukes, *El individualismo*, cap. XVII "El individualismo metodológico", pp. 137-154.

² Aunque ya en Nicolás Maquiavelo se encuentran tesis fuertemente individualistas: la política es concebida fundamentalmente como una serie de estrategias que el Príncipe debe conocer y llevar a la práctica considerando tanto el riesgo como la incertidumbre (Cf. Paulette Dieterlen, "El individualismo metodológico" en *Sociológica* No. 14, pp. 273-289).

³ Modelo que, de Hobbes a Hegel (incluido-excluido), compartiría el iusnaturalismo moderno en contraposición al pensamiento tradicional aristotélico (Cf. Norberto Bobbio, "El modelo iusnaturalista" en *Sociedad y Estado en la filosofía política moderna*, pp. 15-145).

⁴ Para frenar algunas de las posibles y frecuentes objeciones a esta teoría, se ha insistido en que la reducción de la que parte no asume las siguientes implicaciones: que los individuos son necesariamente egoísta, que son "átomos" con una existencia presocial, o que la teoría se extiende a lo que ocurre dentro de sus mentes. (Cf. Jon Elster, *Una introducción a Karl Marx*, pp. 24-27)

llamado *paradigma de la elección racional*. El nacimiento de la teoría económica moderna se ubica alrededor de 1870 cuando las preocupaciones fundamentales ya no son las cuestiones macroeconómicas del crecimiento y la distribución sino los problemas microeconómicos de la toma de decisiones. Si bien la economía no se agota en la microeconomía, aquí nos interesa ubicar cómo, con ayuda de las matemáticas (de las llamadas técnicas *marginalistas*), la economía moderna pudo empezar a formular con precisión los costos y beneficios asociados a los usos alternativos de los recursos escasos.⁵ Uno de sus presupuestos básicos es que cada factor productivo (trabajo, maquinaria, materias primas, etc.) tiene una producción adicional (una productividad marginal) decreciente, aunque positiva. Un problema típico es la pregunta por el número de obreros que debe contratar una empresa antes de que empiece a ser indiferente la contratación de un obrero más.

Entre los conceptos claves de la economía moderna sobresale la concepción del hombre económico como un ser racional. Esto se ha interpretado de diversas maneras, para algunos (los más ortodoxos) conforma un modelo descriptivo, para otros proporciona una guía normativa, y para algunos más es simplemente un recurso hipotético, un agente idealmente racional, que se construye con fines explicativos. Para esta última posición, así como los físicos pueden especular sobre los vacíos perfectos o los gases ideales, los economistas también hacen abstracción de muchas de las complicaciones de la vida real con el fin de construir una teoría comprensiva y con poder explicativo. Sin embargo, más allá de esta controversia (que como veremos se extiende del ámbito de la economía a otras ciencias sociales), en el paradigma de elección racional, el sujeto económico es descrito por un conjunto de creencias y por un conjunto de deseos o preferencias, y dispone de un conjunto de estrategias posibles dadas determinadas restricciones. Aquí, decidir *racionalmente* es resolver un problema de maximización, es decir, elegir

⁵ Cf., *Ibid.*, pp. 28-34.

la estrategia que *mejor* satisface los deseos del agente dadas sus creencias y restricciones. Como veremos, de acuerdo con este paradigma, explicar es demostrar que con su comportamiento los individuos maximizan su función de utilidad.

Esta propuesta se enriqueció y amplió con la teoría de juegos (la cual también llegó a plantear sus límites). Los inicios de la teoría de juegos se ubican en el año de 1928 cuando John von Neumann presenta una ponencia en la Sociedad Matemática de Göttingen, aunque es hasta 1944 que sale a la luz su célebre tratado *Theory of Games and Economic Behavior* que escribe junto con Oskar Morgenstern y que llevó posteriormente al desarrollo de diversos aspectos teóricos sobre tácticas y estrategias. La teoría de juegos se preocupa por establecer los principios que deben guiar la toma de decisiones en determinados contextos que corresponden a una gran diversidad de situaciones (no sólo económicas). El problema es que nuestros actos pueden ser interdependientes en formas muy complejas, por lo que hay que establecer antes y de forma precisa estas interdependencias.⁶

Aunque los juegos acompañan a toda sociedad, incluso a las más "primitivas" (lo que ha justificado la expresión *Homo ludens*), los antecedentes de la teoría de juegos apenas surgen en el siglo XVI cuando se acuña una de sus nociones fundamentales: la probabilidad.⁷ Pero esto era sólo en juegos de azar. Fue hasta el siglo XVIII que Leibniz plantea el problema de juegos en los que se combina el azar y la habilidad, e incluso preve la posibilidad de la simulación de situaciones de la vida real (por ejemplo, los problemas navales

⁶ Cf. Leonid Hurwicz, "The Theory of Economic Behavior", S. Vajda, "Theory of Games" y Abraham Kaplan, "Sociology Learns the Language of Mathematics" en *The World of Mathematics*, ed. J. R. Newman, pp. 1267-1313.

⁷ De hecho, Morgenstern se sorprende de este relativo atraso: "Dado que los juegos han ocupado siempre al hombre en un sentido muy real, es curioso que haya transcurrido tanto tiempo antes de que los juegos se convirtieran en un tema de la investigación científica" (Oskar Morgenstern, "Teoría de los juegos" en *Metodología y crítica económica*, selecc. de Camilo Dagum, p. 159.)

podían representarse y estudiarse en tableros de maniobras).⁸ Sin embargo, el tratamiento matemático de un juego que no es estrictamente de azar (el solitario) data apenas de mediados del siglo pasado; el siguiente registro salta hasta después de la primera década del siglo XX cuando se trató, con ayuda de las matemáticas, de hacer avanzar el conocimiento sobre el ajedrez. Básicamente es hasta el trabajo de von Neumann y Morgenstern que se establece propiamente este campo de estudio al probar un teorema general que garantiza que siempre hay una estrategia óptima disponible para un jugador: la regla *mini-max* (que veremos más adelante).

A diferencia de los inicios del paradigma de elección racional, con la teoría de juegos el agente ya no está solo sino que se enfrenta a otros que, como él, son *racionales* y tienen una capacidad de previsión estratégica no controlable por los demás. En este caso, la decisión de cada uno, al igual que su recompensa, depende de la decisión de todos. Obviamente, a la par que el individualismo, la teoría de elección racional también ha recibido fuertes críticas, sin embargo, a pesar de ellas, se trata de una de las concepciones teóricas que han marcado el desarrollo de las ciencias sociales contemporáneas.

En este capítulo ubicaremos los principales elementos y problemas de dicha teoría. La intención del mismo es ofrecer un panorama general que sirva como piso para el desarrollo de los siguientes capítulos. Para llevar a cabo esta ubicación, contamos con una amplia y detallada exposición elaborada por uno de sus más brillantes defensores que, como ya señalamos, es al mismo tiempo uno de sus más incisivos críticos: Jon Elster.⁹ Aquí también hemos estructuramos el capítulo en tres secciones:

⁸ Esto no excluye obviamente que con anterioridad se hubieran visto algunas semejanzas entre los juegos y la vida real, como es el caso del ajedrez y del go, pero con ello no se trataba de establecer ninguna teoría rigurosa.

⁹ A lo largo de casi toda su obra, Jon Elster expone reiteradamente los principales conceptos de la teoría de la elección racional así como algunas de sus diversas limitaciones. Para la exposición general de este capítulo cf. especialmente Jon Elster, "Introduction" a *Rational*

1. En la primera, presentamos los conceptos básicos de la teoría de la elección racional: la racionalidad optimizadora, la racionalidad paramétrica y la racionalidad estratégica. La intención de esta sección es destacar la complejidad que puede llegar a presuponer la racionalidad instrumental, e incluso cómo dentro de sus propios términos y desde su nivel más básico es ya problemática.
2. En la segunda sección, nos ocuparemos de la estructura de la explicación de la elección racional, en donde se establecen las condiciones que deben cumplir las relaciones entre sus elementos básicos: la acción a explicar, las creencias y los deseos del agente. Como veremos, con estas condiciones se va conformando el modelo estándar de la teoría de elección racional.
3. En la tercera y última sección, se presenta un esbozo de los principales problemas de la teoría de elección racional, algunos de los cuales serán retomados en los siguientes capítulos.

1. Conceptos básicos

En primera instancia, la teoría de elección racional (TER) es una teoría normativa que proporciona *imperativos condicionales* sobre los medios más adecuados para alcanzar lo mejor posible determinados fines.¹⁰ Más allá de las

Choice, ed. Jon Elster, pp. 1-33; y "Explicación intencional" en *El cambio tecnológico. Investigaciones sobre la racionalidad y la transformación social*, pp. 65-81. Asimismo, nos apoyaremos principalmente en Shaun P. Hargreaves y Yanis Varoufakis, *Games Theory: A Critical Introduction*, y Michael D. Resnik, *Elecciones. Una introducción a la teoría de la decisión*.

¹⁰ En sus inicios, y aun hoy para los ortodoxos, la TER es considerada una teoría meramente descriptiva. En este caso, la racionalidad de la elección de los medios para alcanzar un fin descansa en su correspondencia con un patrón de conducta consistente. De esta manera, la teoría puede pronosticar acciones pero no puede explicarlas (establecer sus causas). A pesar de que la concepción normativa comparta algunas de sus formas de consistencia, su estructura

posibles motivaciones de una acción (las cuales obviamente no se pueden reducir a los intereses económicos), la *racionalidad* es instrumental, se ocupa de los resultados: "Si se desea lograr Y se debe hacer X". Esta teoría puede ser complementada con una teoría de la creencia racional ya que, en cada situación, se debe saber cuáles son las creencias que tiene el agente sobre los respectivos hechos relevantes. Como veremos, la *racionalidad* de un acto depende de su relación con los fines y las creencias, y la *racionalidad* de una creencia de su relación con las evidencias disponibles. Por ello, es conveniente desde el inicio hacer explícitas dos precisiones: por un lado, la teoría inicialmente no se ocupa de las metas que tiene el agente sino de la elección de los medios para alcanzarlas; por otro, la situación tiene un carácter eminentemente subjetivo, porque no basta con que las opciones estén objetivamente disponibles, o que ciertas opciones conduzcan objetivamente a determinados resultados, en ambos casos es necesario que el agente tenga razones para creerlo.

Pero la TER no sólo es normativa también puede emplearse para fines explicativos. En una explicación de este tipo, cuando se sostiene que un acto fue realizado *porque* es racional, se deben demostrar dos cosas: en primer lugar, que la acción es racional, es decir, que el agente hizo *intencionalmente* (no de manera accidental o por coincidencia) *lo mejor* para él mismo; en segundo lugar, que hay una relación *correcta* entre deseos, creencias y evidencias.¹¹ Vayamos por partes.

Los elementos básicos de la explicación son tres:

se considera *demasiado pequeña para ser realmente útil*. Por un lado, se llega a reconocer que los animales se comportan no sólo de manera no-automática y no-miope, sino incluso racionalmente de acuerdo con un patrón de consistencia estándar en microeconomía: la curva de demanda con pendiente negativa. Por otro lado, se presupone un comportamiento racional a unidades no-individuales.

¹¹ Esto excluye de la explicación de la TER, además obviamente de las acciones no-intencionales y las no-racionales, a los agentes no-individuales ya que, para el individualismo metodológico, no existen deseos ni creencias colectivas.

- a) el conjunto de acciones factibles, es decir, todas las acciones racionales que satisfacen las restricciones lógicas, físicas, económicas y/o mentales de la situación;
- b) el conjunto de creencias racionales que estructuran causalmente la situación y ligan cada acción con un resultado determinado; y,
- c) el conjunto de preferencias viables, es decir, la jerarquización subjetiva de los resultados esperados de cada acción.

Como dijimos, la situación de elección racional tiene un carácter eminentemente subjetivo ya que el agente debe creer que tiene ciertas opciones y que éstas conducen a determinados resultados. Se trata de una acción intencional que se caracteriza por su relación con el *futuro*. La meta que guía la acción está ausente, todavía no se realiza, solamente es imaginada o representada. A diferencia de la selección natural -en donde la "elección" sólo puede ser entre alternativas *reales*, dadas-, aquí se puede elegir entre posibles aun *no-realizados*. En este sentido, una de las características fundamentales del ser humano es su capacidad para postergar la gratificación, es decir, el hombre puede esperar (rechazar en un momento opciones favorables para acceder más adelante a opciones aun más favorables) o incluso utilizar estrategias indirectas (aceptar en un momento opciones desfavorables para tener acceso más adelante a opciones mucho más favorables).¹² De hecho, los

¹² Frente a la *maximización local* que caracteriza a la selección natural, la elección racional se caracteriza por la *maximización global*. Aunque hay que agregar que los animales algunas veces también son capaces de una conducta maximizadora global, pero esto lo hacen en contextos muy estereotipados que se pueden explicar por la propia selección natural, como es el caso de la presa y el depredador: la presa es capaz de esperar (el huir para reducir un peligro actual, puede aumentar el peligro posterior al indicarle su presencia al predador) y los depredadores son capaces de esperar y utilizar estrategias indirectas (acercarse poco a poco a su presa para no alertarla o desviarse para interceptarla en su huida). Frente a las capacidades programadas para esperar y utilizar estrategias indirectas, el ser humano se caracteriza por su capacidad *generalizada* de maximización global aun en situaciones cualitativamente nuevas (aunque estas últimas como veremos pueden presentar también problemas). Cf. Jon Elster, *Ulysses and the Sirens. Studies in rationality and irrationality*, especialmente el capítulo I "Perfect Rationality: Beyond Gradient-Climbing", pp. 1-35.

ejemplos por excelencia de una elección racional se caracterizan por la asignación de consecuencias negativas a corto plazo pero positivas a largo plazo. Pensemos por ejemplo en alguna elección dentro de la medicina, ya sea la simple inyección de alguna vacuna o, más drástico, la extracción oportuna de un tumor.

Actuar racionalmente, como dijimos, es elegir la acción que se cree llevará al *mejor* resultado posible.¹³ Esto implica a) la identificación de una meta en forma de preferencia establecida (posiblemente enunciada en forma numérica) y b) el control de las variables que determinan la realización de dicha meta. Dependiendo del tipo de situación en la que se dé la elección, y que corresponde básicamente al segundo inciso, es decir, al tipo de control que se tenga sobre las variables (total o parcial), esta *racionalidad optimizadora* puede ser *paramétrica* o *estratégica*. (Fig. 1)

En la racionalidad paramétrica, el agente cree que el medio en el que actúa está conformado por objetos y otros agentes que obedecen a leyes causales. En otras palabras, supone (quizás erróneamente) que el medio es paramétrico -es decir, constante, dado-, y se considera a sí mismo como la única variable.¹⁴ En un medio de este tipo (aunque también en el de la racionalidad estratégica), pueden darse dos casos: cuando el agente tiene una información completa sobre los resultados de los diversos cursos de acción posible y cuando dicha información es incompleta. En el primer caso, el agente puede deducir sin ambigüedad el mejor curso de acción a seguir; como por

¹³ Sin embargo, a pesar de que el caso paradigmático de la explicación de elección racional es en términos de optimalidad, ésta no es sinónimo de racionalidad. Como veremos más adelante, hay situaciones en donde la racionalidad es *satisfaciente*, es decir, se elige una alternativa *lo suficientemente buena* y no la *mejor*. Elster destaca que incluso en el ámbito de la economía, muchos especialistas reconocen a la teoría de la satisfacción como una importante alternativa teórica frente a la optimización, aunque se trate de una teoría que todavía no está desarrollada como para servir de base al trabajo empírico. (cf. nota 18 de la introducción)

¹⁴ En los casos de racionalidad paramétrica, el agente también puede considerar que la conducta de los otros no obedece a leyes causales, sin embargo son siempre considerados menos sofisticados que él. Por ejemplo, si el agente considera que los demás tienen la capacidad de adaptarse al medio, se considera a sí mismo como el único que se adapta a la adaptación de los demás.

ejemplo cuando un ingeniero calcula la trayectoria más corta para que un cohete llegue a la luna. En el segundo caso, cuando la información es incompleta, la situación se puede caracterizar a su vez por dos estados: el riesgo o la incertidumbre. En una situación de riesgo, si bien la información es incompleta sí es cuantificable, es decir, el agente puede asignar probabilidades a los resultados de las distintas acciones posibles y, con base en ello, elegir la opción que maximice la utilidad esperada.¹⁵ Pensemos por ejemplo en un campesino que debe elegir entre dos cultivos para poder cosechar al siguiente año; su decisión debe basarse en las probabilidades conocidas para cada clase de clima y las propiedades conocidas de cada cultivo en cada clase de clima. Supongamos que también son dos los estados posibles del clima, los cuales tienen la misma posibilidad de ocurrir, y que los correspondientes ingresos del campesino son los siguientes:

	CLIMA 1	CLIMA 2
CULTIVO 1	30	25
CULTIVO 2	50	15

El cultivo 2 tiene la más alta expectativa, sin embargo no necesariamente será escogido. Si el campesino es adverso al riesgo, preferirá asegurar los 25 del cultivo 2, que arriesgarse a ganar 50 y quedarse finalmente sólo con 15.

En una situación bajo incertidumbre, si bien el agente puede conocer los posibles resultados de las diversas acciones, no puede asignarles ningún rango de probabilidades; por ello, aunque se puedan eliminar las opciones que llevarían a las peores consecuencias, la elección queda normalmente sin

¹⁵ Como veremos enseguida, si el agente es adverso al riesgo y busca lo seguro, no necesariamente escoge dicha opción.

resolver.¹⁶ Las mejores y las peores consecuencias son extremos que pueden combinarse de diversos modos, sin embargo, cuando se trata de una situación en donde las consecuencias pueden ser de grandes mejoras o grandes peligros, un agente que juegue a lo seguro optará por un criterio *maxi-min*: comparar las *peores consecuencias* y escoger la opción cuya *peor consecuencia* es mejor que la *peor consecuencia* de las otras opciones. En otras palabras, maximizará las utilidades mínimas. En este caso, el ejemplo por excelencia es la apuesta de Pascal que podemos representar de la siguiente manera:

	DIOS ES	DIOS NO-ES
CREO	salvación	vida sin goce
NO CREO	condena	vida con goce

Como el castigo por no creer es infinito, Pascal puede afirmar: "Tendría mucho más miedo de engañarme y encontrar que la religión cristiana es verdadera, que engañarme creyéndola verdadera".¹⁷

Aquí vale la pena hacer dos observaciones. Por un lado, la regla *maxi-min* representa un acercamiento conservador a las decisiones bajo riesgo o incertidumbre, pues se selecciona entre todas las acciones posibles la menos perjudicial. Esto puede ser lo más razonable si las pérdidas potenciales son

¹⁶ La situación bajo incertidumbre puede ser todavía más extrema: el agente, además de no poder asignar probabilidades, ni siquiera puede especificar el conjunto completo de los posibles estados del mundo, como es por ejemplo el caso de elegir entre la energía nuclear y otras formas de producción de energía.

¹⁷ B. Pascal, *Pensieri*, No. 241, p. 167. El problema de Pascal está constituido por dos posibles estados del mundo: Dios es o no es; y dos posibles cursos de acción: actuar con completa indiferencia hacia Dios o actuar de tal manera que, a su debido tiempo, llegaremos a creer en su existencia y preceptos (la elección no puede ser propiamente entre creer o no creer, es más bien entre dos tipos de acciones que, por "contagio", nos convertirán en creyentes o no creyentes). Cf. Ian Hacking, *El surgimiento de la probabilidad. Un estudio filosófico de las ideas tempranas acerca de la probabilidad, la inducción y la inferencia estadística*, especialmente el capítulo 8 "La gran decisión", 84-95.

considerables (pensemos en el caso extremo del infierno eterno en la apuesta de Pascal). Sin embargo, en una situación en donde las posibles ganancias son enormemente superiores a las posibles pérdidas, la regla *maxi-min* impediría que el agente aproveche esas oportunidades. Volvamos al caso del campesino, pero ahora con los siguientes valores:

	CLIMA 1	CLIMA 2
CULTIVO 1	30	25
CULTIVO 2	50	20

En esta situación, si el campesino también elige el cultivo 1 de acuerdo con la regla *maxi-min*, estaría rechazando la oportunidad de obtener 50 para evitar una pérdida de 5. En estos casos, en donde las ganancias son significativamente mayores (100%) que las posibles pérdidas (20%) lo recomendable sería seguir la regla *mini-max*, llamada también *regla del arrepentimiento*: el *arrepentimiento* es la cantidad de oportunidad perdida con la posible acción en cada posible estado. La tabla de arrepentimiento del caso anterior sería la siguiente:

	CLIMA 1	CLIMA 2
CULTIVO 1	20	0
CULTIVO 2	0	5

Aquí, el *arrepentimiento* máximo se ubica en el cultivo 1. La regla *mini-max* establece que hay que elegir el acto cuyo arrepentimiento máximo sea el mínimo, en este caso, el cultivo 2.¹⁸

Por otra parte, la distinción entre elecciones bajo certeza, riesgo o incertidumbre puede ser y ha sido fuertemente cuestionada. En un extremo se encuentran las posiciones que sostienen que sólo en el ámbito de la matemática y de la lógica hay absoluta certeza; en sentido estricto, sólo cuando ya se han realizado los eventos podemos asignarles con propiedad la probabilidad de 1. De esta manera, las decisiones consideradas bajo certeza serían más bien bajo riesgo. Recordemos por ejemplo a Hume cuando distingue los objetos de la razón en *relaciones de ideas* y *cuestiones de hecho*. A diferencia de las verdades demostradas por Euclides que conservarán siempre su certeza y evidencia en tanto expresan una relación que es intuitiva o demostrativamente cierta, los razonamientos sobre las cuestiones de hecho se fundan en la relación causa-efecto. Con base en su crítica al principio de causalidad, puede afirmar:

Aunque no hubiera azar en este mundo, nuestra ignorancia de la causa real de un suceso tendría la misma influencia sobre el entendimiento y engendraría un tipo de creencia u opinión similar.¹⁹

¹⁸ Aquí sólo mencionaremos estas reglas, aunque la racionalidad paramétrica puede ser mucho más compleja conforme se presentan los inconvenientes que surgen de cada una de ellas. Por ejemplo, si al uso de cada regla se le asigna un grado de optimismo o pesimismo, en cuyos extremos se encontraría la regla maxi-max con un agente optimista-desenfrenado, y en el otro la regla maxi-min con un pesimista-conservador, parecería necesario entonces encontrar una solución intermedia, el llamado *índice de optimismo*, para ponderarlo en la toma de decisiones.

¹⁹ David Hume, *Investigación sobre el entendimiento humano*, p. 73. Actualmente, dentro del ámbito de la lógica, también se trabaja el problema de la falta de seguridad en las creencias a partir de la estructura de la probabilidad: la falta de seguridad de una proposición es uno menos su probabilidad. (Cf. Dorothy Edgington, "The Logic of Uncertainty" en *Crítica, Revista Hispanoamericana de Filosofía*, vol. XXVII, No. 81, pp. 27-54)

En el otro extremo están los que sostienen que no puede haber una completa ignorancia sobre las probabilidades que corresponden a los resultados producidos por un acto. De esta manera, las decisiones consideradas bajo incertidumbre serían también bajo riesgo. Este es el caso de la *teoría subjetiva de la probabilidad* que propuso Frank Ramsey y que se base justamente en la relación de las creencias con los deseos y la acción de un agente: las acciones pueden verse como una apuesta que está vinculada al grado de creencia de los enunciados pertinentes al caso. Así como se apuesta al resultado de algún acontecimiento, ya sea en una ruleta o una carrera de caballos, también se puede apostar a la verdad de ciertos enunciados. Finalmente, el agente al actuar realiza una serie de apuestas sobre un conjunto de enunciados y sus posibles relaciones.²⁰

Con estos señalamientos esbozamos algunos de los principales elementos conceptuales que entran en juego en la definición de la racionalidad paramétrica, y vemos que desde su nivel básico la racionalidad debe enfrentarse a los problemas que surgen de la falta de una información completa, como es el caso de la mayoría de las situaciones, ya sea de la vida personal o social, en donde debemos tomar alguna decisión. En otras palabras, en su afán por establecer cuál es la *mejor* acción dadas ciertas creencias y deseos, la racionalidad frecuentemente se encuentra en situaciones bajo riesgo o incertidumbre.

Ahora podemos pasar a ver las características peculiares de la racionalidad estratégica que es aun más compleja en tanto presupone los elementos ya mencionados. A diferencia de la racionalidad paramétrica, en la racionalidad estratégica, el agente no es la única variable ya que interactúa

²⁰ La propuesta de Ramsey data de 1926, pero adquirió un gran reconocimiento después de 1954 a raíz de la publicación del libro *The Foundations of Statistics* de L. J. Savage. Actualmente, esta propuesta es conocida como la teoría *bayesiana*, en honor al pensador del siglo XVIII, Thomas Bayes, cuya contribución fundamental se volvió dos siglos después el núcleo de una de las principales teorías del siglo XX: a partir de unos datos iniciales es posible

con otros agentes que son considerados también racionales. Aquí, las restricciones a las que se enfrenta cada agente no están dadas sino que son interdependientes en la medida en que se conforman por sus propias decisiones. Como vimos, cuando las restricciones son externas o paramétricas, normalmente se pueden evaluar y después decidir qué hacer. En cambio, cuando las restricciones dependen de las decisiones del agente, éste debe tomar la decisión estratégicamente, es decir, anticipando qué harán los otros agentes, los que, a su vez, también deciden a partir de lo que creen que harán los demás.²¹

Las decisiones estratégicas, en donde el agente no tiene un control total sobre las variables que determinan la realización de su meta, son el centro del estudio de la teoría de juegos. Con más de cinco décadas de desarrollo, la estructura de esta teoría se ha vuelto cada vez más compleja aunque nunca podrá lograr formalizar completamente las interdependencias sociales.²² Aquí no pretenderemos esbozar ni siquiera un panorama general de la misma, sólo nos limitaremos a delinear algunas de sus caracterizaciones básicas. La teoría de juegos analiza las decisiones estratégicas a partir de los siguientes cuatro elementos:

asignar probabilidades, las cuales se pueden modificar si se incorpora más información. (Cf. Ian Hacking, *op. cit.*, cap. 9 "El arte de pensar", pp. 96-108)

²¹ Este proceso llevaría a una regresión infinita si no fuera por el concepto *punto de equilibrio*: cuando cada agente no puede mejorar su situación mientras los otros no cambien sus elecciones. Este grupo de elecciones es el óptimo (llamado *óptimo de Pareto* en reconocimiento a quien lo formuló) frente a cualquier otro posible grupo, aunque, como veremos, no siempre se da. Básicamente, la teoría de juegos que analiza este tipo de racionalidad nos permite entender las interacciones sociales a partir de cuatro conjuntos de interdependencias: 1) la recompensa de cada uno depende de la elección de todos, en virtud de la causalidad social general; 2) la recompensa de cada uno depende de la recompensa de todos, ya sea por egoísmo o altruismo; 3) la decisión de cada uno depende de la decisión de todos, que es la contribución específica de la teoría de juegos; y 4) la estructura de preferencias de cada uno depende de las acciones de todos, mediante la socialización y otros mecanismo similares.

²² Como la formalización o es incompleta o es contradictoria, se trata de formalizar en la mayor medida posible y en donde falle se busca complementarla con nuevas decisiones formales. Por ello, para Morgenstern, toda teoría social debe ser dinámica pasando de un formalismo a otro. En este sentido, siempre se pueden inventar nuevos juegos que sirvan como prototipos de nuevos arreglos sociales. (Cf. Oskar Morgenstern, "Teoría de los juegos", p. 182).

- a) el grupo de agentes,
- b) las elecciones o estrategias posibles para cada uno de los agentes,
- c) el estado del mundo producido por la elección de los agentes, y
- d) las preferencias sobre los estados del mundo de cada uno de los agentes.

Fundamentalmente, los juegos son estudiados a partir de la distinción entre juegos cooperativos y juegos no-cooperativos así como por la diferencia entre juegos de suma-cero (constante) y juegos de suma-variable (no-cero).²³ En relación con la primera distinción, la teoría de juegos cooperativos -aunque puede ser muy útil para los análisis normativos-, tiene menos poder explicativo ya que supone que diversos grupos de agentes pueden actuar coordinadamente contra otros grupos, pero no investiga ni las posibilidades ni las condiciones para que se dé dicha cooperación.²⁴ En cambio, la teoría de juegos no-cooperativos es una teoría más satisfactoria ya que sólo postula decisiones racionales individuales. En relación con la segunda distinción, si bien en todo juego la ganancia implica un costo que debe ser sopesado, en el caso de los juegos de suma-cero, en donde el total a ser repartido es fijo, la ganancia de un agente significa la pérdida para los otros y, por ello, se caracterizan por ser juegos de puro conflicto, pensemos en el ajedrez o en un

²³ En la teoría de juegos hay al menos una distinción previa fundamental entre juegos de azar y juegos de estrategia. En los primeros los resultados son independientes de la acción del individuo (como es el caso de la ruleta) y en los segundos el resultados no está controlado sólo por el azar (el cual puede intervenir o no como es el caso del póquer o del ajedrez), ni por el jugador individual, sino por cada jugador hasta cierto punto.

²⁴ De hecho, con el fin de darle fundamento a la teoría de juegos cooperativos, se ha recurrido a la teoría de juegos no-cooperativos en la que se puede demostrar que la decisión de cooperar puede ser un movimiento dentro de un juego no-cooperativo; aunque también se ha recurrido a otros mecanismos (no estratégicos) que llevan a la solución de un juego cooperativo a través de algún tipo de *mano invisible*. Esto último implica suponer que la conducta cooperativa se producirá porque es colectivamente óptima, lo cual es caer en las *redes* del funcionalismo fuertemente criticado por Elster. Para este autor, en las ciencias sociales el modelo explicativo pertinente es el intencional, así como el causal corresponde a la física y el funcional a la biología (Cf. Jon Elster, "Explicación causal" y "Explicación funcional" en *El cambio tecnológico*,

duelo. En cambio, en el caso de los juegos de suma-variable, el total a ser repartido depende de las estrategias elegidas por los propios agentes, por ello pueden ser juegos de pura cooperación o de conflicto y cooperación mixtos, como son la mayoría de nuestras relaciones en la familia o en el ámbito laboral.²⁵

A todas estas especificaciones hay que agregar todavía un elemento más: el tiempo. Especialmente en las situaciones estratégicas, el tiempo puede modificar qué se va a considerar como lo racional. Tomemos como ejemplo uno de los juegos más conocidos: el *dilema del prisionero*. Dos prisioneros se encuentran en celdas separadas y a cada uno se le plantean las siguientes alternativas: si denuncia al otro y éste no lo denuncia será puesto en libertad; si ambos se denuncian mutuamente, a cada uno se le impondrán tres años de cárcel; si no denuncia al otro pero éste sí a él, su pena será de cinco años; y, finalmente, si ninguno de los dos denuncia al otro, sólo tendrán un año de cárcel cada uno. Así, en este juego de suma-variable, se conforma la siguiente tabla:

	Prisionero 2	DENUNCIA	NO-DENUNCIA
Prisionero 1			
DENUNCIA	3 años de cárcel	3 años de cárcel	libertad
NO-DENUNCIA	5 años de cárcel	libertad	1 año de cárcel

pp. 28-64 ; y Paulette Dieterlen, "La crítica de Jon Elster al marxismo funcionalista" en *Marxismo analítico: explicaciones funcionales e intenciones*, pp. 141-164).

²⁵ Los juegos de conflicto y cooperación mixtos conforman el tópico central de la teoría de juegos. En general, entre los juegos más referidos y estudiados sobresalen *El dilema del prisionero*, *La gallina* y *Los seguros*. En su momento recurriremos a cada uno de ellos.

En tal situación, a pesar de que lo mejor para ambos sería que ninguno de los dos denuncie al otro (si cooperan entre ellos sólo tendrían un año de cárcel), para cada uno la estrategia dominante, es decir, la mejor respuesta frente a cualquier decisión que tome el otro, es denunciarlo (no-cooperar): si el otro no lo denuncia, sale libre; y en el caso de que lo denuncie, sólo tendrá tres años de cárcel (frente a los cinco de no haberlo denunciado él también). Obviamente, como el otro prisionero hace el mismo razonamiento, pasará esto último y el resultado (tres años) será *peor para ambos* de lo que hubieran obtenido cooperando (un año).²⁶

De esta manera, en el *dilema del prisionero* lo racional es la defección, el egoísmo. Pero ello sólo si es a un turno único ya que en un juego con repeticiones, la elección racional puede ser la cooperación (obviamente si no tratamos con un miope para quien la ganancia inmediata de la defección puede dominar sobre las ganancias futuras de la cooperación).²⁷ Pensemos por ejemplo en el caso de las organizaciones mafiosas en donde el costo de la deslealtad es pagado con la propia vida. Como este ejemplo y el propio nombre del juego nos remiten a uno de los aspectos más oscuros de la *convivencia social*, es pertinente hacer una aclaración: el *dilema del prisionero* es un juego

²⁶ Cuando el resultado es el *peor para todos* de cualquier otro resultado posible, se le denomina "subóptimo de Pareto". La suboptimalidad es un rasgo perverso de la interacción social. De hecho, uno de los principales problemas de la teoría de elección racional es el paso de la racionalidad individual a la racionalidad colectiva. Por lo que se refiere en especial a las elecciones sociales dentro de un sistema democrático, destaca el reconocido trabajo de Kenneth J. Arrow, *Social Choice and Individual Values* de 1951: de preferencias individuales consistentes no se puede obtener, con base en la regla de la mayoría, preferencias sociales también consistentes. Aquí el problema, conocido como la *paradoja del voto*, es que se viola la regla de transitividad.

²⁷ Sobre el *dilema del prisionero* y la teoría de la cooperación existe un estudio de Robert Axelrod en donde presenta los resultados de un *torneo computarizado* (Cf. Robert Axelrod, *La evolución de la cooperación. El dilema del prisionero y la teoría de juegos*). Como vimos, el *dilema del prisionero* es un juego en el que los jugadores no tienen propiamente intereses contrarios, pueden beneficiarse mutuamente con la cooperación pero siempre existe la posibilidad de que uno abuse del otro o que ninguno coopere. Buscando la mejor estrategia para este juego, Axelrod convocó al mencionado torneo el cual fue ganado por el programa más sencillo: *toma y daca*. La estrategia del programa ganador es comenzar cooperando y después hacer en cada jugada lo que el otro jugador hizo en la anterior. Este resultado

que corresponde a diversas situaciones sociales. Tomemos otro tipo de referencia: cuando en la Ciudad de México se planteó la conveniencia de que los automóviles particulares dejaran de circular un día a la semana, la mayoría de los automovilistas siguieron circulando todos los días esperando que los demás sí dejaran de circular uno. Fue necesaria la implementación oficial del *Programa Hoy No Circula*, es decir, que el Estado hiciera que el costo de la no-cooperación fuera más alto que el de la cooperación, para que efectivamente la mayoría de los automovilistas dejaran de circular un día a la semana.²⁸

Otra de las distinciones básicas de la teoría de juegos es la diferencia entre juegos con estrategia dominante y juegos sin estrategia dominante. Como señalamos, el *dilema del prisionero* tiene una estrategia dominante porque, más allá de que la recompensa de cada uno esté determinada por la decisión de todos, para cada jugador hay un curso de acción o estrategia que es *su mejor opción independientemente* de qué decida hacer el otro. La estrategia dominante en este caso es la no-cooperación (el egoísmo), aun cuando lleve al peor resultado para los dos. Pero no todos los juegos son así, veamos el

permitiría pensar en la posibilidad de una cooperación fundada exclusivamente en el principio de reciprocidad, lo cual retomaremos más adelante con el juego de los seguros.

²⁸ Dentro del individualismo metodológico se ha problematizado la acción racional de los agentes en la medida en que, por lo que se refiere a las acciones colectivas, puede llevar al peor resultado posible en lugar de al mejor. Incluso, se ha planteado que los teóricos del contrato social se equivocaron al creer que el Estado era necesario porque los hombres son irracionales (pasionales); sería justo al revés, porque hay un exceso de racionalidad, el Estado tiene la función de hacer que el costo de la cooperación sea menor que el de la no-cooperación. (Cf. Martin Hollis, "Rational Man and Social Science" en *Rational Action. Studies in Philosophy and Social Science*, ed. Ross Harrison, pp. 1-15) Este grave problema de la racionalidad instrumental que, al basarse exclusivamente en las preferencias individuales, lleva a la irracionalidad social, a subóptimos colectivos, ha sido el centro de atención de diversos autores. Las propuestas para llegar a alguna solución, además de la ya mencionada en la que el Estado tiene el poder de hacer que los costos de la no-cooperación sean más altos que el de la cooperación, han recurrido a la incorporación de diversos recursos éticos, como son las aclaraciones de que el sujeto racional no es necesariamente egoísta, que tiene una *pluralidad de motivaciones*, que también tiene *preferencias morales*, o que se pueden tomar *decisiones imparciales* bajo un "velo de ignorancia", por mencionar sólo algunas de las más relevantes. En general estas últimas propuestas tratan de garantizar tanto la libertad y responsabilidad individual como la racionalidad de la cooperación y de las decisiones colectivas. (Cf. Amartya Sen, *On Ethics and Economics*; John Harsanyi, "Morality and the Theory of Rational Behaviour" en *Utilitarianism and Beyond*; y John Rawls, *A Theory of Justice*)

llamado *juego de los seguros* en donde *la mejor opción depende* de lo que decida hacer el o los otros. En este juego, cada agente sigue teniendo dos opciones: cooperar (ser altruista, A) y no-cooperar (ser egoísta, E); y cuatro posibilidades: todos cooperan (A, A), todos no-cooperan (E, E), uno coopera y el o los otros no (A, E), uno no-coopera y el o los otros sí (E, A). Pero, a diferencia del *dilema del prisionero*, el *juego de los seguros* se caracteriza porque cada agente prefiere cooperar pero sin que se aprovechen de él. Pensemos por ejemplo en alguna tarea colectiva en la que los participantes están convencidos de su pertinencia pero condicionan su participación a la participación de los demás, como puede ser el caso de la realización de alguna empresa familiar.

En el *juego de los seguros* cada jugador clasifica sus alternativas de la siguiente manera: prefiere en primer lugar la situación (A, A), después (E, A), seguida de (E, E) y por último (A, E). Aquí, no hay una estrategia dominante; incluso el altruismo, que el individuo prefiere, depende del altruismo de los otros. El óptimo es (A, A), pero a pesar de ser estable a nivel individual no se puede acceder a él individualmente, es decir, el individuo aunque prefiera ser altruista será egoísta si los otros se comportan egoístamente; por ello, este juego tiene dos puntos de equilibrio: (A, A) y (E, E). Sin embargo, aclaremos, el que este juego no tenga una estrategia dominante, no significa que no tenga una solución, es decir, si los jugadores son racionales y están bien informados (saben las preferencias de los demás) convergen tácitamente en una estrategia, en este caso, cooperar.²⁹ Por ello, en este juego no sólo se deben compartir preferencias sino también el saber sobre ellas. De hecho, aunque por causas diferentes al *dilema del prisionero* -en donde lo que falta es una coordinación (los prisioneros están separados y no pueden comunicarse)-, en el *juego de los seguros* también podemos llegar a un resultado que es *peor*

²⁹ Con este juego se puede destacar cómo el individualismo metodológico no se limita a explicar sólo el egoísmo, también puede explicar la solidaridad y el consenso.

para todos si falta la información requerida. Siguiendo con el ejemplo de la familia pensemos en cualquiera de los múltiples casos en donde un *malentendido* puede desembocar en un *pequeño infierno*.³⁰

Frente a estos dos casos de suboptimalidad, hay todavía una situación más *perversa*: los juegos sin solución. En estos juegos ni siquiera hay una acción estratégicamente racional y se caracterizan por la inestabilidad. Este es el caso del *juego de la gallina*, en donde cada agente tiene un incentivo para hacer E si los demás hacen A y viceversa. Aquí las preferencias son en el siguiente orden: en primer lugar cada agente prefiere la situación (E, A), después (A, A), seguida de (A, E) y por último (E, E). Aquí el ejemplo por excelencia es la prueba de valentía que llegó a caracterizar a una época y que se realizaba a través del enfrentamiento directo de dos coches a gran velocidad, en donde gana quien no cambia de dirección. Obviamente, si ninguno de los dos agentes gira el volante, chocarán de frente y muy probablemente ambos morirán. Por ello, el orden de preferencia de cada uno es mantenerse firme y que el otro flaquee, después que los dos pierdan, seguido de que gane el otro y por último la colisión. En estos casos, en donde no hay una solución racional, cada jugador supondrá algo del otro y actuará en consecuencia; este supuesto, sin embargo, tampoco es racional, en el sentido de derivarse de la hipótesis de que los demás son tan racionales y bien informados como él. De hecho, la situación se vuelve intolerable porque cada jugador al tratar de *adivinar* que hará el otro, sabe que él también está haciendo lo mismo; finalmente, a pesar de ser seres racionales y considerarse así, cada uno tratará al otro como un ser causalmente determinado, reintroduciéndose de esta manera el pensamiento paramétrico dentro de la

³⁰ En la explicación de las suboptimalidades sociales puede llegar a ser difícil determinar si las preferencias son del tipo *dilema del prisionero* o *juego de los seguros*, lo cual también tiene consecuencias prácticas importantes ya que en cada caso se requiere de respuestas distintas: si las preferencias son condicionalmente altruistas se necesita proporcionar la información que lleve a las personas a converger, en cambio, si se trata de preferencias incondicionalmente egoístas se necesita algún tipo de coerción.

racionalidad estratégica. Si bien se trata de un juego sin solución, no es objetivamente indeterminado, finalmente cada uno de los competidores mantendrá o girará el volante, determinación que sólo se podrá explicar recurriendo a una teoría causal.

Un ejemplo menos violento es el voto. Obviamente, a nivel nacional la posibilidad de que un individuo pueda influir en la votación es prácticamente nula, por ello hay que suponer en este caso que el voto seguramente no obedece a un interés propio ni a sus posibles consecuencias. Pero en electorados pequeños la situación es otra: un elector puede considerar que si para la mayoría es irracional votar sólo una cantidad pequeña de personas acudirá a las urnas, por lo que entonces sí es racional ir a votar ya que su voto sí será significativo para el resultado final. El problema es que los demás pueden hacer el mismo razonamiento y como consecuencia muchos electores votarían, volviéndose irracional ir a votar. Pero, otra vez, se puede pensar que los demás también hacen la misma consideración, lo que nos lleva nuevamente a la proyección de pocos votantes en donde sí es racional ir a votar. Y así sucesivamente. Esta regresión sería infinita y sólo se puede interrumpir por la propia acción, ya sea ir a votar o la decisión consciente de no votar, que finalmente responde a una consideración implícita o explícita de cómo se cree (supuesto injustificadamente) que actuarán los demás.

Con esto tenemos que, más allá de las visiones hoy tradicionales del paradigma de elección racional, la intencionalidad no es garantía de racionalidad así como ésta tampoco es sinónimo de optimalidad. Finalmente, el reconocimiento de la irrupción de la incertidumbre y de la existencia de juegos con múltiples equilibrios y de juegos sin solución, exige una evaluación más modesta de los posibles alcances de la teoría.

Incluso el propio Morgenstern expresa la necesidad de tener claros dichos alcances:

[...] la esperanza de encontrar una solución únicamente mejor para los asuntos sociales es vana: tales arreglos no tienen estabilidad [...] La falta de identificación de un arreglo o imputación singular no es una deficiencia de la teoría de los juegos. Más bien se revela aquí una característica fundamental de la organización social, humana, que no puede describirse adecuadamente por otros medios.³¹

El gran supuesto de que la racionalidad dicta qué es lo que debe hacer el agente, y que cuando actúa de modo diferente actúa irracionalmente, es en muchos casos un supuesto espurio. A ello se suma el gran salto que implica el paso de la racionalidad individual a la racionalidad colectiva. Estos espacios problemáticos constituyen el contexto teórico dentro del cual Elster hace su crítica y al mismo tiempo su aportación a la TER. Finalmente, a pesar de que se considera que en las ciencias sociales el tipo de explicación pertinente es la explicación intencional de las acciones individuales con base en los deseos y las creencias del agente, como veremos, para este autor se debe explicar también la propia formación de estos deseos y creencias así como los resultados (muchas veces no previstos) de las interacciones individuales. En otras palabras y aunque estemos a *años luz* de alcanzarla, al menos se puede postular la necesidad de una teoría social que comprenda al mismo tiempo una explicación causal de la formación de deseos y creencias individuales, una explicación intencional de las acciones individuales y una explicación causal de los macroestados como resultado de dichas acciones.³²

³¹ Oskar Morgenstern, "Teoría de los juegos", pp.173-174.

³² Sin embargo, hay que aclarar que, como veremos en los siguientes capítulos, para Elster el planteamiento de una teoría general en ciencias sociales es un *sueño ilusorio*. En lugar de formular teorías generales, Elster nos propone ubicar los *mecanismos*, es decir, las cadenas causales específicas que en cada caso nos permiten explicar *a posteriori* las acciones o interacciones individuales. (Cf. Jon Elster, *Political Psychology*)

2. Estructura de la explicación

Con este primer acercamiento a los conceptos básicos de la TER, podemos pasar a abordar el problema de la estructura de sus explicaciones.³³ Como dijimos, se trata de un tipo de explicación intencional que se caracteriza en general por relacionar tres elementos: la acción a explicar, el conjunto de creencias que tiene el individuo y el conjunto de deseos que se le puede imputar.

Desde Aristóteles se puede rastrear la preocupación por identificar las condiciones que hacen que un comportamiento pueda ser considerado una acción.³⁴ Vivimos en un mundo en donde interactúan entre sí una diversidad de entidades, pero sólo los seres humanos pueden ser considerados agentes en tanto son la causa de su propio comportamiento y por ello, a diferencia de los procesos naturales, se le puede atribuir responsabilidad:

[...] no podemos referir ni reducir nuestros actos a principios distintos de los que hay en nosotros mismos, aquellas acciones que tienen en nosotros su principio dependen, también ellas, de nosotros y son voluntarias. No aduzco otra prueba que la conducta privada de cada uno de nosotros y la de los mismos legisladores. Estos infligen castigos y penas a los que obran mal, de no ser que los actos hayan sido impuestos por la violencia o hayan sido causados por una ignorancia involuntaria.³⁵

³³ Por lo que se refiere a la estructura de las explicaciones de la TER, explícitamente Elster reconoce la deuda intelectual que tiene con Donald Davidson. De hecho, Elster es uno de los autores que, con el trabajo titulado "The Nature and Scope of Rational-Choice Explanation", colabora en el reconocido libro *Actions and Events, Perspectives on the Philosophy of Donald Davidson*, editado por Ernest LePore y Brian P. McLaughlin (pp. 60-72). Este trabajo es posteriormente recuperado por Elster en su introducción al libro *Rational Choice*.

³⁴ Aquí retomamos parte de la reconstrucción de la propuesta aristotélica que hace Eduardo Rabossi en "La filosofía de la acción y la filosofía de la mente" en *Acción humana*, coord. Manuel Cruz, pp. 5-20. Artículo en donde se ubica a Aristóteles como el autor que fija la agenda de la filosofía de la acción, y a Descartes y Hobbes como los que la fijaron para la filosofía de la mente.

³⁵ Aristóteles, *Ética Nicomaquea*, 1113b, p. 1201.

Una acción voluntaria es intencional, está dirigida hacia un fin elegido e implica una deliberación sobre los medios para alcanzarlo. En este sentido, la acción humana se caracteriza por estar orientada de manera racional:

Puesto que el objeto de nuestra preferencia es una decisión que nos lleva a lo que depende de nosotros, la elección consciente podría muy bien ser una aspiración acompañada de deliberación, hacia lo que depende de nosotros, pues tomamos una decisión, luego de una previa deliberación, y tendemos a su realización de acuerdo con esta deliberación.³⁶

Si bien Aristóteles distingue entre deliberación (en tanto cálculo de medios) y sabiduría práctica (en tanto elección de fines correctos), ambas son indispensables para la racionalidad de la acción.

Dentro de la filosofía contemporánea, destaca un artículo publicado en 1963 que se volvería una referencia obligada para diversas áreas de la filosofía. Preguntándose por la relación entre razón y acción, Donald Davidson aceptaba que las explicaciones teleológicas se distinguen de las explicaciones naturales en cuanto no incluyen leyes, pero insistía en que ambas explicaciones recurren a conexiones causales. En la explicación de una acción intencional, que descansa en los motivos o razones que tuvo un agente para hacer lo que hizo, la relación de éstas últimas y la acción es una relación de causa y efecto.³⁷

La razón por la que alguien hace algo comprende dos aspectos: una actitud favorable por parte del agente hacia acciones de una clase determinada (en la que no sólo se incluyen deseos, sino también impulsos, instintos, convicciones morales, principios estéticos, prejuicios económicos, convencionalismo sociales, metas y valores tanto públicos como privados), por un lado, y, por otro, la creencia del agente (que puede ser sólo una suposición

³⁶ Aristóteles, *Ibid.*, 1113a, p. 1200.

o percepción) de que la acción en cuestión es de esa clase. Lo que finalmente Davidson llama la *razón primaria* por la que un agente realizó una acción, y que se presenta como la causa de dicha acción, incluye tanto la actitud favorable como la creencia antes señalada.

Recordemos el ya célebre ejemplo: presiono el interruptor, enciendo la luz, ilumino el cuarto y alerto (sin saberlo) a un merodeador de mi presencia en la casa. Estas son cuatro descripciones distintas de un mismo suceso y sólo bajo cierta descripción las razones que tuve para actuar explican mi acción. Si sostengo que presioné el interruptor *porque* quería encender la luz, esta razón *explica* el acto de presionar el interruptor pero no el de haber alertado al merodeador ni el de haber iluminado el cuarto. De esta manera tenemos que para Davidson

R es una razón primaria por la que un agente realizó la acción *A* en la descripción *d*, sólo si *R* consiste en una actitud favorable del agente hacia las acciones que poseen cierta propiedad y en una creencia suya de que *A* en la descripción *d* tiene esa propiedad.³⁸

Aquí hay que insistir en que la *razón primaria* de una acción es su causa y por eso puede explicarla e incluso justificarla. En otras palabras, una razón primaria no sólo remite a las razones que tuvo un agente para actuar, en el sentido de que hay razones que justifican y presentan como razonable su acción, sino también y sobre todo a que éstas sean las razones por las que se llevó a cabo dicha acción. De hecho, alguien puede tener buenas razones para una acción pero realizarla sin que éstas sean la *razón primaria* de la misma. Pensemos por ejemplo en alguno de los múltiples casos en donde por accidente o de manera fortuita, no intencionalmente, logramos algo que queríamos.

³⁷ Cf. Donald Davidson, "Actions, Reasons, and Causes" en *Essays on Actions and Events*, pp. 3-19.

³⁸ *Ibid.*, p. 5.

Reconociendo estas aportaciones de Davidson, Jon Elster realiza su análisis de la estructura de las explicaciones de la TER. Como se trata de la explicación de una acción intencional, se parte de una triple relación entre la acción (A) a explicar, las creencias (C) y los deseos (D) del agente. Esta relación es definida por las siguientes tres condiciones:

- 1) Dado C, A es el *mejor* medio para realizar D
- 2) C y D *causaron* A
- 3) C y D *causaron* A *en tanto que* razones

Con estas condiciones se establece que C y D no sólo son buenas razones para hacer A sino también que racionalizan y causan A. Para seguir con el ejemplo anterior, nos encontramos en el atardecer leyendo un libro, y como (C) creemos entre otras cosas que hay energía y una instalación eléctrica adecuada y (D) deseamos tener luz para poder continuar con nuestra lectura, lo mejor es (A) presionar el interruptor de la luz. Sin embargo, hay casos en donde las relaciones entre C, D y A no son tan sencillas. Dejando de lado por el momento el problema de la optimalidad expresada por la primera condición, con las otras dos condiciones se excluyen las llamadas coincidencias de primer y segundo orden respectivamente. Una persona puede tener razones para actuar de la manera en que lo hace pero no son esas razones la que la mueven a hacerlo, incluso accidentalmente se puede llegar a hacer algo que uno tiene razones para hacer, y en el caso extremo hasta una conducta compulsiva puede llegar a ser adecuada a una situación. Por otra parte, cuando las razones para hacer una acción sí son su causa, puede ser que la causen de una manera incorrecta. Otro de los conocidos ejemplos de Davidson es el siguiente: un hombre al tratar de matar a otro de un disparo, no acierta en el blanco pero provoca una estampida de cerdos salvajes que finalmente atropella a su víctima y la mata. En este caso, si bien la intención era efectivamente la

muerte del otro, no podríamos decir que lo mató intencionalmente ya que la cadena causal que explica el hecho es otra.

A estas tres primeras condiciones (una de optimalidad y dos de racionalidad o no coincidencia) que caracterizan a una explicación intencional, la explicación de elección racional añade inicialmente dos condiciones más ya que la conducta intencional es considerada racional sólo si los deseos y las creencias son también ellos racionales.³⁹ De esta manera tenemos que:

- 4) El conjunto de creencias (C) es internamente consistente
- 5) El conjunto de deseos (D) es internamente consistente

Hay que aclarar que un sistema de creencias es consistente sólo si existe un mundo posible en el que todas sus creencias son ciertas y creídas al mismo tiempo. Veamos un conocido ejemplo de una acción intencional que sería irracional debido a la inconsistencia del conjunto de creencias: en cierta ocasión, como tenía una herradura en su puerta, se le preguntó a Niels Bohr si la había puesto porque creía que las herraduras traen buena suerte; a lo que respondió (obviamente pudo haber sido en broma): "No, pero me dijeron que traen suerte incluso a quienes no creen en ellas."

De manera similar, el conjunto de deseos es consistente sólo si existe un mundo posible en el que se crea que esos deseos se pueden cumplir y que se pueden cumplir a través del intento de hacerlo. En otras palabras, un deseo sería irracional si se cree que no puede ser realizado (como es el caso del deseo de que todos ganen más dinero que el ingreso promedio) o bien si no puede ser el resultado de una acción voluntaria (como es el caso del sueño o

³⁹ Estas dos condiciones son básicas para la explicación de elección racional y marcan su gran diferencia con la explicación intencional. Como veremos más adelante, puede haber un acción intencional que no sea racional. Por ahora lo relevante es destacar que si los deseos son internamente inconsistentes no hay manera de realizarlos y mucho menos de escoger el mejor camino para ello; aunque puede darse el caso de que, debido a creencias también

del olvido). Asimismo, por lo que se refiere a las jerarquizaciones o preferencias entre los propios deseos, uno de los criterios básicos de consistencia es que sean transitivas, es decir, que si prefiero A sobre B y B sobre C, entonces prefiera A sobre C. Si prefiero ir al cine que ver un video en casa, y ver un video que un programa de televisión, sería irracional que prefiriera ver un programa que ir al cine.⁴⁰

Sin embargo, la explicación de elección racional exige una forma más amplia de racionalidad que la simple consistencia.⁴¹ Al menos por lo que se refiere en especial a las creencias, éstas deben ser racionales en el sentido de

inconsistentes, se crea lo contrario, es decir, que sí se pueden realizar y se pretenda escoger la mejor manera de llevarlos a cabo.

⁴⁰ Si bien las preferencias pueden cambiar con el tiempo, las condiciones de su ordenación se ocupa sólo de las preferencias del agente en el presente. Para ubicar estas condiciones usaremos la expresión "xPy" para indicar que "el agente prefiere escoger x a y" y "xly" para indicar que "al agente le es indiferente escoger entre x e y". Las primeras tres condiciones se refieren a las *asimetrías* entre las preferencias del agente, la cuarta condición a que éstas estén ligadas entre sí, y las últimas cuatro condiciones a su transitividad:

- a) Si xPy, entonces no yPx
- b) Si xPy, entonces no xly
- c) Si xly, entonces no xPy y también no yPx
- d) xPy o yPx o xly, para todos los resultados relevantes x e y
- e) Si xPy y yPz, entonces xPz
- f) Si xPy y xly, entonces zPy
- g) Si xPy y ylz, entonces xPz
- h) Si xly y ylz, entonces xly

Si las preferencias de un agente cumplen estas condiciones, los elementos ordenados según sus preferencias se dividen en *clases de indiferencia*: al agente le son indiferentes los elementos de una misma clase pero establece preferencias entre éstos y los elementos de otra. Esta preferencia permite ordenar las diversas *clases de indiferencia*, ordenación sobre la que se puede establecer la importancia relativa o *utilidad* que tiene para el agente los elementos de cada clase. De esta manera, la utilidad de x es mayor que la de y si xPy, y ambas utilidades son iguales si xly.

⁴¹ Con esto, sin embargo, no se supone que la condición de consistencia sea algo simple. En este nivel también surgen problemas. Aunque no nos detengamos en ello, es pertinente volver a mencionar el gran golpe que le dio Gödel a la pretensión de una formalización exhaustiva y consistente en matemáticas; de hecho, frente a la posibilidad de interpretar su teorema en el ámbito de los asuntos humanos, el propio Gödel señaló en una carta: "Una sociedad completamente no-libre (i.e., una que procede en todas las cosas mediante estrictas reglas de "conformidad") será, en su comportamiento, o inconsistente o incompleta, i.e., incapaz de resolver ciertos problemas, quizá de vital importancia. Desde luego, ambas pueden poner en riesgo su supervivencia en una situación difícil. Una observación similar se puede aplicar también a los seres humanos individuales". (Carta del 15 de marzo de 1961 citada por Hao Wang, discípulo de Gödel, en el libro *A Logical Journey. From Gödel to Philosophy*, p. 4.) Estas

estar fundadas en la evidencia disponible.⁴² En paralelo con las tres primeras condiciones en las que se establece la relación entre D, C y A, esta condición sobre la relación entre creencia y evidencia se subdivide en una de optimalidad y dos de no coincidencia:

- 1b) La creencia tiene un grado *máximo* de plausibilidad inductiva dado por la evidencia
- 2b) La creencia es *causada* por la evidencia disponible
- 3b) La evidencia causa la creencia de *manera correcta*

La evidencia por sí misma no garantiza que tengamos la mejor creencia posible, asimismo ésta puede no estar causada por aquella o estarlo pero de manera incorrecta. Como veremos más adelante, con 2b) se excluyen las creencias causadas por los deseos y no por la evidencia disponible (*wishful thinking*) y con 3b) se excluyen las creencias que resultan de procesos de inferencia con errores compensatorios. Pensemos por ejemplo en alguno de los múltiples casos de la vida cotidiana en donde se sostiene una creencia sin ningún fundamento (como el enfermo que niega los síntomas de su enfermedad y se cree sano) o los casos en donde por coincidencia se tiene una creencia correcta (como cuando un hipocondríaco realmente se encuentra enfermo).

Pero una creencia fundada en la evidencia disponible presupone a su vez la decisión previa de concluir la recolección de información. En este

observaciones de Gödel ponen en duda la posibilidad de que un sistema de reglas, de creencias o de deseos (base de la explicación racional), sea consistente y completo a la vez.

⁴² Al igual que en las creencias, podemos preguntar también por la racionalidad de los deseos más allá de su simple consistencia. Aunque muchos economistas que enfatizan el carácter instrumental de la elección racional lo negarían, Elster ha tratado de desarrollar algunos elementos, los cuales sin embargo no ofrecen una respuesta positiva suficientemente fuerte como para poder ser incorporados a la definición formal de la explicación de elección racional. Como veremos en los siguientes capítulos, frente a una *teoría estricta* de racionalidad en tanto que consistencia, Elster nos propone una *teoría amplia* que comprenda la racionalidad en un sentido más sustancial (Cf. Jon Elster, *Sour Grapes. Studies in the subversion of rationality*, especialmente el capítulo 1 "Racionalidad", pp. 1-42)

sentido, una elección racional de hecho estaría implicando al menos dos decisiones: no sólo elegir la acción a realizar sino también cuándo dejar de recoger la información pertinente. Como de la racionalidad de esto último depende la racionalidad de la propia acción, se debe incorporar también una condición de optimalidad, aunque sea muy débil, por lo que se refiere a la recolección de la información:

N) La cantidad de información recogida se ubica dentro de una franja cuyos límites superior e inferior son definidos por la propia situación

Aunque no se puede establecer un criterio general de optimalidad en la recolección de información, hay que destacar que ésta obedece al deseo y otras creencias del agente e implica un costo y un tiempo, por lo que no puede prolongarse indefinidamente. Cada situación tiene un mínimo y un máximo de información que es *razonable* tener antes de hacer la elección y parecería que en la mayoría de los casos es posible saber si hay que seguir buscando información o dejar de hacerlo. De hecho, para seguir con los últimos ejemplos, cuestionaríamos la racionalidad de un enfermo que no quiere hacerse los análisis correspondientes para poder seguir creyendo que está sano o la del hipocondríaco que continua indefinidamente haciéndose estudios para no tener que aceptar que no está enfermo.

Retomando estas últimas observaciones, por lo que se refiere a la exigencia de una racionalidad más estricta en las creencias, tenemos que:

6) La relación entre C, D y E satisface 1b), 2b), 3b) y N)

Es decir, las creencias que intervienen para decidir realizar una acción como el mejor medio para satisfacer un deseo, deben estar fundadas en y causadas por la evidencia disponible, la cual dependiendo de la situación debe

ser la adecuada. Para seguir con casos de salud, pensemos en la siguiente situación: un médico en la sección de emergencias de un hospital recibe a un enfermo grave, sus acciones están encaminadas a salvarle la vida al paciente, pero sería un irresponsable si procediera sin los análisis y estudios necesarios aunque tampoco puede realizar con calma todas las pruebas posibles dada la inminente muerte. Como dijimos, en la mayoría de los casos se puede saber cuál es la información indispensable (y que sería irracional no recoger) y cuál es menos importante (y que sería irracional recoger dados los costos).

Sin embargo, todas estas condiciones tampoco son suficientes para establecer la estructura de la explicación de elección racional. Esta incluye una última condición. Como existe la posibilidad de una conducta *akrática* o la llamada *debilidad de voluntad*,⁴³ para que ésta quede excluida se debe tener también una condición sobre la relación entre deseos y acción:

7) Dado C, A es la mejor acción con respecto al grupo completo de deseos sopesados.

La conducta *akrática* se caracteriza de la siguiente forma: a) Hay un juicio *prima facie* sobre la bondad de Y; b) Hay un juicio *prima facie* sobre la bondad de X; c) Consideradas todas las cosas hay un juicio de que X es mejor; d) Elegimos Y. Pensemos por ejemplo en el lamentable caso de un alcoholíco que obviamente tiene razones para tomar pero él mismo considera que es mejor abstenerse, y a pesar de ello sigue tomando.

⁴³ En la primera parte del siguiente capítulo nos detendremos especialmente sobre este problema. Como veremos, la conducta *akrática* puede ser incluida dentro de la explicación intencional (aquí nuevamente destaca el trabajo de Davidson) pero no dentro de la explicación de elección racional. Sin embargo, así como señalamos en la nota anterior que Elster propone la necesidad de una *teoría amplia* frente a la *teoría estricta* de la racionalidad, veremos también que frente a la *teoría de la racionalidad perfecta* propone la necesidad de una *teoría de la racionalidad imperfecta*. (Cf. Jon Elster, *Ulysses and the Sirens. Studies in rationality and irrationality*, especialmente el capítulo II "Imperfect Rationality: Ulysses and the Sirens", pp. 36-111).

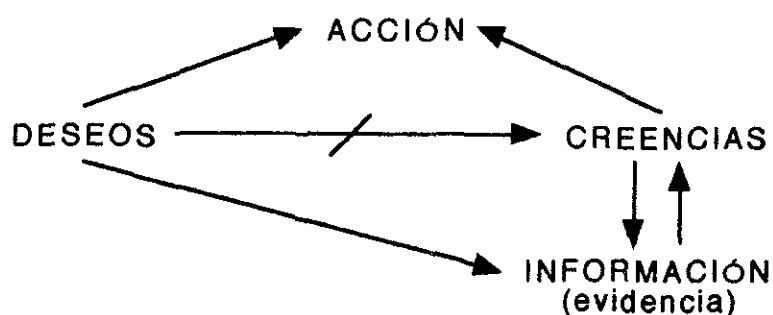
Para finalizar este apartado destaquemos cómo estas siete condiciones de la explicación de elección racional satisfacen tres diferentes grupos de requerimientos:

El primero se refiere a la optimalidad: dadas las creencias, la acción es la mejor manera de satisfacer los deseos; dada la evidencia, las creencias son las mejores que se pueden formar; y, dados los deseos, la cantidad de evidencia es óptima.⁴⁴

El segundo grupo se refiere a la consistencia: las creencias y los deseos tienen que estar libres de contradicciones internas.

Por último, el tercer grupo se refiere a la causalidad: la acción no sólo tiene que ser racionalizada por los deseos y las creencias, sino que también debe ser causada por ellos y causada de *manera correcta*; asimismo, las creencias deben ser causadas por la evidencia y deben ser causadas de *manera correcta*.

De esta manera queda conformado el siguiente modelo de la TER:



De acuerdo con lo que hemos visto, en este esquema las flechas que unen a los distintos elementos del modelo se pueden interpretar tanto como

⁴⁴ Las condiciones de optimalidad son el foco de atención, casi exclusivo, de la explicación de elección racional en las ciencias sociales. Esto parecería justificado en la medida en que no se busca explicar específicamente la acción individual sino más bien la conducta de grandes cantidades de personas en circunstancias externas similares. En estos casos, las condiciones de consistencia se dan por satisfechas y se presenta como algo casi imposible y sin mucho sentido pretender cubrir las condiciones de causalidad. Sin embargo, a partir de la gran difusión

causalidad como optimalidad. La flecha con una diagonal indica que la influencia directa de los deseos en la formación de creencias no está permitida (caso de *wishful thinking*).

3. Límites y fracasos de la teoría

Los principales problemas de la TER surgen de uno de sus grandes supuestos: la definición unívoca de la noción de elección racional. No olvidemos que para esta propuesta decidir racionalmente es elegir la estrategia que *mejor* satisface los deseos del agente (dadas sus creencias y restricciones); por ello, la teoría falla si no se puede establecer unívocamente cuál es la elección racional. Pero, incluso en el caso de que se pueda establecer cuál es la elección racional, la teoría también falla si la conducta que es realizada difiere de la normativamente prescrita. En otras palabras, como sintetiza Elster, la TER puede fallar porque no ofrece los requerimientos racionales o porque la gente no se comporta de acuerdo a ellos. (Fig. 2)

El primer problema surge de la no-univocidad o incluso la no-existencia de racionalidad en cualquiera de las condiciones de optimalidad que acabamos de ver: en la acción, en la creencia o en la evidencia. Veamos cada una de ellas.⁴⁵

La racionalidad de la acción puede ser no-unívoca: con creencias y deseos fijos, las mejores alternativas pueden serle indiferentes al agente ya sea porque tienen una misma característica o son tan distintas que se

de conductas auto-destructivas en la sociedad contemporánea, Elster no ve tan fácilmente justificable el usual descuido de la séptima condición.

⁴⁵ Aunque puede haber problemas de no-univocidad tanto en las creencias como en la evidencia, no son problemas que surgen propiamente en la toma de decisiones prácticas. Por ello, Elster ni nosotros aquí nos ocuparemos de ellos. Por otra parte, como ya señalamos, dejaremos para los siguientes capítulos los problemas relacionados con la racionalidad de los deseos.

compensan al ser evaluadas.⁴⁶ Pensemos en dos latas idénticas de sopa o dos departamentos uno muy bien ubicado y el otro con una excelente orientación. Obviamente, para fines normativos esto no es tan grave, ya que como se trata de dos opciones iguales y máximamente buenas, no se necesitaría de una teoría que nos determine cuál escoger. Sin embargo, para fines explicativos esta indeterminación es problemática sobre todo cuando las alternativas son muy diferentes entre sí.

Además de la no-univocidad, también puede darse la no-existencia de una acción racional: aquí el problema surge por la incapacidad del agente para comparar correctamente todas las alternativas y en este sentido sus preferencias son incompletas o están desconectadas. Puede ser el caso de alternativas que son nuevas para él por lo que desconoce sus posibles relaciones y consecuencias, o también porque dichas alternativas implican una comparación interpersonal de utilidades que el agente considera que no puede realizar. Pensemos por ejemplo en el caso de un viajero que se siente mal y debe elegir un médico del directorio telefónico o una voluntaria que debe distribuir los donativos recabados entre diversos damnificados.

De igual manera, puede no-existir una creencia racional: en una elección bajo incertidumbre no se pueden ligar las mejores acciones con los mejores resultados, por ello tampoco se pueden establecer cuáles son las mejores alternativas. Pensemos en un enfermo que debe decidir a cuál protocolo de investigación solicitar su ingreso o un empresario que debe escoger entre diversos proyectos de innovación tecnológica a cuál se le otorgará el financiamiento. Aquí, hay que insistir, el problema surge por la incertidumbre y no por la incapacidad del agente como en el caso anterior. Esta falta de racionalidad, como vimos, también aparece en los juegos no-cooperativos sin

⁴⁶ La indiferencia también puede surgir entre estrategias mixtas en punto de equilibrio. Este caso es más interesante y problemático en virtud de la importancia que tienen estas estrategias en la teoría de juegos: ante la menor incertidumbre o desconfianza, el (los) agente(s) puede(n)

solución: ya sea porque el juego carezca de un punto de equilibrio o porque tenga varios, la indeterminación estratégica surge de la imposibilidad de determinar qué harán los otros jugadores. El juego por excelencia es el de la gallina.

Finalmente, puede no-existir una cantidad óptima de evidencia: este aspecto obedece a la imposibilidad de estimar el valor marginal esperado en la búsqueda de información. A excepción de algunas situaciones muy estereotipadas (como son algunas formas de diagnóstico médico), no se puede saber con anticipación lo que probablemente se llegará a conocer si se continúa con la búsqueda de información. Incluso, pensemos en el ámbito de los negocios o de la política en donde la información puede volverse rápidamente obsoleta. Si bien, como hemos insistido, un agente racional tomará finalmente una decisión en lugar de diferirla indefinidamente, esto no implica que dicha decisión sea una elección racional.

Estas dificultades en los requerimientos racionales tanto en la acción como en la creencia y en la evidencia, no necesariamente invalidan a la TER. Para Elster, aunque no se logre el ideal de *eliminar todas las opciones menos una*, la teoría puede ayudar a eliminar algunas, es decir, a reducir el grupo de opciones a las mejores alternativas aunque no llegue a establecer cuál es la mejor elección. Esta ayuda varía de caso en caso. Puesto que frente a alternativas poco diferenciadas, el agente puede ser indiferente, y ante alternativas altamente diferenciadas, puede ser incapaz de compararlas, parece que la teoría es más útil en situaciones medianas: cuando las alternativas sí son significativamente diferenciadas pero no demasiado. Por otra parte, parece tener más poder para las decisiones paramétricas y menos urgentes que en las decisiones estratégicas y más urgentes.

abandonar el principio de utilidad máxima esperada y, siguiendo un criterio *maxi-min*, romper precisamente el punto de equilibrio. Recordemos por ejemplo el juego de los seguros.

En general, las limitaciones de la TER pueden estar señalando al menos dos cosas. Para fines normativos, una forma de *decisionismo*: se puede elegir sin tener razones para ello. Ante este *decisionismo* el camino se cierra y no queda nada más que agregar. Sin embargo, para fines explicativos, las limitaciones de la TER pueden estar indicando la necesidad de una teoría causal: establecer cuál es el mecanismo por el que una alternativa, que no puede ser determinada por consideraciones de elección racional, es escogida finalmente para su realización. Con esto último, la acción ya no sería el resultado de dos filtros sino de tres: las restricciones que definen el conjunto factible; la racionalidad que excluye algunas alternativas; y el mecanismo específico por el que queda seleccionada una de las opciones restantes.

Frente a estas observaciones, hay todavía una crítica más radical a la TER. Como dijimos, la teoría no sólo falla si no puede decidir unívocamente cuál es la elección racional, también falla si, a pesar de poderla establecer, la conducta realizada por el agente difiere de la normativamente prescrita. En esta segunda situación nos encontramos con que aun cuando las personas puedan ser racionales, a veces no lo son. Se podría argüir, retomando la última observación, que hay un mecanismo distinto a la elección racional que lleva a escoger la alternativa a realizar. Pero, incluso en esta propuesta de los tres filtros, tendríamos el problema de que se puede violar o falsear la racionalidad (ubicada como segundo filtro).

La no-racionalidad puede ocurrir, al igual que en la situación anterior, en cada uno de los enlaces de la cadena explicativa: en la propia acción, en la formación de creencias y en la recolección de evidencia.⁴⁷

En el nivel de la acción una forma muy común de irracionalidad es la *debilidad de la voluntad*, la cual queda excluida con la séptima condición de la estructura de la explicación de elección racional y que veremos más

detenidamente en el siguiente capítulo. Pero también son frecuentes la sobrevaloración de las probabilidades bajas y la subvaloración de las altas en situaciones de elección bajo riesgo; la conducta mágica en donde con una acción se pretende lograr un resultado que propiamente no resulta de ella; así como las conductas contraproducentes en donde la acción instrumental interfiere precisamente con el fin que se pretende alcanzar. Pensemos por ejemplo, en una persona recalcitrantemente adversa al mínimo riesgo o cuando alguien trata intencionalmente de olvidar alguna experiencia desagradable. Entre las conductas mágicas, Elster destaca el caso de un panfleto metodista del siglo XVIII en el que se exhortaba a la gente a acercarse a la religión. En este panfleto se podía leer: "El alma que viene no necesita el temor a no ser elegida, porque está dispuesta a venir". En este "calvinismo cotidiano", fruto de la paradoja de la predestinación, se sostendría entonces que con el hecho de ir a la reunión uno podría (mágicamente) hacer que fuera ya elegido.

Por lo que se refiere al nivel de la formación de creencias, destacan dos tipos de irracionalidad. Una "caliente" o "motivada" como es el caso de las formas de auto-engaño e ilusiones: la mala fe sartreana o las frecuentes sobrestimaciones de las propias habilidades. La otra "fría" en donde se hacen afirmaciones ignorando por ejemplo las propiedades estadísticas de la información disponible: generalizar a partir de una pequeña muestra o minimizar tendencias.

Finalmente, en el nivel de la recolección de evidencia, como ya mencionamos, es irracional que con el fin de tener razones para seguir sosteniendo una creencia, se recoja evidencia por abajo del límite inferior o por arriba del límite superior de la franja pertinente que es definida por la propia situación. Recordemos el caso del enfermo que prefiere no hacerse los

⁴⁷ Como dijimos, dejamos para los siguientes capítulos los problemas relacionados con la racionalidad de los deseos. En este caso, como veremos, el problema surge en la formación de los deseos en donde sobresalen las preferencias adaptativas y contraadaptativas.

estudios necesarios o del hipocondriaco que continua indefinidamente haciéndose análisis.

A pesar de estas fallas y limitaciones, para Elster la TER con su pretensión de explicar la conducta humana sigue teniendo un poder explicativo mayor que otras alternativas teóricas. Especialmente frente al *estructuralismo* que en una versión extrema se podría replantear en los siguientes términos: las restricciones iniciales (primer filtro) reducen el conjunto factible al grado de no dejar espacio para la elección.⁴⁸ Pensemos por ejemplo en las explicaciones muy difundidas hace algunas décadas en donde se le daba un peso excesivo a las condiciones sociales, como puede ser el pertenecer a una determinada clase social.

Para Elster una alternativa más interesante podría ser la teoría de las *normas sociales* en la que la racionalidad instrumental encaminada a escoger los mejores medios para realizar los fines, es considerada sólo como una norma social particular de las sociedades occidentales modernas. Esta teoría se podría replantear en los siguientes términos: como la explicación de la acción humana es a partir de las normas sociales, éstas ocuparían el lugar de la racionalidad individual (segundo filtro). Aquí uno de los principales problemas es la implícita indiferencia a la expansión del conjunto de las acciones realizables (definido por el primer filtro). Para la TER si hay nuevas posibilidades y una de ellas se presenta como una mejor alternativa, la acción necesariamente se modificará. De hecho, a lo largo de la historia tenemos una multiplicidad de ejemplos en donde una conducta tradicional es finalmente descartada dada la posibilidad de nuevas oportunidades. Por otra parte, no siempre se cumplen las normas, se tendría que establecer qué determina que una norma sea fuerte y estable. Sin embargo, y esto le interesa especialmente rescatar a Elster, se trata de una teoría que, a pesar de sus múltiples problemas, acertadamente señala que no todas las acciones están inmediata y

⁴⁸ Cf. nota 24.

automáticamente ajustadas a las oportunidades.⁴⁹ Como mencionamos antes, el caso del voto a nivel nacional supone que no obedece a un interés propio ni a sus posibles consecuencias, ya que la posibilidad de que un individuo pueda influir en los resultados finales es prácticamente nula (incluso es más alta la probabilidad de que se sufra un accidente en el trayecto a las urnas).

Finalmente, también hay que considerar a la teoría de la *racionalidad satisfecha* o *limitada* a la que también ya nos hemos referido.⁵⁰ Pero aquí se podría argumentar que la *satisfacción* es una forma de la conducta maximizadora que toma en cuenta algunos elementos que el modelo ortodoxo no considera: los costos (que pueden ser tanto en la recolección y procesamiento de la información pertinente, como en el tiempo que se necesita para ello). Recordemos al médico de urgencias que ante la gravedad del paciente no puede esperar a tener toda la información pertinente sobre el caso para empezar a actuar. Aquí el problema es establecer los niveles de satisfacción y por qué éstos pueden variar entre diversos agentes. Por ejemplo, si bien antes de aceptar un compromiso matrimonial se debería adquirir una mayor información que cuando se va a comprar un automóvil, la pregunta por *cuánta* información se debe reunir antes de comprometerse es incontestable.

Por último, retomemos la insistencia de Elster en relación a que, más allá de los indudables problemas que conlleva la TER, en nuestra vida cotidiana asumimos en general que la gente es racional, se comporta racionalmente o al menos así le gustaría comportarse; tratamos a las personas como si fueran seres racionales y, de hecho, la comunicación presupone la premisa tácita de un mínimo reconocimiento entre los interlocutores; al tratar de

⁴⁹ Cf. Jon Elster, *The Cement of Society. A Study of Social Order*, especialmente capítulo 3 "Social norms", pp. 97-151. Otro autor que ha insistido en la no exclusividad del interés propio como principio explicativo de la conducta humana es Amartya Sen. Este autor destaca por su propuesta de incorporar teóricamente la noción de compromiso sin negar que la evaluación razonada siga siendo la base explicativa de la acción. (Cf. especialmente Amartya K. Sen, "Los tontos racionales: una crítica de los fundamentos conductistas de la teoría económica", en *Filosofía y teoría económica*, comps. Frank Hahn y Martin Hollis, pp. 172-217).

⁵⁰ Cf. nota 13.

entender al otro suponemos que tiene deseos y creencias consistentes y que actúa de acuerdo a ellos.⁵¹

De esta manera, la TER se presenta como una teoría (normativa y explicativa) en la que se defiende una concepción fuerte de racionalidad que no permite imputar irracionalidad a la ligera. Asimismo, a partir de sus críticas internas, se va perfilando la construcción de una propuesta en la que el propio concepto de racionalidad, si bien continúe teniendo un papel privilegiado, ya no sea exclusivo. En el siguiente capítulo nos centraremos en dos casos extremos en donde se presenta como lo más racional la abdicación de la propia razón.

⁵¹ La alternativa a esto, enfatiza Elster, no sería la irracionalidad, que se predicaría en un fondo general de racionalidad, sino propiamente el caos.

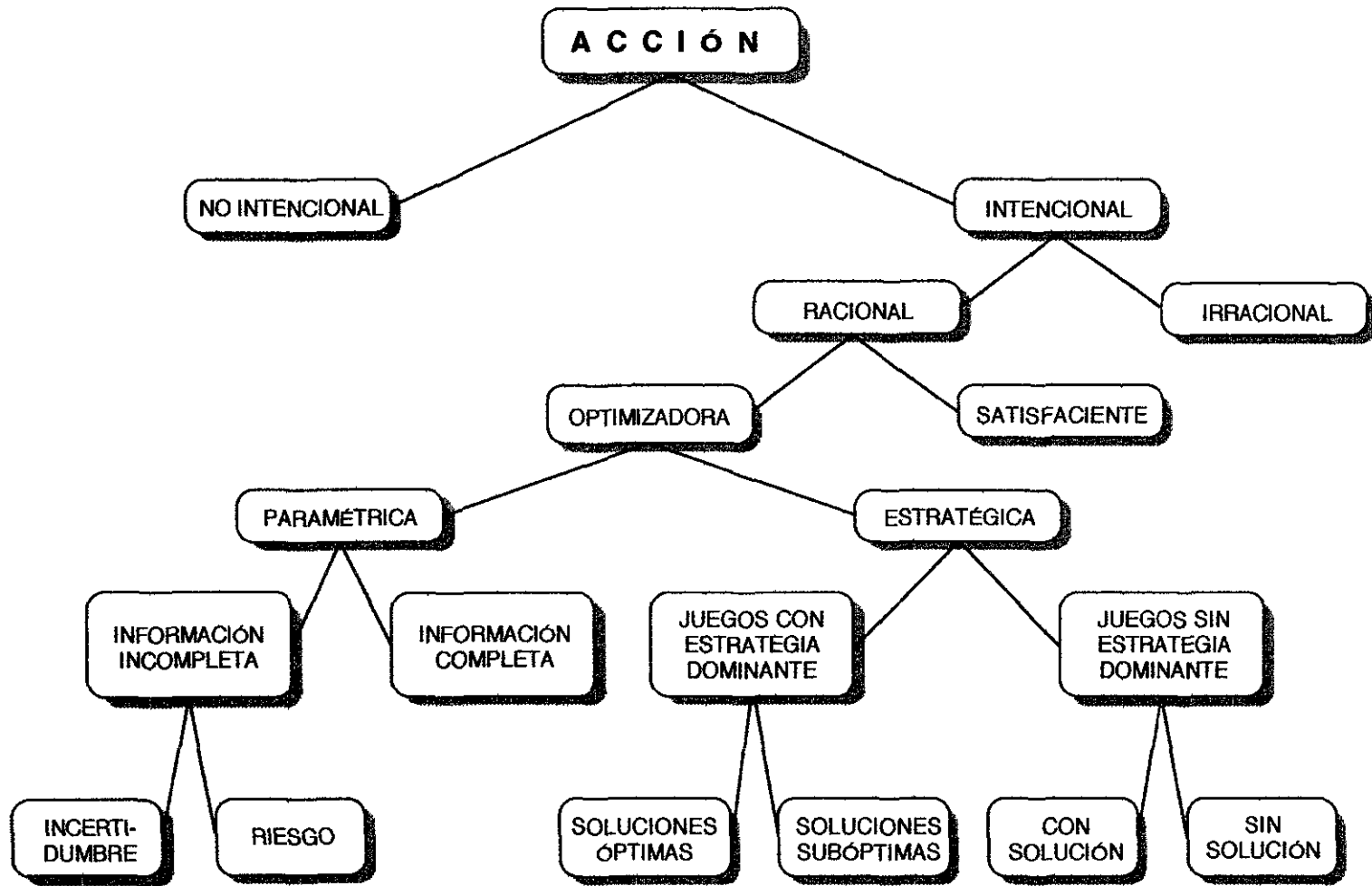


FIG. 1 - RED CONCEPTUAL DEL ANÁLISIS DE LA INTENCIONALIDAD

	NEGACIÓN DE LAS CONDICIONES DE OPTIMALIDAD	NEGACIÓN DE LO NORMATIVAMENTE PRESCRITO
ACCIÓN	<p>por indiferencia entre las mejores alternativas</p> <p>por incapacidad para comparar alternativas</p>	<p>debilidad de la voluntad</p> <p>sobrevaloración y subvaloración de probabilidades</p> <p>conductas mágicas</p> <p>conductas contraproducentes</p>
CREENCIAS	<p>bajo incertidumbre no se pueden ligar las acciones a los resultados</p>	<p>autoengaño</p> <p>desconocimiento de las propiedades estadísticas de la información</p>
EVIDENCIA	<p>dada la imposibilidad de estimar el valor marginal esperado de buscar la información</p>	<p>recolección de más o menos evidencia de la definida por la propia situación</p>

FIG. 2 - PRINCIPALES FALLAS DE LA TEORÍA DE ELECCIÓN RACIONAL

III. ENTRE SIRENAS

En el capítulo anterior ubicamos los principales elementos y problemas de la TER. A partir de ese panorama general nos podemos detener ahora en dos de sus problemas fundamentales. Como dijimos, la racionalidad en la TER es eminentemente normativa y falla cuando no nos indica qué hacer o cuando no hacemos lo que nos indica. Estos dos aspectos marcan los límites a las pretensiones de esta teoría, advirtiendo sobre los peligros de mitificar a la razón. Frente a las posibles *adicciones* a la razón, en este capítulo veremos que entre la racionalidad y la irracionalidad hay espacio para considerar tanto a una racionalidad imperfecta como a una racionalidad indeterminada. De acuerdo con ello, el capítulo está estructurado en dos secciones:

1. En la primera, presentamos el problema de la racionalidad imperfecta el cual surge del hecho de que si bien el ser humano puede actuar racionalmente con frecuencia no actúa así, pero lejos de quedarse tranquilamente en la irracionalidad tiene la capacidad de darse cuenta de ello y tratar de contrarrestarlo.
2. En la segunda sección, nos ocuparemos del problema de la racionalidad indeterminada el cual surge en situaciones de elección que están caracterizadas por la incertidumbre. Frente a la necesidad de elegir entre diversas opciones ya sea bajo incertidumbre o con falta de preferencias, lo más racional es abdicar de la razón.

1. Ulises y las sirenas

Sin duda entre las características distintivas del ser humano destaca su capacidad de maximizar globalmente, sin embargo, no siempre actúa conforme a ello. No es difícil encontrar casos en los que un individuo se abstiene conscientemente de hacer algo que le es beneficioso o incluso que realiza un acto voluntario que lejos de beneficiarle más bien acaba por dañarlo. Pensemos por ejemplo en el ahorro o el ejercicio físico para la primera situación y para la segunda los actos de beber o fumar. Como ya mencionamos en el capítulo anterior, la llamada *debilidad de la voluntad* consiste en obrar contra nuestro mejor juicio o hacer lo que uno cree (considerados todos los factores) que no debería hacer. Este problema puede rastrearse también desde Aristóteles, quien parte del análisis de la opinión común en donde

[...] el hombre, señor de sí mismo, se identifica con el que sigue exactamente las prescripciones de la razón, y el hombre que no tiene dominio sobre sí se identifica con el que se sale de los caminos de la razón; que el que no se domina hace por pasión cosas que él sabe malas; que el hombre que tiene dominio sobre sí, conociendo los malos deseos, se resiste por razón a seguirlos.¹

Para el caso de la akrasia, Aristóteles recurre a la analogía de un cuerpo que ha sufrido parálisis: cuando se quiere mover a la derecha se mueve hacia la izquierda. Lo mismo pasaría en el caso de la incontinencia: los deseos van en un sentido contrario al de la razón. De esta manera, a diferencia del reino

¹ Aristóteles, *Ética Nicomaquea*, 1145a, pp. 1251. Frente a la tesis socrática de que la akrasia libre y consciente es imposible, “el que conoce el bien hace el bien” (posición aun hoy defendida por algunos, cf. por ejemplo D. F. Pears “La akrasia y el razonamiento práctico” en *Segundo Simposio Internacional de Filosofía*, pp. 9-21) podemos recordar también uno de los pasajes de la obra de Eurípides en donde Fedra reconoce: “[...] hace ya largo tiempo que reflexioné una noche en las causas de la corrupción humana, y me parece que no todos los hombres cometen las faltas más graves por sus escasas luces, porque en muchos se observa juicio recto; preciso es, por tanto, confesar que, aun conociendo lo bueno, no lo seguimos, unos por pereza y otros porque posponemos la virtud al deleite.” (Eurípides, *Hipólito* en *Tragedias*, p. 156)

animal que se dirige sólo por el instinto, el hombre es un ser contradictorio en cuyo interior se libra una batalla.

[...] hay aún otra potencia del alma que parece estar privada de razón, a pesar de participar de ella en algún grado. Pues, en el hombre moderado o sobrio y en el intemperante, hacemos caso de la razón, es decir, de la parte del alma dotada de razón. En verdad es ella la que le recomienda precisamente la conducta mejor. Pero, según el sentir común, uno percibe también una especie de instinto que repugna la razón, la combate y le hace frente.²

Dentro de la filosofía contemporánea, destaca nuevamente la propuesta de Donald Davidson a quien se le ha reconocido el haber elaborado una adecuada descripción no sólo del fenómeno sino también de sus condiciones causales de posibilidad.³ De acuerdo con este autor, una acción revela debilidad de la voluntad o incontinencia si cumple con los requisitos de la siguiente caracterización:

Al hacer *x*, un agente actúa de manera incontinente si y sólo si:

- (a) el agente hace *x* intencionalmente;
- (b) el agente cree que hay una acción alternativa y que le es asequible;
- (c) el agente juzga que, hechas todas las consideraciones, sería mejor hacer *y* que hacer *x*.⁴

La preocupación de Davidson por describir la debilidad de la voluntad obedece al hecho de que la existencia de acciones incontinentes pone en crisis una concepción que se presentaba como autoevidente: cuando una persona

² Aristóteles, *Ética Nicomaquea*, 1102b, p. 1184.

³ Cf. Donald Davidson, "How is Weakness of the Will Possible?" en *Essays on Actions and Events*, pp. 21-42.

⁴ Además del carácter intencional de la acción, Davidson enfatiza que la acción alternativa y no necesariamente debe estar disponible, lo relevante es que el agente crea que sí lo es. Como destacamos en el capítulo anterior, las situaciones de elección son eminentemente subjetivas.

actúa intencionalmente, actúa de acuerdo con lo que considera que es lo mejor. En otras palabras, la debilidad de la voluntad socava el hasta entonces indiscutible vínculo entre intencionalidad y racionalidad. Si se acepta esta liga, parecería que no se pueden sostener al mismo tiempo los siguientes principios:

P1. Si un agente quiere hacer x más que lo que él quiere hacer y , y se cree libre para hacer o bien x o bien y , entonces intencionalmente hará x si hace intencionalmente o bien x o bien y .

P2. Si un agente juzga que sería mejor hacer x que hacer y , entonces él quiere hacer x más que lo que él quiere hacer y .

P3. Hay acciones incontinentes.

El trabajo de Davidson se centra en demostrar que la posible contradicción entre los principios P1-P3 es sólo aparente. Como hemos insistido, ya desde Aristóteles se plantea que las acciones intencionales obedecen a los deseos y las creencias del agente. Para tratar de enfrentar el problema de la incontinencia se podría argüir que un deseo y una creencia sobre cuál es la acción que llevará a su satisfacción, no implica de manera *inmediata* la acción misma; por ejemplo nos podríamos quedar simplemente con el juicio de que es *deseable realizar dicha acción* sin realmente llevarla a cabo. Sin embargo, destaca Davidson, esta cuña entre la inferencia de qué es lo mejor y la propia acción tampoco resuelve el problema ya que habría que enfrentar entonces otra pregunta: ¿por qué a veces el deseo y la creencia sí causan una acción, y en otras ocasiones sólo llevan a un juicio sobre qué es lo deseable hacer?

Una de las grandes aportaciones de este autor fue encontrar que los principios P1-P3 pueden dejar de presentarse como contradictorios si se parte de un reconocimiento básico: las razones para actuar son irreductiblemente múltiples y entran en conflicto cuando cada una conduce al agente a acciones

mutuamente incompatibles.⁵ De hecho, hablamos de debilidad de la voluntad cuando alguien sostiene (por una razón) que una acción es la mejor, pero realiza (por otra razón) otra cosa. Esta nueva manera de acercarse al problema de la incontinencia permite abordarlo como un problema de irracionalidad. Por ello, llega a afirmar Davidson:

el *ákrates* no sostiene creencias lógicamente contradictorias ni su falla es, necesariamente, una falla moral. Lo que está mal es que el hombre incontinente actúa y juzga irracionalmente, pues esto, seguramente, es lo que debemos decir de un hombre que va en contra de su propio mejor juicio.⁶

Hay que destacar, sin embargo, que el problema de la debilidad de la voluntad no es propiamente la diversidad de razones, sino el que no se distinguen y jerarquicen los juicios que comprenden el conjunto de las razones. Finalmente, a partir del trabajo de Davidson, un acto incontinente se caracteriza por ser un acto intencional en el que el agente sí tiene una razón para realizarlo, aunque no tiene una razón para no dejar que prevalezca su *mejor* razón. En este sentido, más allá de los diversos motivos que pueda tener cada individuo, un agente racional es el que realiza la acción que juzga como la *mejor* con base en *todas* las razones pertinentes disponibles.⁷

En el capítulo anterior señalamos que las condiciones que caracterizan a la explicación intencional son insuficientes para la explicación de elección racional. Recordemos que a las tres primeras condiciones que corresponden a la explicación intencional, la TER agrega otras cuatro más, la última

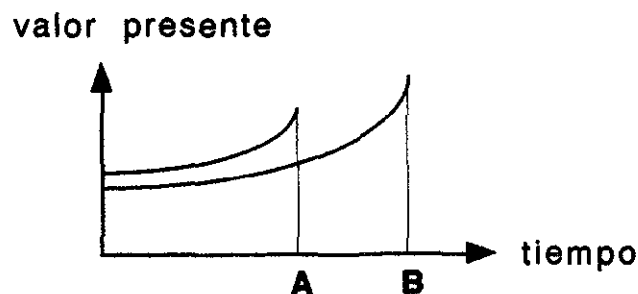
⁵ Es de destacar el que Davidson no reduzca el problema de la debilidad de la voluntad a los casos en donde somos "dominados por la bestia que llevamos dentro", "no escuchamos el llamado del deber" o "caemos en la tentación". De hecho, los logros de su trabajo se deben precisamente a que aborda el problema de la incontinencia desde la filosofía de la acción en general y no desde el ámbito específico de la filosofía moral. Este distanciamiento le permitió reconocer los conflictos sin pretender solucionarlos a partir de un principio moral último o una distinción entre lo deseable *prima facie* y lo absolutamente deseable.

⁶ Donald Davidson, *ibid.*, p. 41.

encaminada especialmente a excluir los casos de conducta akrática. Esta exclusión obedece a que en una acción incontinente se rompen las conexiones causales entre los deseos y las creencias del agente y su acción: los deseos no determinan la conducta en tanto que razones sino más bien, sostiene Elster, como una especie de "turbulencia psíquica". Efectivamente, si en la conducta akrática *prevalece* la razón *más débil*, ésta debe ser en algún sentido la más fuerte, aunque no como razón sino más bien como una excitación emocional o fuerza motivacional. Pasemos a ver esto con más detenimiento.⁸

Hacer y a pesar de que nuestro mejor juicio nos dicta hacer x, puede corresponder a una gran gama de situaciones en las que se manifiesta la debilidad de la voluntad. En especial aquí nos centraremos en uno de los casos analizados por Elster: cuando y es un beneficio menor a corto plazo y x un beneficio mayor a largo plazo.

En primer lugar, hay que aclarar que en sí misma la preferencia por el presente no es una condición necesaria ni suficiente de la debilidad de la voluntad: tomando en cuenta todos los factores, alguien puede preferir el beneficio a corto plazo. Pensemos en una situación que puede ser representada por la siguiente gráfica:



⁷ Para Davidson, esta manera de actuar puede concebirse como un *principio de continencia* que, en contraposición a la irracional debilidad de la voluntad, seguiría todo agente racional.

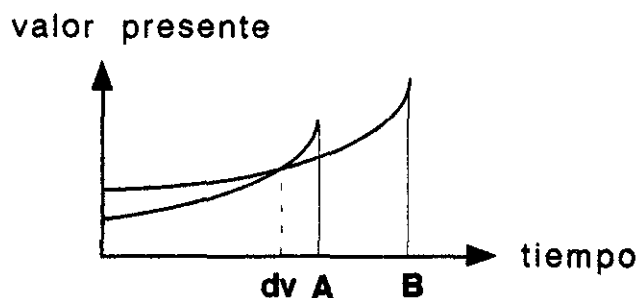
⁸ Cf. Jon Elster, *The Cement of Society. A Study of Social Order*, especialmente capítulo 1 "Collective action", pp. 17-49.

Para simplificar la situación, supongamos que no hay elementos de incertidumbre y que el agente debe tomar la decisión en el momento en que se hace accesible A. Las curvas de la gráfica representan respectivamente el valor presente de A y de B antes de que sean accesibles. Como se puede observar, estos valores se caracterizan por:

- a) declinar exponencialmente en relación a su distancia en el futuro, y
- b) hasta el momento en que se hace accesible A, el valor de A es mayor que el de B

Por ello, a pesar de que puede ser atractiva la opción B en tanto implica un mayor beneficio, si el agente determina sus preferencias con base en los valores presentes, preferirá en *todo momento* la opción A ya que su valor presente es mayor. De esta manera, queda claro que la preferencia por un beneficio menor a corto plazo no es necesariamente un caso de debilidad de la voluntad. Pensemos por ejemplo en un bebedor o un despilfarrador consistente que prefiere, considerados todos los factores, el placer o la diversión del momento que los beneficios futuros de una vida saludable o con cierta estabilidad económica.

Sin embargo, cuando los valores representados en las curvas no cumplen con la primera característica, es decir, no declinan exponencialmente en relación a su distancia en el futuro, sino que lo hacen con mayor rapidez, la situación es otra y nos podemos encontrar frente a un caso de debilidad de la voluntad. Veamos la gráfica siguiente:



La debilidad de la voluntad (dv) está representada por el punto de intersección de las curvas: mucho tiempo antes de tener que elegir se prefiere B pero poco antes de tomar la decisión se cambia la preferencia a A. De esta manera, tenemos que si bien se prefiere (considerando todos los factores) el mayor beneficio a largo plazo, llegado el momento se escoge el menor beneficio a corto plazo. Pensemos ahora en alguien que sí prefiere los beneficios futuros de una vida saludable o con cierta estabilidad económica pero que sucumbe a la "tentación" del momento de escoger el alcohol o el derroche.

Al respecto hay que hacer dos observaciones. En primer lugar, el beneficio mayor a largo plazo puede estar representando la suma de una multiplicidad de beneficios posteriores y no necesariamente uno, como es el caso de tener (mantener) una buena salud. Asimismo, la decisión puede ser una elección que debe tomarse reiteradamente y no sólo una vez, como es el caso del alcohólico que se enfrenta cada día y cada momento a la disyuntiva de beber o no.

Como vimos, las acciones en las que se manifiesta la debilidad de la voluntad se caracterizan por ser acciones intencionales en las que el agente "va en contra de su mejor juicio". En este sentido, a partir del trabajo de Davidson, el problema de la incontinencia se convierte en un problema de irracionalidad. Irracionalidad que, no olvidemos, se predica en un fondo general

de racionalidad.⁹ Al grado de que se pueden dar casos muy peculiares ejemplificados magistralmente por Ulises.¹⁰ Recordemos cómo, después de vivir un año con Circe en la isla Eea, el héroe emprende el viaje de regreso a Itaca y debe superar un primer peligro: la seducción mortal del canto de las sirenas. Para enfrentarlo con éxito, antes de llegar a las islas de las sirenas, Odiseo le dice a sus compañeros de viaje:

No conviene que sean tan sólo uno o dos los que sepan los augurios que Circe me ha hecho, la diosa divina; os los voy a contar para que, conociéndolos, todos perezamos o bien evitemos la muerte y la parca. Me ordenó lo primero que de las Sirenas divinas rehuyamos la voz y el florido pradal en que cantan. Solamente yo puedo escucharlas, más es necesario que me atéis fuertemente con lazos de nudo difícil, de pie al lado del mástil y se aten al palo las cuerdas. Si a vosotros suplico y ordeno soltéis tales nudos deberéis, todavía, con muchos más nudos atarme.¹¹

Para Elster, en este pasaje de la *Odisea* en donde Ulises pide a sus acompañantes que lo aten para no dejarse llevar por el canto de las sirenas, se puede observar la necesidad de una teoría de la *racionalidad imperfecta*:

Ulises no era por completo racional, pues un ser racional no habría tenido que apelar a este recurso; tampoco era, sencillamente, el pasivo e irracional vehículo de sus cambiantes caprichos y deseos, pues era capaz de alcanzar por medios indirectos el mismo fin que una persona racional habría podido alcanzar de manera directa.¹²

Aquí lo relevante es que Ulises no sólo es débil, sino que sabe de su debilidad y por ello puede hacer algo para enfrentarla. Ante un problema de incontinencia, el agente tiene la posibilidad de encontrar una solución racional:

⁹ Cf. la última nota del capítulo anterior.

¹⁰ Para la exposición general de esta sección cf. Jon Elster, *Ulysses and the Sirens. Studies in rationality and irrationality*, especialmente el capítulo II "Imperfect Rationality: Ulysses an the sirens", pp. 36-111.

¹¹ Homero, *Odisea*, canto XII, p. 195.

¹² J. Elster, *Ulysses and the Sirens*, p. 36.

atarse a sí mismo. Con esta estrategia, un individuo se *compromete previamente* actuando sobre el medio en el que se encuentra con el fin de reducir sus opciones disponibles: Ulises manda que lo aten para que, llegado el momento, no pueda moverse aunque quisiera. *Atarse a sí mismo* es una manera privilegiada de resolver el problema de la debilidad de la voluntad ya que por medios indirectos se logra actuar "de acuerdo con el mejor juicio", es decir, racionalmente.¹³ Pensemos en un fumador que para ayudarse a dejar de fumar destruye todos los cigarros que tiene a la mano o un comprador compulsivo que con el fin de dejar de comprar cosas superfluas se asegura de ir al mercado sólo con el dinero justo para adquirir lo necesario.

Entre los elementos básicos para poder bosquejar los lineamientos generales de la idea de *atarse a sí mismo*, destacan las siguientes condiciones:

- i) *Atarse a sí mismo* es llevar a cabo cierta decisión en el tiempo t_1 para aumentar la probabilidad de llevar a cabo otra decisión en el tiempo t_2 .

Con esta condición se establece que la finalidad con la que se realiza una primera acción es producir un cambio en las probabilidades de una acción posterior. Como ya señalamos, esto no necesariamente implica una sola decisión, puede darse a partir de una diversidad de decisiones.

¹³ Obviamente, *atarse a sí mismo* no es la única manera de enfrentar el problema de la debilidad de la voluntad. En vez de modificar el medio exterior se puede tratar de modificar el "espacio interior" de la persona como es el caso de recurrir a apuestas privadas indirectas, a la planificación del carácter y a la responsabilidad para con uno mismo. Sin embargo, más allá de los problemas que conllevan cada una de estas estrategias, sin duda cuando se actúa sobre el medio exterior los resultados son más duraderos ya que se elimina el margen de decisión del propio agente. Lo cual, por otra parte, también puede tener sus inconvenientes: estando Ulises atado y sus marineros con cera en los oídos, difícilmente hubieran podido hacer frente a un imprevisto.

ii) El acto en t_1 tiene el efecto de provocar un cambio en el conjunto de opciones que se dispondrá en t_2 . Si el nuevo grupo factible incluye al anterior, entonces no cuenta como *atarse a sí mismo*.

Con el compromiso previo se pretende aumentar las probabilidades de que se realice una determinada acción en el futuro, por ello la acción realizada debe reducir el grupo factible inicial al excluir algunas de sus opciones para el futuro.

iii) El llevar a cabo la decisión en t_1 debe tener como efecto el desencadenar algún proceso causal en el mundo exterior.

Precisamente porque el sujeto reconoce su debilidad de la voluntad, trata de eliminar su margen de decisión en t_2 . En este sentido, con esta estrategia el agente deposita temporalmente su voluntad en alguna estructura externa: al modificar el entorno se espera que éste modifique el comportamiento.

iv) La resistencia a llevar a cabo la decisión en t_1 debe ser menor que la resistencia que se habría opuesto a llevar adelante la decisión en t_2 , si no hubiese intervenido la decisión en t_1 .

Como hemos insistido, el ser humano se caracteriza por maximizar globalmente, es decir, tiene la capacidad de aguardar y utilizar estrategias indirectas (dar un paso atrás para poder dar después dos pasos adelante). El compromiso previo con el que se pretende contrarrestar la debilidad de la voluntad, se puede entender como una estrategia indirecta (metaestrategia) que permite alcanzar los resultados que de manera directa no podríamos o sería más difícil obtener.

v) El acto de *atarse a sí mismo* debe ser un acto de comisión, no de omisión.

El compromiso previo es una respuesta intencional y racional frente al problema de la incontinencia. *Atarse a sí mismo* implica un acto que voluntaria y causalmente está dirigido a impedir que en el futuro se lleve a cabo un acto de debilidad de la voluntad.

Con este breve acercamiento a la estrategia de *atarse a sí mismo*, queremos destacar cómo los hombres no son *ni ángeles ni animales*, es decir, se reconoce que no son plenamente racionales aunque tampoco fundamentalmente miopes. Desde esta posición poco ortodoxa dentro del ámbito de la TER, se parte de un hecho básico: los hombres tienen la capacidad de enfrentarse estratégicamente a su propia miopía, en palabras de Elster, son *imperfectamente racionales*.

Como dijimos, el *atarse a sí mismo* es una estrategia privilegiada para contrarrestar la debilidad de la voluntad ya que se centra en la modificación del medio exterior con el fin de eliminar el margen de decisión del propio agente. En este sentido, se trata de una estrategia en la que *racionalmente se abdica de la razón*: como no confiamos que en su momento tomaremos la decisión correcta, decidimos previamente quitarnos la opción de decidir. Esta estrategia indirecta se puede usar en muy diversas situaciones y con una gran variedad de mecanismos; asimismo, recurren a ella no sólo los individuos sino también los grupos sociales. Veamos un ejemplo de cada caso.

En el capítulo anterior destacamos que los juegos acompañan a toda sociedad, al grado de justificarse la expresión *Homo ludens*. Sin embargo, el placer por el juego se puede volver un problema grave, igual que el alcoholismo o el tabaquismo. Como un caso peculiar de uso de la estrategia

atarse a sí mismo para enfrentar individualmente la debilidad por el juego, encontramos en un periódico español la siguiente nota:

Casi cinco mil personas solicitaron el pasado año [1997] a la Consejería de Gobernación y Justicia que les fuera prohibida la entrada en los bingos o casinos de Andalucía. Para ello, fueron inscritos en el listado de la Dirección General de Juego de la Junta, entidad que remite una vez al mes esta lista a los bingos y casinos, para que puedan identificar a sus clientes y nieguen la entrada a las personas que se encuentran en ella.¹⁴

Efectivamente, en Andalucía cualquier individuo puede voluntariamente autoprohibirse la entrada a los locales de juego inscribiéndose directamente en un enlistado que para tal efecto dispone la Dirección General de Provincias del Gobierno de la Junta de Andalucía o levantando una "autodenuncia" ante la policía local del municipio que le corresponda. La presencia en estos locales de un ciudadano que esté incluido en el *registro de los prohibidos* puede acarrearle al establecimiento una multa de hasta cinco millones de pesetas. Cabe aclarar que esta autoprohibición puede ser levantada por el propio solicitante cuando lo desee.

Por lo que se refiere al caso de las sociedades que se *atan a sí mismas*, el ejemplo por excelencia son las constituciones:

Las constituciones son cadenas con las que los hombres se amarran a sí mismos en los momentos de lucidez para no morir de mano suicida en el día de la locura.¹⁵

Las asambleas constituyentes tienen la función excepcional y privilegiada (aunque no por derecho sino por "accidente" histórico) de definir las reglas básicas que deberán seguir las generaciones futuras. En situaciones

¹⁴ ABC, 15 de septiembre de 1998, p. 48.

¹⁵ Palabras de John Potter Stockton en el debate sobre el comportamiento del Ku Klux Klan de 1871, citado en Jon Elster, *Argomentare e negoziare*, p. 22.

históricas excepciones e impredecibles, la ruptura radical con el pasado deja libre el camino para amarrar el porvenir. En este sentido, a través de una asamblea constituyente, una “nación” puede *atarse a sí misma* confiando ciertos poderes a la rama judicial y estableciendo que dichas reglas sólo pueden ser cambiadas por una mayoría calificada.¹⁶ Incluso se pueden plantear límites a la propia regla de la mayoría, como es el caso de las llamadas *democracias protegidas*: la regla de la mayoría (que es una de las reglas del juego democrático que debe ser aceptada por todo el grupo) no puede ser derogada por mayoría.¹⁷

En esta sección hemos visto que para Elster el caso de Ulises y la estrategia de *atarse a sí mismo* exige incorporar en la TER una teoría de la *racionalidad imperfecta*. Pero éste es sólo un primer paso por el que, poco a poco, la sobrevalorada racionalidad perfecta con la que inició el paradigma de elección racional va adquiriendo su justa dimensión. Por ahora podemos terminar este apartado con una de las conclusiones del autor que retomaremos en el siguiente capítulo:

Una plena característica de lo que significa ser humano debiera incluir al menos tres rasgos. El hombre puede ser *racional*, en el sentido de sacrificar deliberadamente la gratificación actual por la gratificación futura. El hombre a menudo no es racional, y antes bien muestra *flaqueza de voluntad*. Aun cuando no es racional, el hombre sabe que es irracional y puede *atarse a sí mismo* para protegerse contra la irracionalidad. Esta “segunda mejor o imperfecta” racionalidad se ocupa a la vez de la razón y de la pasión. Lo que se pierde, tal vez, es el sentido de la aventura.¹⁸

¹⁶ Como menciona Elster, Lally-Tolendal en la Asamblea Constituyente de París (1789-1791) sostenía: “cambiar [la Constitución] no debe ser ni fácil ni imposible”. (*Ibidem.*)

¹⁷ Cf. Norberto Bobbio, “La regola della maggioranza e i suoi limiti” en *Soggetti e potere. Un dibattito su società civile e crisi della politica*, ed. V. Dini. Aquí Bobbio insiste en la necesidad de poner límites a la regla de la mayoría: “Que la democracia pueda ser asesinada por la misma democracia puede ser desagradable pero no es por ello menos realista.” (*Ibid.*, p. 23)

¹⁸ Jon Elster, *Ulysses and the Sirens*, p. 111. En el siguiente capítulo nos ocuparemos especialmente del problema de la relación entre razón y pasiones.

2. Las sirenas de la razón

En el capítulo anterior mencionamos dos grandes limitaciones de la TER. (Fig. 2) Por un lado, planteamos el problema de la indeterminación, es decir, cuando la teoría no puede proporcionar predicciones singulares, recordemos el caso extremo del juego de la gallina. Por otro, la impropiedad, es decir, cuando la teoría sí proporciona predicciones pero falla porque los individuos actúan irracionalmente, como es el caso de la debilidad de la voluntad. Estas dos grandes limitaciones de la TER pueden a su vez estar ligadas, de hecho, el problema de la indeterminación puede llevar al de la impropiedad. En esta sección veremos que el no reconocimiento de la indeterminación de la TER puede llevar a una conducta irracional.¹⁹ Básicamente, como sintetiza Elster, el problema es que en general

sentimos una gran renuencia a admitir la incertidumbre y la indeterminación en los asuntos humanos. Antes que aceptar los límites de la razón, preferimos los rituales de la razón.²⁰

Rituales de la razón que incluso la difusión de la TER ha reforzado al simplificar en extremo tanto las presunciones sobre los agentes como del mundo en el que actúan.

Para empezar recordemos la maravillosa paradoja del Asno de Buridán, planteamiento que sin figurar en sus escritos se atribuye al rector de la Universidad de París del siglo XIV. El problema es el siguiente: un asno frente a dos haces de heno esencialmente iguales y a la misma distancia, no tendría ninguna razón para preferir uno más que el otro, y, al no poder escoger entre

¹⁹ Para la exposición general de esta sección cf. Jon Elster, *Solomonic Judgements. Studies in the Limitations of Rationality*, especialmente el capítulo II "Taming Chance: Randomization in individual and Social Decision", pp. 36-122.

²⁰ Jon Elster, *Ibid.*, p. 37.

ellos, acabaría por morir de hambre.²¹ Aunque se trata de una situación empíricamente imposible, esta paradoja plantea el problema de que no puede haber una elección sin preferencias.²² Sin embargo, frente a este planteamiento se podría argüir que el asno finalmente sí tiene una preferencia que no está contemplada, -no morir de hambre-, la que lo llevaría a comer *cualquiera* de los dos haces de heno. Otra posibilidad es aceptar que no todas las elecciones deben ser siempre razonables (con base en las preferencias). Retomando la sugerente expresión de Sen, podríamos decir que si el asno muere de hambre por no poder elegir entre dos haces de heno, seguramente era un gran *tonto racional*.²³

Pero la indeterminación no sólo puede surgir por la ausencia de preferencias. Ya Maquiavelo tuvo que reconocer el lugar que ocupa la fortuna en los asuntos humanos:

Muchos han creído y creen todavía que las cosas de este mundo las dirigen la fortuna y Dios, sin ser dado a la prudencia de los hombres hacer que varíen, ni haber para ellas remedio alguno; de suerte que, siendo inútil preocuparse por lo que ha de suceder, lo mejor es abandonarse a la suerte. En nuestra época han acreditado esta opinión los grandes cambios que se han visto y se ven todos los días, superiores a toda humana previsión. Meditando en ellos me han hecho a veces inclinarme algo en favor de esta creencia; sin embargo, como nuestro libre arbitrio existe, creo que de la fortuna dependa la mitad de nuestras acciones, pero que nos deja dirigir la otra mitad o algo menos.²⁴

²¹ Cf. Nicholas Rescher, "Choice Without Preference. A Study of the History and of the Logic of the Problem of 'Buridan's Ass'", en *Kant-Studien*, vol. 51, No. 2, 1959-1960, pp. 142-175, quien ha rastreado la historia de esta paradoja desde el mundo griego y que de alguna u otra forma está implícita en la mayor parte de los debates entre deterministas e indeterministas, ya sea en cosmología, teología o ética.

²² El problema son las no-preferencias del hipotético asno y no las no-diferencias de los haces de heno. De hecho, el caso puede aplicarse a objetos diferentes pero cuyas diferencias son irrelevantes respecto a su deseabilidad o que no se conocen al momento de elegir.

²³ Obviamente, en la paradoja del Asno de Buridán, la raíz del problema de la elección sin preferencias es un problema teórico, no práctico.

²⁴ Nicolas Maquiavelo, *El príncipe* en *Obras Completas*, p. 355. Recordemos que para Maquiavelo hay dos formas de ascender a príncipe: por genio o por fortuna. En relación a esta última advierte: "A los particulares que únicamente a la fortuna deben el llegar a ser príncipes,

A la memoria también viene la célebre expresión atribuida a Julio César después de cruzar el río Rubicón que marcaba el límite de su provincia:

Marchemos a donde nos llaman los signos de los dioses y la iniquidad de los enemigos. *Jacta alea est.*²⁵

Parece que los logros de un individuo se construyen a partir de una combinación entre esfuerzo y suerte. Pero en general, para bien o para mal, el azar y la contingencia juegan un papel crucial en la vida de toda persona. Más allá de nuestras mejores creencias y rigurosos cálculos, muchos de nuestros actos dependen de factores ajenos a nuestro control. Los principales acontecimientos que marcaron (en retrospectiva) la dirección de una vida son frecuentemente fortuitos. Aquí tal vez sea conveniente hacer una distinción entre fortuna y suerte.²⁶ Ambas, en positivo o negativo, están ligadas a acontecimientos fortuitos que afectan nuestros intereses, pero se pueden distinguir por la *previsibilidad* racional que supone la primera y que en cambio excluye la segunda. De esta manera, a diferencia con la fortuna, sería un contrasentido esperar "golpes de suerte". Pongamos dos ejemplos. Se ha calculado que para sufrir un accidente aéreo, se debería abordar durante cuatro mil años un determinado vuelo; por ello, no tenemos suerte cuando aterrizamos sin novedad, aunque tendríamos muy mala suerte si nuestro avión se estrellase. En el caso de una rifa, el ganador tiene suerte y el que pierde sólo es desafortunado.²⁷

cuéstaes poco trabajo ascender, pero mucho el mantenerse: suben sin ningún obstáculo y llegan pronto; pero, al llegar, empiezan los inconvenientes." (*Ibid.*, pp. 315-316)

²⁵ Cayo Suetonio, *Los doce Césares*, vol. 12, p. 41.

²⁶ Cf. Nicholas Rescher, *La suerte. Aventuras y desventuras de la vida cotidiana*.

²⁷ Rescher hace también dos aclaraciones. Por un lado, la no previsibilidad de la suerte no está ligada necesariamente a bajas probabilidades, más bien a que los acontecimientos sean favorables o desfavorables por *casualidad*. Si alguien juega a la ruleta rusa con una sola bala, a pesar de las probabilidades a favor, tendrá suerte si vive para contarle ya que sólo por casualidad no le tocó la cámara cargada. Por otra parte, un resultado negativo producto de la

Lo imprevisto que caracteriza a la suerte, está ligado a la ignorancia del afectado. En este sentido, Dios en tanto ser omnisapiente no podría tener ni buena ni mala suerte. A esto hay que agregar que se trata de un azar que no es necesariamente objetivo. Pensemos por ejemplo en el caso de un camino que de repente se divide en tres y no sabemos a dónde conduce cada uno. Es poco probable que escojamos el camino correcto, si lo hacemos sería sólo por casualidad y diríamos que hemos tenido suerte (aunque no sea por casualidad que cada uno de los caminos conduzca a donde lo hacen).

El indudable papel que tiene lo impredecible y lo azaroso en nuestras vidas, tiene como trasfondo precisamente la previsibilidad y la determinación. En este sentido, la razón no necesariamente se contrapone a la suerte, se puede reconciliar con ella si es capaz de reconocer sus propios límites. En otras palabras, no se trata de escoger entre la racionalidad y lo fortuito, sino más bien de reconocer que *no todo lo real es racional*. Incluso, podríamos recordar que como especie somos el resultado de una evolución altamente improbable.

Pero el problema es que normalmente no nos gusta enfrentarnos a la incertidumbre. Cada día salimos de casa confiados en que regresaremos y el hecho de concluir un día más lo vemos como algo "natural", aunque tal vez sea cuestión de buena suerte. Por lo general, destaca Elster, tratamos de evitar la incertidumbre porque nos impediría no sólo hacer planes a futuro sino también realizar una elección que podamos justificar como un acto basado en la razón.²⁸ Por ello, frente a esta extendida inclinación, se ha interesado por analizar los casos en donde se recurre por ejemplo a loterías para la toma de decisiones bajo incertidumbre, ya que presupone una inusual disposición a

falta de habilidad o perseverancia del agente no es cuestión de mala suerte sino de mala fortuna ya que es un resultado previsible. Pensemos en un estudiante que no estudió lo suficiente para un examen. Aunque también se pueden dar casos mixtos, como es por ejemplo el caso de un conductor imprudente que sufre un accidente inusitado

admitir los límites de la razón: ante la falta de razones para hacer una elección se opta por elegir una de las opciones posibles por medio del azar. En este sentido, el asno de Buridán que no tiene ninguna razón para elegir entre los dos montones de heno, bien podría recurrir a una moneda al aire para decidir con cuál empezar a comer.

Cuando se usan las loterías para resolver problemas de decisiones bajo incertidumbre, se presupone normalmente que éstas son justas, entendiéndose por ello que son aleatorias e imparciales. Más allá del problema de que se pueda disponer de un instrumento intrínsecamente aleatorio -es decir, en el que todos los posibles resultados tengan la misma probabilidad de salir-,²⁹ la aleatoriedad intrínseca no es necesaria ni tampoco suficiente para considerar que una lotería es justa. Más que la aleatoriedad intrínseca, lo importante es que el instrumento al que se recurre para realizar la lotería sea considerado imparcial. Por lo que se refiere en especial a la imparcialidad, uno de los principales problemas es que, cuando no se sabe el procedimiento con el que se genera la aleatoriedad, sólo se le puede juzgar a partir de los resultados. En este juicio interviene normalmente la difundida creencia de que la secuencia obtenida mediante un proceso aleatorio tiene más probabilidades de ser una secuencia intrínsecamente aleatoria que una secuencia regular. Pero parece que es justo al revés, es más probable que de un proceso aleatorio surja una distribución uniforme. Entre los ejemplos que proporciona Elster podemos mencionar el patrón de los bombardeos alemanes en Gran Bretaña durante la

²⁸ Aunque a veces también podemos apreciar la incertidumbre e incluso propiciarla. Elster pone el ejemplo extremo de imaginarnos cómo cambiarían nuestras vidas si pudiéramos saber su inevitable desarrollo y desenlace. (Cf. Jon Elster, *Solomonic Judgements*, pp. 38-39)

²⁹ La aleatoriedad intrínseca parece ser altamente problemática. Incluso, ningún generador de números por medio de una computadora es "realmente aleatorio": funciones aritméticas predeterminadas son las que le permiten generar la secuencia a partir de un número pseudoaleatorio llamado "semilla" (por ejemplo la hora). Parece que la única manera de obtener un valor "realmente aleatorio" es recurriendo a la medición del *spin* de un electrón de un átomo tomado al azar, el cuál, independientemente de su traslación alrededor del núcleo del átomo, puede girar sobre sí mismo con igual posibilidad en positivo o negativo (con o contra las manecillas del reloj).

Segunda Guerra Mundial, azar que para un lego aparece como una regularidad o tendencia a los cúmulos.³⁰

Frente a la diversidad de problemas que pueden surgir en la realización de una lotería como medio para la toma de decisiones bajo incertidumbre, puede ser suficiente la aleatoriedad epistémica. En otras palabras, aunque el mecanismo al que se recurra no sea objetivamente aleatorio, puede ser aceptado como imparcial si de acuerdo a lo que sabemos, todos los resultados son igualmente probables. Incluso en el caso extremo de recurrir a un dado cargado, si no sabemos cuál es su tendencia, todos los números seguirían teniendo la misma probabilidad de salir, ningún número podría tener más probabilidades que los otros. De hecho, el azar epistémico es fundamental en los *sorteos naturales*, es decir, cuando se recurre a un acontecimiento que es aleatorio para los interesados en tanto no tienen información ni control sobre de él.³¹ Pensemos por ejemplo en la posibilidad de escoger una de las tres sendas a partir del vuelo de la primera ave que aparezca. Aunque se puede llegar a desagradables sorpresas, como es el caso del juez hebreo Jefté quien le prometió a Yahve que, si le permitía ganar la batalla contra los ammonitas, le sacrificaría a la primera persona que apareciera regresando a casa: la primera persona que vio fue su única hija que le salía a su encuentro para festejar la victoria.³²

Hay que aclarar que no todas las loterías deben partir de un proceso aleatorio (ya sea objetiva o epistémicamente) equiprobabilístico, es decir, en el que todos los resultados tienen las mismas probabilidades de salir. Aunque ésta sea la norma, puede haber excepciones a través de las cuales se

³⁰ Aunque también puede ser a la inversa, es decir, si no se aprecia regularidad suponer que es el resultado de un proceso aleatorio. Pero en general, dada una sucesión no es posible determinar si se ha generado mediante un proceso aleatorio o no aleatorio.

³¹ Obviamente cuando se recurre a los sorteos naturales para dirimir intereses enfrentados, podría haber problemas en relación a que efectivamente el azar epistémico es para todos los interesados. En estos casos uno puede proponer el procedimiento y el otro tendría el derecho a establecer a cada resultado una de las decisiones posibles.

³² Cf. *Biblia de Jerusalén*, Jueces 11:29, p. 284.

pretenda favorecer o castigar a algún grupo participante. Entre los ejemplos de sorteos calibrados destaca el procedimiento a partir del cual se llevó a cabo un reparto de tierras en Georgia en 1832: todos los ciudadanos tenían derecho a una oportunidad, pero un grupo favorecido (huérfanos, veteranos de guerra y padres de familia) tenía dos oportunidades. Con este tipo de sorteo se trataba de ayudar o retribuir a ciertos sectores especiales, aunque tampoco se quería que el nuevo asentamiento estuviese integrado exclusivamente por ellos.

Por último, hay que destacar que a pesar de que las loterías pueden ser un excelente instrumento para la toma de decisiones bajo incertidumbre, a lo largo de la historia con frecuencia se han usado para descubrir la *voluntad de Dios*. Al menos desde el Antiguo Testamento hasta la temprana edad moderna se hacía presente la sentencia de que “Se echan los suertes en el seno, pero la decisión viene de Yahve”.³³ Obviamente, en estos casos el resultado del sorteo no es considerado como un suceso aleatorio sino más bien como un acto intencional divino.³⁴

Más allá de la cuestión religiosa, pasemos a ver cuándo se ha recurrido a las loterías para la toma de decisiones. En primer lugar, parece que se recurre a las loterías más frecuentemente en las democracias o estados populares, aunque no sea el procedimiento de selección más usual (sobre todo en las democracias modernas en donde claramente predominan las elecciones sobre los sorteos). Desde la democracia ateniense hay casos en donde se recurre a la selección por sorteo dada la imposibilidad práctica de que todo el mundo participe en los asuntos cotidianos de gobierno. Asimismo, predomina el uso de las loterías para la asignación de tareas administrativas, legales o políticas o para la asignación de tareas o cargas al pueblo (es menos habitual para la asignación de recursos escasos). Ejemplos de estos son la selección de

³³ *Biblia de Jerusalén*, Proverbios 16:33, p. 877.

³⁴ Incluso, señala Elster, se podría aventurar que el sorteo, por ejemplo para la selección de funcionarios en la antigua Atenas, tuvo probablemente un origen religioso aunque después se

jurados o de soldados en Estados Unidos. En México, es conocida la asignación supuestamente aleatoria para participar en las mesas de votaciones en periodos electorales así como el sorteo que exime en cada generación a algunos de sus miembros de realizar el servicio militar.

La razón fundamental para recurrir a las loterías es la indeterminación. Si los candidatos son considerados iguales entonces se puede echar a la suerte. Incluso, frente a la igualdad de votos o de puntos (“bondad” medida “objetivamente”) se puede recurrir a las loterías para superar el empate. Por ejemplo en el caso de candidatos políticos o el ingreso de estudiantes. Aunque también se puede recurrir a las loterías cuando los costos de una selección exhaustiva son superiores a los beneficios de elegir al mejor. En este caso, las selecciones obedecen más a los beneficios sociales que a las necesidades o méritos de los solicitantes. Entre los ejemplos más socorridos se encuentra el de la custodia de un menor, en donde se debe sopesar cuáles son para el niño los costos y beneficios de una selección exhaustiva, pasando a segundo término los derechos o necesidades de los padres.

Las loterías también se utilizan en situaciones extremas, cuando ni siquiera es posible hacer el cálculo de costos y beneficios dada la inconmensurabilidad del caso. Se trata de situaciones en donde la comparación ya no es costosa o difícil sino imposible o poco fiable. A nivel individual destacan los casos en donde el agente no tiene un orden de preferencias completas o no puede asignar probabilidades a los resultados de sus diversas acciones. A nivel social destaca el problema de las comparaciones interpersonales; recordemos el teorema de Arrow en el que se cuestiona la posibilidad de construir un orden social que respete las ordenaciones individuales: no podemos reunir las comparaciones

haya vuelto una institución laica. Por otra lado, a pesar que desde el siglo XII la doctrina cristiana oficial no favorece esta interpretación de las loterías, tampoco pudo eliminarla.

interpersonales en una ordenación consecuyente pretendiendo que ésta sea la comparación social.

Pero el hecho de que se recurra a las loterías para la toma de decisiones no significa que aquellas sean racionales o moralmente necesarias. De hecho, señala Elster, si los candidatos son igualmente buenos se podría recurrir a otros métodos de designación escogiendo alguna cualidad o característica: el más feo, el más alto, el más gordo, etc. Sin embargo, normalmente se prefieren las loterías por su preponderancia en situaciones de indeterminación ya que son simples, mecánicas y universalmente aplicables. Por otra parte, las propiedades que se podrían escoger para seleccionar a los candidatos están a su vez relacionadas con otros criterios que podrían no quererse utilizar, por ejemplo que las personas bien parecidas ganen más llevaría a una desviación innecesaria de la igualdad; asimismo la elección de personas feas puede llevar a reforzar actitudes irracionales que consideran esta característica como una minusvalía que justifica una compensación. En cambio, la incertidumbre de las loterías puede tener un efecto de incentivo en el sentido de que la ignorancia del futuro puede llevar a mantener un comportamiento socialmente útil (aunque también puede llevar a un comportamiento contrario dependiendo de la honestidad y complejidad de la organización social). De hecho, estos efectos incentivadores pueden justificar el uso de loterías incluso cuando se dispone de criterios racionales explícitos. Por ejemplo, para la selección de una especialidad en Medicina parecería que las personas con experiencia laboral deberían tener una prioridad de admisión, sin embargo también se podría recurrir a la loterías para evitar que las personas pierdan años de estudio acumulando puntos.

La incertidumbre de las loterías evita dos tipos de efectos incentivadores indeseables: las decisiones discrecionales y las elecciones predeterminadas. Por ejemplo, si no se recurre a una lotería para establecer la fecha de unas elecciones y la decisión la toman quienes pueden beneficiarse de ellas,

escogerán las fechas que maximice sus posibilidades de victoria; incluso si se recurre a un procedimiento imparcial para tomar precauciones sobre ello, el gobierno por ejemplo seguramente aplicará su política económica en función de dicha fecha.

Los efectos incentivadores de la aleatoriedad se producen en distintos niveles: los funcionarios no pueden usar su poder discrecional para beneficiar a sus favoritos o castigar a sus enemigos; así como tampoco los elegidos o receptores potenciales pueden sobornar o amenazar a los funcionarios. Sin embargo, por otra parte, la incertidumbre también puede llevar a la ineficiencia: si es por sorteo nadie tiene el incentivo de invertir tiempo y esfuerzo preparándose por ejemplo para unos cargos o unas becas que se asignan aleatoriamente. En este sentido, la asignación por loterías puede acabar teniendo más aspectos negativos que positivos.

Otro valor de las loterías es que fomentan la justicia: cuando no hay diferencias significativas entre los candidatos se utiliza una lotería ya que la alternativa, es decir, utilizar diferencias irrelevantes, sería injusto. Aquí, por justicia se entiende que los casos iguales deben tratarse de igual manera; e incluso se puede dar un paso más y sostener que, aun cuando existan diferencias importantes, a las personas hay que tratarlas por igual, es decir, a pesar de las diferencias relevantes, todas las personas por ejemplo son igualmente dignas. Aunque no necesariamente hay que aceptarlo en estos términos, también se puede defender el recurrir a loterías ponderadas con base en un criterio *maximin*: se pueden aceptar oportunidades desiguales si con ello se aumentan las oportunidades de los que están en una peor situación. Este tipo de loterías puede llevar a aumentar las probabilidades de todos para obtener algún bien, es decir, puede haber desigualdades que al final hacen que los bienes sean menos escasos (por ejemplo ingreso a estudios médicos). Pero la productividad no es la única razón para desviarse de la igualdad. Se pueden hacer distinciones por la necesidad, contribuciones y merecimientos,

además de que en cada caso se pueden establecer también loterías ponderadas estratificadas (por ejemplo en la repartición de un órgano tomar la expectativa de vida como un criterio para ponderar probabilidades: un anciano uno de cinco y un joven cuatro de cinco).³⁵

Si bien en algunos casos de toma de decisiones puede haber buenas razones para recurrir a las loterías, ¿por qué se utilizan tan poco? Tal vez porque, como insiste Elster,

Las elecciones visiblemente arbitrarias suelen repelernos. Incluso cuando no tenemos razón alguna para decidir un procedimiento u otro, nos gustaría que el resultado estuviese determinado por razones.³⁶

Para tener ambas cosas a la vez, se puede vincular la elección a la causalidad natural con la esperanza de que ésta refleje algún fin o destino, la voluntad de Dios, o la interconexión natural entre todas las cosas. De hecho, como ya señalamos, a lo largo de la historia se ha dado un gran solapamiento entre las loterías y los procedimientos adivinatorios, desde la oración hasta la astrología pasando por la brujería. Aunque los teólogos establezcan distinciones entre estas diversas prácticas -algunas de ellas se consideradas legítimas, pero la gran mayoría se consideran blasfemas y supersticiosas-, en el acervo popular van juntas y forman parte de la creencia de que el universo no es aleatorio y es posible desvelar sus secretos. El objetivo de estas técnicas era tanto cognoscitivo como práctico. Ante un mundo de incertidumbre y de miseria, la gente no sólo deseaba conocer las causas de sus desventuras sino también qué hacer ante ellas. Parece que la idea de que el sufrimiento puede acometernos de manera ciega y azarosa resulta difícil de soportar; y aunque lo

³⁵ Aunque aquí no nos ocuparemos de ello, hay que destacar que la justicia es uno de los grandes temas de la filosofía. Se trata de un término problemático que tampoco se resuelve simplemente con el recurso a las loterías ponderadas, en este caso ¿sería justo elegir al anciano en lugar del joven?.

³⁶ Jon Elster, *Domar la suerte*, p. 145.

más satisfactorio pueda ser creer que existe alguien a quien culpar de nuestras desventuras, en última instancia es preferible pensar que los culpables somos nosotros mismos a creer que no hay nadie a quien culpar (incluso se ha recurrido a la idea de vidas pasadas). Esta explicación aunque resulte desfavorable para nosotros, por lo menos da algún significado a los sucesos. Finalmente, los seres humanos, que se caracterizan entre otras cosas por buscar significados, se sentirían desasegurados con la idea de que los sucesos no significan nada.³⁷ De hecho, entre las causas del ocaso de la magia a finales del siglo XVII, destaca la creciente "capacidad para tolerar la ignorancia", capacidad que caracterizaría la actitud científica.

Ante la incertidumbre bajo la que se deben tomar muchas decisiones, tanto individuales como colectivas, las loterías se deberían usar con mayor frecuencia, obviamente sin pretender que expresen el destino o la voluntad de Dios. Pero parece que a la sociedad contemporánea le molesta el reconocimiento de la incertidumbre, la ignorancia y la indeterminación. La llamada teoría bayesiana, por ejemplo, puede verse como la expresión del deseo de tener razones para todo. Por ello, advierte Elster,

La idea de que ante la ausencia de información específica todos los resultados deben considerarse igualmente probables no puede justificarse en lógica debido al problema de los estados individuales del mundo. Tiene no obstante firmes fundamentos psicológicos en el deseo de forzar una solución determinada para todos los problemas decisorios; no resulta fácil ser capaz de tolerar la ignorancia y la ambigüedad.³⁸

³⁷ Los seres humanos no sólo buscamos significados, también razones, al grado de desear tener razones para todo lo que hacemos, llegando incluso en casos extremos a crearlas cuando no hay ninguna. Asimismo, tenemos la pretensión de que estas razones sean claras y concluyentes para tomar decisiones moderadas y no terminantes (en general no nos gusta tomar decisiones terminantes, tal vez por el potencial arrepentimiento asociado a ellas). En el caso de una decisión de este tipo, se da la tendencia a reajustar los argumentos que llevaron a ella, de manera tal que la opción elegida aparezca como claramente superior a las otras.

³⁸ *Ibid.*, p. 148. La necesidad de significado y razones estaría ligada a la renuencia a admitir la ignorancia y el agnosticismo. Curiosamente, las ciencias sociales en lugar de analizar esta necesidad más bien se han preocupado por satisfacerla. Elster llega incluso a sostener: "el

En general, como hemos señalado, la gente quiere tener razones para lo que hace y que las razones sean lo que determine sus elecciones. Sin embargo, cuando se propone recurrir a las loterías no se pretende excluir a la razón, ésta se utiliza pero en una etapa anterior en el proceso de decisión. Ya Pascal habían señalado cómo lo más racional puede ser decidir no utilizar la razón para tomar una decisión final:

El supremo paso de la razón radica en el reconocimiento de que hay una infinidad de cosas que la sobrepasan. Es muy débil si no alcanza a reconocerlo [...] Nada hay tan conforme a la razón como la denegación de la razón.³⁹

Recordemos que Ulises sigue el consejo de Circe de amarrarse al mástil, abdicando racionalmente de la razón, ya que no puede confiar en que tomará la decisión correcta cuando llegue el momento de oír a las sirenas. En el caso del recurso de las loterías para la toma de decisiones, que también implica una abdicación racional de la razón, no hay propiamente ninguna decisión correcta. A pesar de esta gran diferencia, ambas abdicaciones pueden estar relacionadas. Como destaca Elster, el recurrir a las loterías puede ser una forma racional de resistirse a las sirenas de la razón. Veamos en que consistiría lo que puede llamarse una adicción a la razón: si bien hay situaciones en donde no hay ninguna decisión óptima, pensemos por ahora en una situación en donde sí se podría alcanzar una decisión óptima, pero sus costos (en tiempo, dinero y esfuerzo) son más altos que sus beneficios; un adicto a la razón lejos de recurrir por ejemplo a una lotería para zanjar el problema, seguiría buscando razones hasta llegar a una que le permita

funcionalismo y el psicoanálisis inventaron significaciones donde no las había. La teoría de la decisión, particularmente la bayesiana, nos ha enseñado a darles razones a las decisiones que eran esencialmente injustificables desde el punto de vista racional." (Jon Elster, *Political Psychology*, p. 14).

³⁹ B. Pascal, *Pensieri*, Nos. 267 y 272, pp. 150-151.

justificar la decisión como una “decisión racional”. El ejemplo patético es el pleito por la custodia de los niños: con el fin de proteger el interés del niño, se buscan las pruebas de la idoneidad o no de los padres, aunque en el intervalo el daño causado al menor por esta búsqueda supera los beneficios esperados de obtenerla. En este caso, cuando ambos padres son “responsables”, sería más racional resistirse a las sirenas de la razón y dejar la decisión a la suerte.

Recordemos el célebre pasaje bíblico en donde al rey Salomón se le presentan dos mujeres y cada una afirma ser la madre de un niño que está vivo atribuyéndole a la otra la maternidad de otro niño que está muerto:

Dijo el rey: “Partid en dos al niño vivo y dad una mitad a una y otra a la otra.” La mujer de quien era el niño vivo habló al rey, porque sus entrañas se conmovieron por su hijo, y dijo: “Por favor, mi señor, que le den el niño vivo y que no le maten.” Pero la otra dijo: “No será ni para mí ni para ti: que lo partan.” Respondió el rey: “Entregad a aquélla el niño vivo y no le matéis; ella es la madre.”⁴⁰

El sabio Salomón supo descubrir cuál de las dos mujeres se preocupaba realmente por el niño.⁴¹ En los casos de divorcio en donde los dos padres pretenden quedarse con la custodia de los hijos, no siempre es fácil alcanzar una decisión basada en los intereses de los niños, no sólo por el alto costo emocional que el proceso puede implicar para ellos sino también por los casos en donde *honestamente* los dos tutores están en igualdad de condiciones.⁴²

⁴⁰ *Biblia de Jerusalén*, Libro Primero de Los Reyes 3:25-28, p. 369.

⁴¹ Aunque también podríamos preguntarnos por la situación en donde la supuesta impostora hubiese sido un poco más astuta, es decir, que las dos mujeres se hubieran opuesto al sacrificio del niño.

⁴² Las leyes sobre la tenencia infantil ha variado históricamente de país en país. En el siglo pasado destaca la *regla de preferencia paterna* en la legislación británica (sólo un padre abusivo podía perder su derecho legal a la tenencia de los niños) y la *presunción basada en la culpa* en los Estados Unidos (la parte “inocente” del divorcio cuidaría mejor de los hijos). En nuestro siglo poco a poco se volvió dominante la regla de preferencia materna (sobre todo en el caso de los niños pequeños, sólo una madre inepta perdería el derecho a su custodia). A partir de la década de los setenta, la preferencia materna se ha ido sustituyendo en muchos países por la decisión *según los intereses del niño*. (Cf. J. Elster, *Solomonic Judgements*, especialmente capítulo III “Solomonic judgements: against the best interests of the child”, pp.

Otro caso está vinculado al resquemor que puede producir la toma de decisiones terminantes, que nos parecen probablemente definitivas. Para reducir la tensión que provoca el temor y la previsión de arrepentimiento, se ajusta el peso de los criterios para hacer que una opción sea claramente superior a las otras. Incluso cuando racionalmente se utilizan loterías para la toma de decisiones, existe la posibilidad de que surja posteriormente el arrepentimiento si se llega a tener más información sobre la situación. Este problema puede llevar a buscar más información de lo que es preciso.

Por último mencionemos que la incapacidad para recordar la perspectiva *ex ante* puede volverse también un problema. A pesar de que una decisión resulte *ex post* errónea, pudo haber sido la mejor que se podía tomar en el momento de la elección. El ejemplo es el militar que para confundir al enemigo elige su plan de ataque tirando una moneda al aire, seguramente si falla su plan será fuertemente criticado. Por otra parte, también se debe tomar en cuenta que la igualdad *ex ante* puede ser muy frágil: es fácilmente aceptada mientras la moneda está en el aire, pero después la parte perdedora puede protestar porque sus razones no fueron suficientemente escuchadas.

La razón básica para usar sorteos en las decisiones es la honestidad. La honestidad nos exige que reconozcamos la presencia de la incertidumbre y la inconmensurabilidad, en vez de negarla o eludirla. Algunas decisiones serán arbitrarias y epistémicamente aleatorias al margen de lo que hagamos, por mucho que intentemos basarlas en razones. El azar regula gran parte de nuestras vidas, por mucho que intentemos eludirlo. La domesticación del azar nos permite controlar los factores aleatorios del universo en la medida de lo posible, y ahuyentar a la vez el autoengaño. Los requerimientos de causación personal y autonomía se concilian en el uso consciente del azar para tomar decisiones cuando falla la argumentación

123-174). En México está estipulado que en caso de disputa, el juez es el que decide cuál de los padres tendrá la custodia de los hijos considerando el interés superior de estos últimos (aunque parece que por usos y costumbres, salvo casos extremos, hay una preferencia a que la custodia se otorgue a la madre, en particular cuando los hijos son pequeños). Cf. *Código Civil para el Distrito Federal*, especialmente artículo 283, p. 27.

racional. Aunque esta desolada visión puede perturbarnos, es preferible a una vida construida sobre la confortante falsedad de que siempre podemos saber qué hacer.⁴³

Entre los ejemplos históricos del uso de loterías destaca sin duda la tragedia de Masada (74 d. c.). Como sabemos, los últimos sobrevivientes de la rebelión hebrea, sitiados en la célebre fortaleza construida por Herodes, rechazaron el ofrecimiento de los romanos de perdonarles la vida a cambio de su rendición. Frente a las previsibles cadenas, torturas y violaciones, los hombres de la comunidad, liderados por Eleazar ben Ya'ir, decidieron llevar a cabo un suicidio colectivo como su último acto libre. Cada hombre se encargó de su familia y de quemar las últimas provisiones; una vez cumplida esta tarea se reunieron para escoger por lotería a los diez que tenían que ejecutar a todos los demás; finalmente los últimos diez también echaron a la suerte quien se encargaría de los restantes nueve y de sí mismo. Como dice Flavius Josephus, tal era la confianza mutua que se tenían que para ellos no había diferencia entre uno y otro.⁴⁴

Entre los ejemplos contemporáneos de toma de decisiones difíciles, además de la ya mencionada custodia de los niños en caso de divorcio, destaca la asignación de órganos para transplantes. En ambos casos hay criterios básicos para la toma de decisiones, pero en situaciones extremas la racionalidad no puede establecer qué hacer o sólo podría hacerlo a un costo demasiado elevado.

En este capítulo vimos dos aspectos de la racionalidad que socavan las pretensiones iniciales de la TER: la racionalidad imperfecta y la racionalidad indeterminada. Con ello, sin embargo, no se pretende desechar la TER sino ubicarla en su justo espacio, estableciendo sus alcances y límites. Finalmente,

⁴³ J. Elster, *Solomonic Judgements*, p. 121.

⁴⁴ Cf. F. Josephus, *The Jewish War*, Libro VII, Capítulo IX, ed. gral. G. Cornfeld, p. 500. Obviamente las mujeres y los niños no participaron en la toma de decisión, aunque no todas se

la pretensión de racionalidad que llega a la adicción a la misma, en tanto no puede reconocer la falta o la sobra de racionalidad, se vuelve ella misma una forma de irracionalidad. En este sentido, la ponderación de las posibles aportaciones de la TER debe también incorporar sus límites para no volverse una expresión del mito de la razón.

resignaron a la determinación masculina; dos mujeres con cinco niños se escaparon y escondieron mientras se llevaba a cabo el suicidio de 960 personas.

IV. LA RACIONALIDAD DE LOS DESEOS

En el capítulo dos destacamos los elementos básicos con los que se caracterizan las acciones racionales. Las primeras cinco condiciones conforman lo que Elster denomina la *teoría estricta de la racionalidad*, la cual se presenta como insuficiente, en la medida en que, a excepción de un mínimo requisito lógico de consistencia, no comprende una caracterización de las creencias y los deseos que son las razones de la acción y que permiten atribuirle racionalidad. De esta manera, frente a la *teoría estricta*, en donde la racionalidad se identifica con consistencia (consistencia en el sistema de creencias y en el de deseos, así como entre estos dos últimos y la acción para la cual son razones), Elster nos propone una *teoría amplia de la racionalidad* en donde las creencias y los deseos sean racionales en un sentido más sustancial.¹

El capítulo está estructurado en dos secciones:

1. En la primera, recuperamos la sexta condición de la estructura de la explicación de elección racional, con la que se exige a las creencias una forma más amplia de racionalidad que la simple consistencia, y que remite finalmente a la noción de *juicio*. A semejanza con dicha condición de racionalidad para las creencias, nos ocuparemos de las posibles características de una racionalidad amplia de los deseos, la cual, a pesar de no estar incorporada a dicha estructura, remite finalmente a una noción básica como es la *autonomía*.
2. En la segunda sección, presentamos el cuestionamiento del modelo estándar de la TER en tanto los deseos, que son junto con las creencias

una de las razones de la acción, a su vez no descansan en ningún elemento que pueda racionalizarlos. Así como la racionalidad de las creencias responde a la evidencia disponible, Elster nos propone la incorporación de las emociones como un elemento que puede racionalizar a los deseos.

1. Racionalidad amplia

Hay que iniciar con una observación. Cuando se formula la pertinencia de una *teoría amplia* de la racionalidad no se pretende con ella abarcar todas las propiedades que tal vez nos gustaría que tuvieran las creencias y los deseos, como puede ser respectivamente la verdad y la bondad ética. Esto obedece a la siguiente consideración: si bien es pedir demasiado poco que la condición de racionalidad de las creencias y los deseos sea la consistencia, exigir como condición de racionalidad verdad en las creencias y bondad ética en los deseos, sería pedir demasiado. Precisamente, para Elster, “entre la teoría estricta de lo racional y la teoría completa de lo verdadero y de lo bueno hay suficiente espacio y necesidad para una teoría amplia de lo racional.”²

En esta teoría, además de la consistencia, se valora la racionalidad de las creencias y de los deseos, no por la verdad o bondad ética, sino por la *manera en que se conformaron* estos dos elementos que, como hemos visto, racionalizan la acción. De hecho, aunque una creencia sea verdadera o un deseo sea adecuado a la moral, tendríamos problemas para afirmar que son racionales si se formaron por factores causales irrelevantes o por una causalidad psíquica ciega. En este sentido, hay que enfatizar que en la teoría amplia, en donde se pretende incorporar una noción más sustancial de racionalidad, no se trata de ubicar contenidos “racionales” sino de enfocar la historia causal de las creencias y los deseos. De esta manera, los posibles

¹ Cf. Jon Elster, *Sour Grapes*, especialmente el capítulo I “Rationality”, pp. 1-42.

² Jon Elster, *Ibid.*, p. 15.

“errores” en estas historias los volverían irracionales más allá de la verdad o bondad de sus contenidos.

Aunque es difícil determinar cuándo una historia causal es “correcta”, al menos se pueden descartar algunos de los elementos que conducen a una causalidad “errónea”. En consecuencia, más allá de sus contenidos específicos y exigiendo algo más que la simple consistencia en los deseos y las creencias, tenemos que a partir del análisis de sus historias causales podemos distinguir algunos tipos de “distorsión” o “perversión” en los mismos, lo que llevaría finalmente a aceptar que si bien pueden fungir como “razones” de una acción no pueden sin embargo atribuirle racionalidad a la misma.

i) Creencias

En relación con las creencias, retomemos brevemente lo señalado en la sexta condición de la estructura de la explicación de elección racional: las creencias deben estar *fundadas correctamente* en la evidencia disponible, la cual presupone también, aunque sea en forma débil, una recolección *razonable* de la información pertinente.

Por una parte, como la racionalidad de una creencia descansa en su relación con las evidencias disponibles (no propiamente con el “mundo”), una creencia puede ser verdadera y no racional, y a la inversa, ser racional y no verdadera; pensemos por ejemplo en algunos casos históricos como puede ser la creencia visionaria de que algún día el hombre llegaría a la luna o la creencia de que con el método de la sangría se pueden aliviar ciertas enfermedades. Además de esta distinción entre la verdad y lo racional, como no basta cualquier relación entre las creencias y las evidencias, sino que estas últimas deben causar “correctamente” a las primeras, se deben excluir los casos en donde las creencias no son causadas por la evidencia sino por los

deseos³ así como cuando son el resultado de procesos de inferencia con errores compensatorios. Recordemos los ejemplos del enfermo que niega los síntomas de su enfermedad y se cree sano, y del hipocondríaco que por coincidencia se encuentra realmente enfermo.

Por otra parte, como vimos, la recolección de información también debe ser adecuada a la situación. Recordemos los casos del enfermo que no quiere hacerse los análisis correspondientes para poder seguir creyendo que está sano, y del hipocondríaco que continúa indefinidamente haciéndose estudios para no tener que aceptar que no está enfermo. Aquí podríamos también agregar el lamentable caso de los médicos que realizaron un trasplante de riñones con SIDA: frente al inminente deterioro de los órganos y su deseo de aprovecharlos, no esperaron los resultados de la prueba del SIDA y “confiaron” en los antecedentes del donador y de su esposa en los que no se reportaba la presencia del VIH.⁴

Finalmente, la sexta condición de la estructura de la explicación de elección racional, en la que se incorpora la *racionalidad amplia* de las creencias, remite en última instancia a la noción de *juicio*: la capacidad del agente para evaluar y sintetizar la información pertinente relacionada con una determinada situación de elección. Si bien éste ha sido uno de los grandes tópicos que ha acompañado a la filosofía, nosotros no nos detendremos especialmente en él.⁵ Para los fines del presente trabajo sólo queremos destacar que la racionalidad de las creencias, -en tanto presupone no sólo una

³ Cf. Salma Saab, “El problema del auto-engaño y el modelo de niveles de creencias” en *Límites de la subjetividad*, coord. Mariflor Aguilar, pp. 33-47 en donde se propone que la intervención de los deseos en la formación de creencias sí puede llegar a ser legítima. Para ello, la autora recurre a un modelo biológico-funcional que permitiría ubicar la intervención de los deseos en un nivel subintencional. Con esto, sin embargo, no se refuta, incluso refuerza, la idea de que “conscientemente” se debe excluir la intervención de los deseos en la formación de creencias. (Cf. también la observación que hace Elster para el caso de las emociones, *infra* cita de la nota 46)

⁴ Cf. *La Jornada*, 27 de febrero de 1999, p. 50.

⁵ Entre los trabajos que proponen una caracterización de las creencias podemos destacar el texto de Bernard Williams, “Decidirse a creer” publicado en *Problemas del yo*, pp. 181-200, así como el libro de Luis Villoro, *Creer, saber, conocer*.

historia causal “correcta” en relación a la evidencia disponible, sino también la decisión racional de concluir la recolección de información-, está ligada a cada caso y situación. En este sentido, se aleja de las grandes pretensiones de universalidad y univocidad, y recupera más bien la particularidad y la pluralidad. Por ello, cuando se incorporó la subcondición N), señalamos que era una condición de optimalidad débil: no se puede establecer un criterio general de optimalidad en la recolección de información, la cual obedece a los deseos y otras creencia del agente. De hecho, el gran supuesto es que cada situación tiene un *mínimo* y un *máximo* de información que es *razonable* tener antes de hacer la elección. Este “saber” si hay que seguir buscando información o dejar de hacerlo es un requisito que remite al *juicio* del agente.

ii) Deseos

Por lo que se refiere a los deseos, en la estructura de la explicación de elección racional no se llegó a registrar ninguna condición extra. Ahora nos ocuparemos de ello y veremos que, a semejanza con la noción de *juicio* para las creencias, en la *racionalidad amplia* de los deseos la noción clave es la *autonomía*.⁶ Aquí parece todavía más difícil que en el caso de la racionalidad amplia de las creencias, determinar cuándo un deseo es formado de *manera correcta* (es decir no “distorsionado” por procesos causales irrelevantes) así como distinguir la autonomía de los deseos de su bondad ética. Sin embargo, aunque sea de manera negativa, se pueden hacer ciertas diferenciaciones básicas. Para el primer aspecto, Elster se centra en el examen de algunos modos no autónomos de formación de deseos, y para el segundo se preocupa por destacar que cuando se critica un deseo por no ser autónomo se recurre a

⁶ Para este apartado cf. Jon Elster, *Sour Grapes*, especialmente capítulo III “Sour Grapes”, pp. 109-140, texto en el que el autor retoma un trabajo anterior, “Sour Grapes – Utilitarianism and the Genesis of Wants”, publicado en Amartya Sen y Bernard Williams (eds.), *Utilitarianism and Beyond*, pp. 219-238.

elementos distintos de cuando se critica un deseo por no ser ético.⁷ De hecho, un deseo puede ser cuestionado tanto por su origen (deseos no autónomos) como por su contenido (deseos no éticos).

Para la racionalidad amplia de los deseos, nos centraremos especialmente en el problema de la *autonomía*, en donde sobresale el mecanismo llamado *uvas verdes*, es decir, cuando las preferencias se adaptan a lo que se cree posible. Obviamente, este no es el único mecanismo, hay otras variedades como son el conformismo (adaptar los deseos a los de otras personas) y lo que se podría llamar la "inercia común". Asimismo, no hay que olvidar los fenómenos inversos, es decir, las preferencias contraadaptativas, el anticonformismo y la obsesión por la novedad. Diversidad de mecanismos en los que podríamos reconocer cómo "Nuestras mentes nos juegan toda suerte de trucos, y nosotros hacemos lo propio con ellas".⁸

Dada la dificultad de una definición positiva, partimos de una noción de juicio que está vinculada a la idea de que en la formación de las creencias no intervienen "distorsiones" o "ilusiones", así como de una noción de autonomía en la que la formación de deseos no cae en ciertos mecanismos que, como veremos, conforman preferencias irracionales. A pesar de que estas condiciones son poco satisfactorias y no pueden llevar a una caracterización completa, sí le permiten a Elster ubicar algunas de las principales condiciones que deben cubrir la formación de creencias y deseos para poder ser considerados racionales. De hecho, aunque su caracterización sea pobre, -al grado de no poder ser incorporada a la estructura de la explicación de elección racional-, con estos elementos se plantea la necesidad de empezar a delinear o esbozar una racionalidad amplia que no se limite a la simple consistencia de las creencias y deseos.

⁷ Obviamente con esto no se excluye la posibilidad de que un deseo sin "bondad ética" pueda convertirse también en un deseo no autónomo.

⁸ Jon Elster, *Sour Grapes*, p. 111.

Antes de pasar a ver los cuatro casos en los que se conformarían preferencias irracionales, hay que hacer dos aclaraciones. En primer lugar, los mecanismos *distorsionantes* que analizaremos a continuación pueden considerarse como dispositivos cognoscitivos o como dispositivos afectivos, es decir, sus estados finales pueden ser descritos tanto desde la perspectiva de las creencias como desde la de los deseos.⁹ Por otra parte, es necesario tener presente una distinción básica entre deseos y pulsiones. Si bien las pulsiones pueden modelar los deseos y también las creencias, en sí mismas no pueden ser consideradas deseos en virtud de que no son conscientes ni conocidas por la persona que las experimenta.¹⁰ Básicamente, se conciben a las pulsiones como fuerzas psíquicas no conscientes que se generan en función de la búsqueda de placer a corto plazo, y en este sentido, son opuestas a los deseos que no sólo son conscientes, sino que pueden privarse del placer a corto plazo para alcanzar un beneficio mayor a largo plazo (dar un paso atrás para poder dar dos hacia adelante).¹¹

Después de estas dos aclaraciones, podemos pasemos a ver los cuatro casos en donde, ya sea por pulsiones o defectos cognoscitivos, surgen distorsiones en la formación de deseos y creencias:

a) *Formación de preferencias adaptativas*. Los deseos se ajustan a las posibilidades pero no deliberadamente sino dentro de un proceso causal

⁹ De hecho, como veremos, la falta de racionalidad puede obedecer a procesos cognoscitivos equivocados o a influencias indebidas que proceden de algún impulso afectivo.

¹⁰ Por la misma razón, es decir, en tanto su influencia no es consciente ni conocida por la persona que la experimenta, las pulsiones tampoco podrían ser consideradas como una especie de metadeseos que modelan los deseos de primer orden.

¹¹ Recordemos dos de los principales pilares de la especulación psicoanalítica : 1) la conciencia no puede ser el carácter más universal de los procesos anímicos, sino sólo una función particular de ellos; y 2) las pulsiones son fuerzas que representan los requerimientos que hace el cuerpo a la vida anímica y pueden reducirse a dos básicas: la pulsión de amor (que busca producir unidades cada vez más grandes y conservarlas) y la pulsión de muerte (cuya meta es disolver nexos y destruir las cosas del mundo). (Cf. Sigmund Freud, *Obras Completas*, especialmente "Más allá del principio de placer", vol. XVIII, p. 24 y "Esquema del psicoanálisis", vol. XXIII, pp. 146-147)

inconsciente. Esta adaptación obedece a la pulsión de reducir la frustración que se siente al tener deseos que no pueden satisfacerse. El ejemplo por excelencia es la fábula de la zorra y las uvas que veremos detenidamente más adelante.

- b) *Cambio de preferencias por estructuración.* El atractivo de las opciones cambia cuando la situación en la que se da la elección es reestructurada sin que racionalmente se observen diferencias. Aquí no interviene ninguna pulsión sino rígidos procesos cognoscitivos. Podemos poner como ejemplo el siguiente caso: si al ir a comprar dos boletos para ir al cine perdemos un billete de 50 pesos, esto no impide (obviamente si se dispone de más recursos), continuar con el plan original, sin embargo, si ya hemos comprado los boletos con un costo de 50 pesos y los perdemos, no estamos dispuestos a volverlos a comprar.
- c) *Racionalización de la esperanza.* Este es el caso de las creencias que se conforman por los deseos del agente, lo que lleva a pensar que el mundo es realmente como se quiere que sea.¹² Al igual que en la formación de preferencias adaptativas, se trata de un proceso "caliente" pero aquí el resultado final es una creencia y no un deseo (aunque se trata de fenómenos que a veces se pueden sustituir). Por ejemplo, el deseo de un ascenso laboral puede generar en una persona la creencia de que la promoción es inminente.
- d) *Error inferencial.* Las creencias se basan en juicios e inferencias infundadas que proceden de defectos en el aparato cognoscitivo, se trata de una vía "fría" que finalmente también lleva a la formación de creencias

irracionales.¹³ Su mecanismo causal se asemeja al cambio de preferencias por estructuración, pero por lo que se refiere al efecto que produce se acerca más a la racionalización de la esperanza. En algunos casos la creencia no es racional ni siquiera en sentido estricto, aunque en la mayoría de los casos más que la consistencia lo que se ve afectado es el juicio. Por ejemplo, considerar a las probabilidades como si fueran aditivas: de un individuo que se juzga muy probable que sea de izquierda pero poco probable que sea economista, se juzga moderadamente probable que sea un economista de izquierda.

En virtud de que nuestro interés se centra en lo que Elster denomina una teoría amplia de la racionalidad, cuyas características no pudieron incorporarse a las condiciones de la estructura de la explicación de elección racional, aquí nos ocuparemos exclusivamente de las preferencias adaptativas. En paralelo al caso de Ulises, en donde la técnica de *atarse a sí mismo* exige una teoría imperfecta de la racionalidad, el mecanismo llamado *uvas verdes* que está tipificado en la fábula de la zorra y las uvas exige una teoría amplia de la racionalidad que incluya la formación de los deseos. Para empezar recordemos la conocida fábula en la versión de La Fontaine:

Cierta zorra gascona, otros dicen que normanda, de hambre casi muerta, colgando de una parra vio unas hermosas uvas, cubiertas de piel bermeja. ¡Gran banquete se hubiera dado la bribona! Pero no pudiendo llegar a ellas, dijo:

-¡Puah, están verdes! ¡Quédense para los gañanes!
¿Qué mejor podía hacer que desdeñarlas?¹⁴

¹² Como vimos, el llamado *wishful thinking* queda excluido con la subcondición 2b) de la estructura de la explicación de elección racional: la creencia debe ser *causada* por la evidencia disponible.

¹³ Como vimos, con la subcondición 3b) de la estructura de la explicación de elección racional (la evidencia debe causar de *manera correcta* la creencia) quedan excluidas incluso las creencias que resultan de procesos de inferencia con errores compensatorios.

¹⁴ Jean de La Fontaine, *et. al.*, *Las mejores fábulas*, p. 134. La versión original se atribuye a Esopo: "Quiso una zorra hambrienta, al ver colgando de una parra hermosos racimos de uvas,

En esta fábula, en donde la zorra, después de fracasar en su intento por obtener un racimo de uvas que está fuera de su alcance, dice que realmente no deseaba las uvas porque están verdes, se tipifica la actitud que puede tener un agente para disminuir su decepción.¹⁵ Este mecanismo de las *uvas verdes*, en donde se adaptan las preferencias a las posibilidades para reducir la disonancia, se caracteriza por ser un proceso puramente causal de adaptación que se da a *espaldas* del agente. De hecho, para la zorra las uvas están realmente verdes, y verdes no las desea.¹⁶

Aquí el problema no es tanto que las preferencias de la zorra cambien, de hecho las preferencias de un agente pueden cambiar no sólo con los resultados de sus acciones sino también ante las nuevas elecciones que debe tomar. Más bien el problema del mecanismo de las *uvas verdes* es que la zorra no es consciente del ajuste de sus preferencias a las posibilidades que cree tener, en otras palabras, el problema es que sus preferencias se conforman por impulsos.¹⁷ Por ello, se trata de un mecanismo que es irracional en virtud de

atraparlos con su boca; más no pudiendo alcanzarlos se alejó diciéndose a sí misma: -¡Están verdes! Asimismo ciertos hombres que no pueden llevar adelante sus asuntos por culpa de su incapacidad, culpan a las circunstancias." (*Ibid.*, p. 20)

¹⁵ Esta fábula es importante no sólo para la comprensión de la conducta individual sino también para la concepción de esquemas de justicia social. Con esta fábula se puede cuestionar uno de los pilares de la teoría utilitarista: la satisfacción de los individuos no es un criterio sólido para la justicia y elección social, ya que las preferencias individuales pueden ser conformadas por un proceso previo a dicha elección. En otras palabras, la elección social que se basa en las preferencias individuales puede ser cuestionada en virtud de que las personas tienden a adaptar sus aspiraciones a sus propias posibilidades. En este sentido, destaca Elster, para el utilitarismo no habría ninguna pérdida de bienestar si a la zorra se le excluyera del consumo de uvas ya que de todos maneras las cree verdes. Sin embargo, sería difícil justificar esa asignación invocando las preferencias de la zorra ya que la causa de su creencia de que las uvas están verdes es su convicción de que está excluida de su consumo. Como ejemplo de una polémica entre los que defienden una postura utilitarista y sus críticos, destacan por un lado John Harsanyi (cf. "Morality and the theory of rational behaviour" en *Utilitarianism and Beyond*, pp. 39-62) y por otro Amartya Sen (cf. el ya mencionado artículo "Los tontos racionales").

¹⁶ Uno de los principales problemas del mecanismo de las *uvas verdes* es que no sólo sirve para disminuir la decepción del agente ante la no realización de sus deseos, sino también puede impedir que éste consiga lo que desea cuando ello sí está verdaderamente a su alcance.

¹⁷ Como dijimos, las pulsiones son fuerzas psíquicas no conscientes que se generan en función de la búsqueda de placer a corto plazo. En este sentido son opuestas a los deseos conscientes

que las preferencias dejan de ser independientes de las circunstancias y el agente no interviene conscientemente en su modulación. En este sentido, la zorra pierde su autonomía y deja de ser un agente para convertirse en un simple receptor pasivo de su entorno.

Para esbozar mejor este mecanismo es pertinente distinguirlo de otros mecanismos, uno opuesto y otros que tienen causas similares o producen efectos semejantes.

1. *Las preferencias contradaptativas.* Este mecanismo es opuesto al de las *uvas verdes*. Si las *preferencias adaptativas* se caracterizan por “hacer de la necesidad una virtud”, las *preferencias contraadaptativas* por “dulce es el fruto prohibido”. Igual que en las *uvas verdes*, los deseos están determinados por el conjunto viable aunque aquí de una manera inversa: se desea lo que no se puede tener. Frente a la zorra resignada o conformista, aquí tenemos al insatisfecho permanente.
2. *Cambio de preferencias a través del aprendizaje.* Como dijimos, el problema con las *uvas verdes* no es el cambio de preferencias, sino que éste se dé de manera inconsciente para disminuir la frustración. Por ello, hay que distinguir el cambio de *preferencias adaptativas* del cambio de preferencias que surge a partir de la experiencia y la nueva información.
3. *Precompromiso.* En las preferencias adaptativas la alternativa preferida en el conjunto viable es también la opción preferida dentro del conjunto de alternativas concebibles. Esta característica puede producirse como resultado de un precompromiso, es decir, cuando se conforma

que pueden privarse de placer a corto plazo para alcanzar cierto beneficio o ganancia a largo plazo (tienen la capacidad de esperar y usar estrategias indirectas -un paso atrás y dos adelante). Como veremos más adelante, la conciencia es el criterio para distinguir el fenómeno de las *uvas verdes* de la *planificación del carácter*.

deliberadamente el conjunto viable con el propósito de excluir ciertas elecciones posibles. Como ya vimos, este es el caso de Ulises y, como veremos más adelante, también el de la planificación del carácter.¹⁸

4. *Manipulación*: El mecanismo de las *uvas verdes* lleva al agente a contentarse con lo poco que tiene, resignación de la que ciertamente se pueden beneficiar otros como un efecto colateral. Las preferencias adaptativas suponen una causalidad estrictamente endógena y, en este sentido, se oponen a las explicaciones exógenas en términos de los beneficios reportados a los otros. Por ejemplo, para un gobernante puede ser conveniente que los súbditos sean resignados, pero éstos son resignados porque es conveniente para ellos mismos.

5. *Planificación del carácter*. Como hemos visto, el mecanismo de las *uvas verdes* es un proceso causal de adaptación que se efectúa “a espaldas” del agente, por ello hay que distinguirlo de la conformación intencional de los deseos. En ambos casos existe una tensión inicial entre lo que el agente puede hacer y lo que podría gustarle hacer, pero a diferencia de las preferencias adaptativas en donde finalmente se degradan las opciones rechazadas (las uvas están verdes), en la planificación del carácter se recurre a *estrategias conscientes* en donde se acentúan los aspectos positivos de las alternativas escogidas.¹⁹

¹⁸ Un ejemplo interesante es el matrimonio: hay personas que se casan para impedirse (o al menos dificultarse) mutuamente el abandonarse, pero también hay personas que precisamente no se casan para estar seguras que el lazo que los une (o el deseo de no dejarse) no se vuelva una preferencia adaptativa. “No parece posible [señala Elster] asegurar al mismo tiempo que las personas permanezcan unidas por razones justas y adecuadas y que no se dejen el uno al otro por razones equivocadas.” (Jon Elster, *Sour Grapes*, pp. 114-115)

¹⁹ La distinción entre los mecanismos de las *uvas verdes* y la planificación del carácter recupera no sólo la diferencia entre preferencias modeladas por impulsos y preferencias modeladas por metapreferencias, sino también la diferencia entre adaptarse por resignación no consciente y adaptarse a lo inevitable por medio de la elección (como es el caso de los estoicos o el budismo zen).

6. *Cambio previo de pesos atributivos.* Las preferencias adaptativas implican un cambio en la valoración de los alternativas: una vez hecha la elección, la opción escogida adquiere una superioridad que antes no tenía en relación a las otras opciones. Esto hay que distinguirlo del cambio de valoración previo a la elección, en el cual se sopesan los pros y contras de acuerdo a una nueva información. Recordemos al asno de Buridán, sin duda no se habría muerto de hambre si hubiese podido tener alguna perspectiva desde la cual uno de los montones de heno fuese superior al otro.
7. *Adicción.* Las preferencias adictivas al igual que las adaptativas son inducidas por la situación de elección y cambian endógenamente. Pero la adicción se explica más por la naturaleza del objeto que por la tendencia a adaptarse a cualquier objeto que se presente. Asimismo, a diferencia de las *uvas verdes*, la adicción puede ir acompañada de un yo dividido que lucha en pro y en contra de ella, además de que el deseo del objeto de la adicción puede ser cada vez más fuerte.
8. *Preferencias dependientes de estado.* Las preferencias adaptativas, conformadas causalmente por la situación, se deben distinguir también de los cambios de preferencias conformados por elecciones presentes o pasadas. En otras palabras, a pesar de que en algunos casos puedan darse simultáneamente (incluso su diferencia puede depender más bien de cómo se describa la situación), hay que distinguir los cambios de preferencias que *dependen del estado* de la situación de elección y los que dependen de las *posibilidades* de la misma. Por ejemplo, las preferencias de lecturas pueden estar orientadas por el libro que se esté actualmente leyendo aunque también pueden obedecer a los libros a los que podemos tener acceso.

9. *Racionalización.* Frente a las *uvas verdes* en tanto mecanismo para reducir disonancia que actúa sobre las preferencias, hay un mecanismo alternativo que actúa sobre los elementos cognoscitivos modelando la percepción más que la valoración de la situación.²⁰ Por ejemplo, cuando a un empleado le niegan un ascenso, puede superar la disonancia diciendo que sus superiores temen sus capacidades (error de percepción) o que el puesto de todos modos no merecía la pena (error de formación de preferencias).

Cuestionando los alcances de la teoría estricta de la racionalidad, en esta sección hemos destacado que tanto la racionalidad de los deseos como la de las creencias pueden ser distorsionadas por las pulsiones. Por una parte, como vimos, la formación de preferencias adaptativas es el ajuste de los deseos a las posibilidades, no como adaptación deliberada (planificación de carácter) sino como resultado de un proceso causal que ocurre de manera inconsciente. Detrás de cada adaptación estaría la pulsión de reducir la tensión o la frustración de experimentar deseos que no pueden satisfacerse. Por otra parte, desde el lado de las creencias, la racionalización de la esperanza se caracteriza por la conformación de las creencias de acuerdo con los deseos, lo que lleva finalmente a pensar que el mundo es en realidad como queremos que sea.

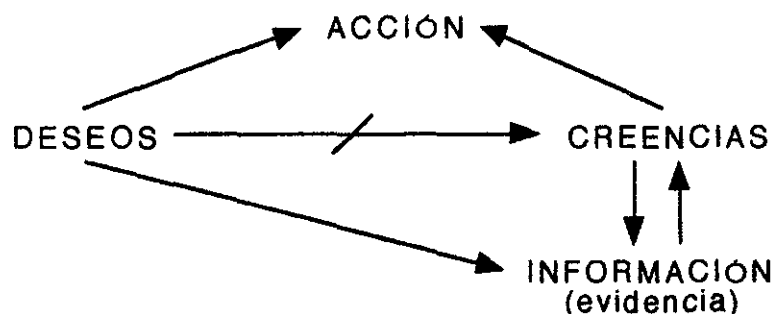
La ubicación de estos mecanismos cuestiona la pertinencia de atribuir racionalidad a una acción de acuerdo simplemente a la consistencia en el conjunto de deseos y creencias del agente, que conforman las razones de dicha acción. Como señalamos, la propuesta de una racionalidad más amplia remite para el caso de las creencias a la noción de juicio y para el de los deseos a la noción de autonomía. Estas nociones, que se presentan como la

²⁰ En algunos casos puede ser difícil distinguir las preferencias adaptativas y las percepciones adaptativas: la zorra puede ser engañada tanto por un error de percepción como por el cambio de preferencias. Asimismo, por lo que se refiere a las preferencias y percepciones

piedra de toque de la racionalidad amplia, se alejan de cualquier posible "encasillamiento" formal.²¹ Si bien sólo se ha esbozado de manera negativa su posible definición, se trata de términos ligados eminentemente a la diversidad y a la pluralidad. Si tuviéramos que asumirlos en algún concepto podríamos recurrir a la idea de que el agente racional, lejos de un simple egoísta consistente, implica más bien una capacidad crítica para no sólo recrear su mundo sino también a sí mismo, además de actuar en consecuencia.

2. La incorporación de las emociones

Además de la autonomía, hay al menos otro aspecto desde el cuál nos podemos preguntar por la racionalidad de los deseos. Recuperemos el modelo estándar de la acción racional que presentamos en el capítulo dos:



Recordemos que en este modelo las flechas pueden interpretarse no sólo como una relación de causalidad sino también como una relación de optimalidad. Por un lado, la acción es causada por los deseos y las creencias

contradaptativas, tendríamos que al "dulce es el fruto prohibido" le correspondería "la hierba siempre es más verde en el jardín del vecino".

²¹ Parecería que a Elster le interesa más bien destacar que no puede haber un fundamento último de racionalidad. Aunque llega a reconocer que también en el juicio se puede "colar la ambigüedad", enfatiza que es la autonomía la que está ligada especialmente a las "idiosincrasias del carácter" del agente. Finalmente, en relación al "carácter" del agente no tendría mucho sentido tratarle de atribuir racionalidad en la medida en que surge de elecciones pasadas, muchas de las cuales son arbitrarias o al menos basadas no en la optimización sino en la satisfacción.

del agente; las creencias por la información que se dispone; la cantidad de información recogida por los deseos y las creencias. Por otro lado, la acción es el mejor modo que tiene el agente para realizar sus deseos dadas sus creencias; las creencias son las mejores que el agente puede tener dada la información que dispone; y también la cantidad de información es óptima, dados los deseos del agente (por ejemplo la importancia que tiene para él la decisión) y aquello que cree en relación a los costos y las ventajas vinculadas a la recolección de información. Asimismo, no olvidemos que si bien los deseos pueden influenciar indirectamente a las creencias, una influencia directa no está permitida.²²

Ya hemos señalado algunos de los principales problemas que surgen cuando se aplica este modelo a la explicación del comportamiento humano. Ahora podemos agregar otra objeción: en este diagrama los deseos se comportan como un motor inmóvil. Podemos ver cómo de los deseos salen distintas flechas pero ninguna va hacia ellos. En este sentido, el modelo estándar de la acción racional no se pregunta por el mecanismo causal de los deseos, y algo todavía más fundamental, no se pregunta tampoco por su racionalidad. Al respecto destaca Elster,

Aquellos que estudian la acción y la decisión estarían naturalmente de acuerdo sobre el hecho de que los deseos deben tener un origen causal, también si generalmente definen esta causa en tal modo que está fuera de su disciplina. En cambio, normalmente, rechazarán como sin significado la idea de que las preferencias pueden ser más o menos racionales, por lo menos cuando se trata de ir más allá de la necesidad de satisfacer algún criterio formal de coherencia interna. En particular, no aceptarían nada que desempeñe para los deseos el papel que la información desempeña para las creencias, y que podría, por así decir, racionalizar los deseos.²³

²² Aunque habría que aclarar que tampoco están permitidas algunas influencias indirectas, como por ejemplo cuando el *wishful thinking* determina la manera en que recogemos la información.

²³ Jon Elster, *Piú tristí ma piú saggi? Razionalità ed emozioni*, p. 57.

De esta manera, además de la propuesta de una teoría amplia de la racionalidad, este autor también sugiere una modificación al modelo estándar de la acción racional para que se incorpore el problema de la racionalidad de los deseos: la inclusión de las emociones. De hecho, en diversos trabajos²⁴ el autor ha reflexionado en torno al papel que ocupan las emociones en el comportamiento humano y, en especial, a la relación que existe entre emociones y racionalidad. Con el fin de delinear algunas de las distinciones conceptuales básicas, retoma los estudios recientes sobre las emociones en el campo de la psicología así como algunas indicaciones sobre la relación de las emociones con las normas sociales que ha presentado en otros textos.²⁵

Para Elster, el estudio de las emociones es el *hijastro* de las ciencias sociales. La mayor parte de los trabajos serios vienen de la psicología y, aunque con varias objeciones (como son el paradigma experimental y la falta de información), las hipótesis que son *imperfectamente* verificadas pueden tener un gran interés intrínseco además de generar distinciones que pueden ser útiles independientemente de las hipótesis propias de la psicología. Por ejemplo, la distinción entre atribuciones globales y específicas permite comprender por qué las personas tienden a culpar a los otros de sus propias fallas, en este caso el problema es que normalmente se hacen atribuciones globales: el sujeto en lugar de reconocer que cometió simplemente una *mala acción* (y que lo llevaría a un sentimiento de culpa soportable) se ve a sí mismo como una *mala persona* y, como esta autoacusación le llevaría a un sentimiento de vergüenza intolerable, proyecta la responsabilidad en los otros.

²⁴ Para esta sección cf. especialmente, además del libro ya citado *Piú tristi ma piú saggi?*, el artículo que le antecedió "Sadder but wiser? Rationality and the emotions" así como el libro *Economics. Análisis de la interacción entre racionalidad, emoción, preferencias y normas sociales en la economía de la acción individual y sus desviaciones*.

²⁵ Cf. Jon Elster, *The Cement of Society. A Study of Social Order*, especialmente capítulo 3 "Social Norms", pp.97-151.

Sin embargo, a pesar de las indudables aportaciones de la psicología, para Elster la principal fuente para entrar a la vida emotiva del hombre son los moralistas clásicos y los grandes dramaturgos y novelistas que presentan a las emociones como la clave (que actúa a nuestras espaldas) para entender el comportamiento humano. Pero, también aquí, las relaciones entre las emociones y los comportamientos son débiles, y ello no tanto porque en los hombres no se presenten esas reacciones (por ejemplo, amar cada vez más a las personas a las que les hacemos un bien y odiar a las personas a las que les hemos hecho algún mal), sino porque frecuentemente se presentan también los comportamientos opuestos. Como ha señalado el autor en distintos contextos,²⁶ el instrumento metodológico principal para estudiar la causalidad psicológica es el uso de un *mecanismo* (parcial) y no de una *teoría* (general): a veces creemos que las uvas que no podemos alcanzar no están maduras, a veces creemos que el pasto del vecino es más verde, pero no podemos saber cuándo. Por ello, más que el recurso a grandes teorías, se necesita en cada caso reconstruir *a posteriori* la cadena causal correspondiente.

Aunque la literatura pueda ofrecernos más elementos que cualquier descripción factual en tanto muestra el funcionamiento interior de los mecanismos emotivos y no sólo sus apariencias exteriores, también la introspección tiene límites no sólo porque la excitación que acompaña a las emociones vuelve difícil una observación cuidadosa sino también porque las emociones que se quisieran observar son precisamente las que no pueden ser observadas. De hecho, parecería que hay ciertos límites naturales en el estudio de las emociones. Y advierte,

No le faltarán al lector las ocasiones para notar en lo que sigue el carácter fluido, incierto, ambiguo y vago de las intuiciones que tenemos sobre las emociones. Es posible que yo mismo haya

²⁶ Cf. Jon Elster, *Political Psychology*, especialmente "Introduction: Why Political Psychology?", pp. 1-34.

creado algunas de estas ambigüedades, pero la falta de consenso en este campo parece indicar que el problema es más general.²⁷

Junto con esta advertencia, nos propone un mapa conceptual²⁸ de las motivaciones humanas a partir de la siguiente tricotomía: las motivaciones pueden derivar de la racionalidad instrumental, de las pasiones o de las normas sociales.²⁹ Aquí por racionalidad instrumental se entiende restringidamente la racionalidad guiada por el interés personal, aunque se reconoce que el no-egoísmo puede desempeñar un papel importante en la explicación del comportamiento humano, para los casos que se analizan en relación a la incorporación de las emociones parece que sólo tiene un carácter marginal. Asimismo, se presupone que tanto las pasiones como las normas sociales pueden sobrepasar los intereses personales y que las emociones deberían poder triunfar no sólo sobre los intereses sino también sobre las normas sociales. Dentro de esta propuesta en donde se considera que las motivaciones humanas se pueden reducir a los intereses, las normas y las pasiones, aquí nos centraremos especialmente en el papel que pueden jugar las emociones. Pero antes hagamos algunas observaciones generales.

²⁷ Jon Elster, *Piú tristi ma piú saggi?*, p. 20.

²⁸ El propio Elster reconoce que "Como todos estos mapas, también el mío es un poco arbitrario" (*Ibidem.*). Incluso, remite al capítulo 4 de *Political Psychology* en donde, al analizar el pensamiento de A. Tocqueville, trata de demostrar que gran parte del mismo tiene un carácter arbitrario y especulativo, pero también advierte que tiene "cierto método en medio de su desvarío. Al analizar las instituciones sociales y el cambio social, Tocqueville estableció una cantidad relativamente pequeña de mecanismos psicológicos bien definidos. Estos, más que sus hiperbólicas y a menudo inconsistentes pretensiones explicativas, constituyen la razón de que valga la pena leerlo y releerlo hoy." (Jon Elster, *Political Psychology*, p. 191)

²⁹ Otra tricotomía que también ha propuesto el autor es la siguiente: las motivaciones pueden derivar de la razón, del interés o de la pasión. Aquí las normas sociales se incluyen en la categoría de las pasiones (que necesitan de su asistencia) y la categoría de la razón incluye las motivaciones imparciales (deontológicas y consecuencialistas). Esta tricotomía es especialmente útil para el estudio del comportamiento político. Por ejemplo, en la elaboración de una constitución, aunque los intereses y las pasiones traten de presentarse como si fueran la razón, se pueden identificar las tres motivaciones claramente. Como es en el caso de la Convención Federal de 1776 y la Asamblea Constituyente Francesa de 1789-91. (Cf. Jon Elster, *Argomentare e negoziare*)

En primer lugar, estos tres tipos de motivaciones no regulan diferentes tipos de conductas, de hecho cualquier conducta puede verse afectada simultáneamente por las tres motivaciones.³⁰ En nuestra interacción con otras personas partimos generalmente de que desean ser racionales aunque en la práctica esto frecuentemente no se realice (e incluso haya casos en donde se promueva). Para Elster, las dos fuente principales de irracionalidad son las emociones y las normas sociales ya que pueden hacer que las personas actúen en contra de sus propios intereses.³¹

A diferencia de la conducta racional que es útil y está orientada hacia el futuro, las normas sociales se caracterizan por ser imperativos que no están centrados en los resultados (lo que no excluye que a veces la falta de previsión puede traer aparejadas consecuencias positivas). Las condiciones para que las normas sean sociales son básicamente que sean compartidas con otros miembros de la comunidad (ya sea en su conjunto o un pequeño grupo) y que se hagan cumplir en parte por las sanciones aplicadas a otras personas (que pueden ir desde un pequeño gesto de desaprobación hasta serias formas de ostracismo social).³² Estas sanciones se caracterizan por las emociones que generan en el observador de la violación de la norma, como puede ser la ira, la indignación o el desprecio; así como por la vergüenza que siente el que es blanco de esas emociones. De hecho, el principal vehículo que hace que se

³⁰ Justo la articulación de la interacción de estas tres motivaciones es la tarea de lo que Elster llama la *egonomics* cuya característica distintiva es la utilización de las herramientas analíticas derivadas de la teoría económica. Pero a diferencia del agente económico que es racional, unitario, con intereses propios y supuestas preferencias inmutables, "La *egonomic*, en cambio, explora la posibilidad de que los agentes sean irracionales, no egoístas ni unitarios y que sus preferencias estén sujetas a formas sistemáticas de cambio." (Jon Elster, *Egonomics*, p. 37)

³¹ Las emociones pueden llevar a actuar no sólo en contra de los propios intereses sino también en contra de las normas sociales. Como veremos, la fuerza de las emociones es tal que es la que permite que se impongan las normas y convenciones sociales.

³² Las normas sociales se distinguen de otros tipos de normas como son las morales, legales y privadas, así como también de los hábitos y neurosis compulsiva. Algunos ejemplos de normas sociales serían las normas de consumo, las que regulan la conducta "contraria a la naturaleza", el uso del dinero y la reciprocidad.

cumplan las normas sociales es la percepción del sentimiento de vergüenza que acarrearía su violación (incluso si no hay observadores).

Sin embargo, hay que aclarar, la relación de las emociones con las normas sociales no se reduce a su sostén y respaldo; las normas sociales regulan grandes áreas de nuestra vida, y existen también normas que regulan qué emoción es o no apropiado expresar en determinadas situaciones, por ejemplo, alegría en un nacimiento y pena en un funeral. Asimismo, las emociones también pueden ser espontáneas, es decir, no depender de las normas sociales. De hecho, las emociones son capaces de sobrepasar tanto el interés como las normas sociales, y como estas últimas, también a veces pueden servir a los intereses del agente. Entre las principales características de las emociones se encuentran las siguientes:

1. Excitación fisiológica o psicológica, ya sea intensa o suave (pero no cero), que habitualmente se manifiesta externamente, por ejemplo en la postura corporal, el tono de voz o la expresión facial;³³
2. Habitualmente se les asocia con determinadas tendencias a la acción: golpear la causa del enojo, destruir por envidia, huir por miedo, esconderse por vergüenza, danzar de alegría, etc.;
3. Asimismo, una emoción puede ser negativa o positiva dependiendo de si es buena o mala para la persona que la está sintiendo en ese momento (no por sus consecuencias), y

³³ Una de las perspectivas desde la que se ha analizado no sólo el problema de las emociones sino también el de las reglas morales y sociales, es el de la neurociencia. Algunos más cautelosos que otros, plantean los vínculos de dichos aspectos con el cerebro. En especial, se ha detectado que la corteza prefrontal afecta el comportamiento social, tanto en el aprendizaje de reglas como en el uso de las emociones para la toma de decisiones. Sin embargo, si bien no pueden negarse sus posibles vínculos, hay que tener cuidado con los "reduccionismos materialistas". Entre los que han expuesto este planteamiento se encuentran Eduardo Rabossi, ponencia presentada en el XIV Congreso Interamericano de Filosofía (Puebla, 17 agosto de 1999); un caso extremo es Paul. M. Churchalan en su libro *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*, y en otro tono, desde una perspectiva evolucionista de la racionalidad, también el ya mencionado trabajo de Robert Nozick, *La naturaleza de la racionalidad*.

4. A diferencia de los sentimientos (generales) y los estados de ánimo (disposiciones), las emociones tienen un objeto, es decir, *se refieren a algo* en particular, se sienten *por* algo.

En general, las consecuencias de las emociones sobre la conducta dependen de su carácter específico y de su objeto intencional. Como veremos, un aspecto fundamental es que las emociones tienen la capacidad de alterar y distorsionar la apreciación del objeto que las provocó, por lo que, destaca Elster,

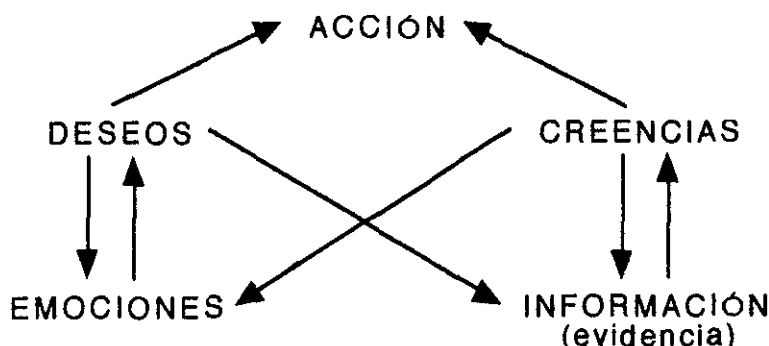
*El objeto de una emoción es el retrato emocional distorsionado de su causa. Esta retroalimentación desde las emociones a sus orígenes cognitivos es un elemento clave para la dinámica de las emociones y explica cómo pueden crecer y descontrolarse.*³⁴

El problema básico que trata de enfrentar Elster es que frecuentemente se considera que la racionalidad y las emociones son diametralmente opuestas. Se supone que las emociones interfieren en nuestra capacidad de formar creencias racionales o de hacer elecciones racionales. Sin embargo, para nuestro autor, esta descripción es inadecuada: además de que desempeñan un papel crucial para darle sentido y rumbo a la vida, las emociones pueden estar sujetas a criterios de racionalidad así como también pueden facilitar la cognición en lugar de obstruirla.

En relación al modelo estándar de la explicación de la acción racional, como ya dijimos, Elster nos propone incorporar a las emociones. A pesar de que no conforman una unidad homogénea y que se debe en cada caso recuperar sus particularidades, con esta incorporación de las emociones, los deseos adquieren una causalidad y se puede representar el problema de su

³⁴ Jon Elster, *Economics*, p. 119.

racionalidad. A partir de esta propuesta tendríamos el siguiente modelo de la explicación de la acción racional.



Veamos las flechas que van a las emociones y parten de ellas. En primer lugar se está marcando que las emociones dependen de las creencias por la necesidad de dar una valoración cognitiva a la situación que las suscita, antes incluso de que pueda surgir cualquier reacción emotiva.³⁵ Pensemos, por ejemplo, en el caso de los celos como respuesta a la creencia de que el objeto amado se interesa por otro.

Que las emociones dependen de la actividad cognitiva es verdad sin excepción, ya que la evidencia de los sentidos debe ser interpretada como dotada de significado antes de que sea posible reaccionar. La relación de dependencia opuesta es contingente y depende de la habilidad de las personas para aislar los propios procesos del pensamiento de los emotivos.³⁶

Además del necesario vínculo con las creencias, muchas emociones, aunque no todas, surgen como resultado directo de la satisfacción o frustración de ciertos deseos. Y a la inversa, una emoción puede suscitar el deseo de cambiar o de mantener la situación que la ha producido. Incluso, en algunos casos, se pueden observar ambos mecanismos, uno después del otro: la

³⁵ Desde otra perspectiva, también en la filosofía analítica se ha trabajado la diversidad de las emociones. Para el estudio del papel que desempeñan las actitudes proposicionales en algunas emociones, destaca el trabajo de Olbeth Hansberg, *La diversidad de las emociones*.

³⁶ Jon Elster, *Piú tristí ma piú saggi?*, p. 58.

satisfacción o frustración de un deseo causa una emoción que a su vez causa la aparición de un nuevo deseo. Para seguir con el ejemplo de los celos, pensemos en el caso extremo que desemboca en el asesinato del objeto inicialmente amado.

En este modelo modificado, las nuevas relaciones tienen la misma interpretación que en el modelo estándar, es decir, las relaciones de las emociones con las creencias y los deseos tienen una interpretación causal. Pero surge la pregunta: ¿pueden ser interpretadas también como condiciones de optimalidad? Para responder esta cuestión, Elster divide el problema en dos partes. Primero, ¿qué significado puede tener la idea de que una emoción es racional dadas ciertas creencias y ciertos deseos? (aunque como en el caso de la autonomía, se ocupa más de las emociones irracionales que de las racionales). Segundo, ¿qué significado puede tener la idea de que un deseo es racional dadas ciertas emociones? Obviamente, no se pretende dar aquí una respuesta acabada a estos problemas, sólo esbozaremos el posible horizonte de solución.³⁷ Asimismo, antes de pasar a delinear las posibles respuestas a estas interrogantes, hay que aclarar que nos limitaremos a las emociones vinculadas con la frustración y la satisfacción de los deseos para mantener la discusión ligada al modelo estándar de la TER, el cual, como hemos visto, descansa su explicación de la acción en los deseos y las creencias del agente.

i) Emociones racionales

La primera pregunta sobre la racionalidad de una emoción dadas ciertas creencias y deseos, se puede plantear de una manera más general: ¿qué significado puede tener la idea de una emoción racional? Elster considera seis

³⁷ El propio Elster advierte que no pretende dar una visión final del problema, y como en otros trabajos, señala que está consciente de que el tratamiento está en un estado inicial y que las caracterizaciones presentadas son sólo indicaciones para futuros desarrollos.

posibles respuestas, aunque obviamente ninguna puede ser considerada en sí misma como *la respuesta*:

- a) las emociones racionales son simplemente un cierto tipo de acción racional
- b) las emociones son racionales si son del tipo más adecuado a la situación
- c) las emociones son racionales si tienen la intensidad más adecuada a la situación
- d) las emociones son racionales si se fundan en creencias racionales
- e) las emociones son racionales si se fundan en deseos racionales (autónomos) y
- f) las emociones son racionales si nos hacen felices

Pasemos a ver algunas de las principales observaciones que hace el autor en relación a cada una de estas posibles respuestas:

a) En el modelo estándar, tanto la acción como la información se consideran racionales según los deseos y las creencias del agente, por ello la información puede considerarse también como la abreviatura de una acción: recoger información. Como vimos en el segundo capítulo, cada acción está acompañada de una "acción sombra" anterior: el proceso de recolección de información. De ahí surge la pregunta por si las emociones, en el caso de ser evaluadas como racionales a la luz de las creencias y de los deseos, puedan igualmente ser conceptualizadas como acciones. Por ejemplo, para Robert Solomon las emociones serían muy similares a las acciones: las emociones están dirigidas a alcanzar algún fin y se escoge una emoción de la misma manera que se escoge una línea de conducta. Pongamos el siguiente ejemplo de cólera, que no se explicaría por a quién está dirigida o cuál es su causa, sino por su finalidad: un profesor, al que sus estudiantes le reprochan sus frecuentes inasistencias, monta en cólera porque los alumnos no hicieron las lecturas de la sesión; en este caso la cólera le permitió al profesor manipular y detener las quejas de sus estudiantes. Sin embargo, esta propuesta no es muy

convinciente ya que se presupone algún tipo de saber sobre el efecto probable de la emoción, saber que al mismo tiempo no podría ser consciente ya que induciría a su vez a fingir la emoción.³⁸

b) Las emociones podrían asemejarse más a las creencias, las cuales, como vimos, aunque no son escogidas por un acto de voluntad, se definen como racionales si tienen una determinada relación con la información disponible. Una emoción podría ser más o menos apropiada o adecuada a la situación que la causa. Sin embargo, esto no puede significar que la emoción corresponde a las expectativas ajenas o que suscita su aprobación, ya que entonces estaríamos frente a un problema que remite más bien a las normas sociales que a la racionalidad. Tampoco puede significar que hay una conexión necesaria entre la emoción y las condiciones que la generan, porque entonces las emociones nunca podrían ser irracionales o inapropiadas. De esta manera, se debe buscar una noción intrínseca de *apropiada* y, sobre todo, de *no-apropiada*. En consecuencia, Elster nos propone que una emoción sería irracional o inapropiada si hay una divergencia (generada por otra emoción) entre el objeto intencional de la emoción y su causa: por ejemplo, el hombre que entra en cólera con la familia porque fue amonestado en el trabajo, o la persona en la que la envidia por un bien ajeno se transforma en indignación ya que se termina creyendo que seguramente fue obtenido de manera inmoral.³⁹

³⁸ Si bien la conexión entre una emoción y su efecto podría obedecer a un mecanismo de refuerzo, Elster advierte: “[...] podemos en cierta medida escoger la intensidad de nuestras emociones, controlando las expresiones exteriores que frecuentemente tienden a reforzarlas. Una vez que ha surgido una emoción, podemos, por así decirlo, darle nuestro asentimiento o negárselo. Pero todos estos fenómenos son de segundo orden. No constituyen una argumentación a favor de la tesis según la cual las emociones son acciones.” (Jon Elster, *Piú tristi ma piú saggi?*, p. 63)

³⁹ En el primer caso, la emoción se asocia con un objeto que *no es la causa* de la misma, en el segundo, la emoción actúa sobre la *percepción* del objeto que sí la generó. Hay que agregar que habría otras maneras en que no son apropiadas las emociones, pero serían ejemplos en los que más que hablar de irracionalidad tendríamos reacciones *patológicas*: sentirse culpable por no haberle dado un “aventón” a alguien que después nos enteramos que en el trayecto sufrió un accidente, o sentir alegría ante el dolor inmerecido de otra persona.

c) Las emociones podrían ser inapropiadas o irracionales si son desproporcionadas en relación a sus objetos intencionales. A diferencia del caso anterior, aquí el problema radica en el grado de intensidad.⁴⁰ Un ejemplo sería el siguiente: si bien es apropiado sentir cierto miedo ante la posibilidad de que un coche a exceso de velocidad nos atropelle aunque seamos cuidadosos al cruzar la calle (sobre todo en la Ciudad de México), este miedo se volvería inapropiado si es de tal intensidad que nos impidiera cruzar la calle. Sin embargo, aquí se puede decir que la emoción es irracional porque se hace referencia a ciertos criterios de coherencia que las emociones deben satisfacer: por ejemplo, si nuestro miedo es mayor que en otras ocasiones en donde creíamos correr el mismo riesgo, o si nos convencemos de que el riesgo es menor de lo que pensábamos y nuestro miedo aumenta. Por ello, afirma Elster, “no alcanzo a ver otras razones por las que se pueda afirmar que las emociones son desproporcionadas en relación a su objeto. No existe un nivel (o una gama de niveles) intrínsecamente ‘justos’ de cólera, alegría, envidia o dolor que puedan ser generados por una situación dada.”⁴¹

d) Las emociones podrían ser irracionales si se basan en creencias irracionales sobre la situación que las generan.⁴² Pensemos en el clásico caso de los celos infundados. Sin embargo, aquí hay que destacar que no todas las creencias irracionales tienen causas emotivas (como hemos vistos, puede ser el caso de una irracionalidad “fría” debida a procesos cognitivos deficientes), así como tampoco todas las creencias irracionales que tengan un origen

⁴⁰ En este inciso se deben excluir las fobias (que más que irracionales pertenecerían al ámbito de lo patológico), las emociones inadecuadas a su causa (que corresponde al inciso anterior), la desproporción en virtud de las características *reales* del objeto (por ejemplo el miedo ligado a un objeto que erróneamente creemos que es peligroso), así como el papel que puedan tener las normas sociales en la definición de cuál es la intensidad apropiada de un determinada emoción.

⁴¹ Jon Elster, *Piú tristí ma piú saggi?*, p. 68.

emotivo son casos de irracionalidad motivada: las emociones pueden ser sólo factores del desencadenamiento independientemente del sentido o dirección de las creencias, pensemos en el caso de un vendedor enamorado que por estar pensando en la mujer amada se equivoca al hacer las cuentas, sin que ello sea necesariamente en su propio beneficio. Por otra parte, la irracionalidad motivada puede surgir de deseos que a su vez no tienen ningún fundamento emotivo: se puede partir de un deseo tranquilo (grado cero de excitación) de ir a una obra de teatro, este deseo puede inducir la creencia (contra la evidencia disponible) que todavía se pueden conseguir los boletos, finalmente, creencia y deseo pueden llevar a inducir la emoción de la esperanza, que refuerza a su vez la creencia y la propia esperanza de conseguir las entradas.

e) Las emociones no sólo se basan en las creencias sino también en los deseos. Por ello, de la misma manera en que una emoción puede ser alterada por la irracionalidad de las creencias, nos podemos preguntar por si las emociones pueden ser alteradas también por deseos irracionales. No hay un consenso sobre qué puede entenderse por deseo irracional, incluso como señalamos, habría quien niegue que esta idea tenga algún sentido. Pero Elster propone dos hipótesis: en primer lugar, como vimos, un deseo podría ser definido como irracional si es heterónimo (es producido por un mecanismo psíquico no autónomo como es el caso del conformismo y su contrario el anticonformismo); en segundo lugar, como veremos en el siguiente inciso, los deseos irracionales podrían ser aquellos que conducen a desilusiones o sufrimientos inútiles.⁴³

⁴² Esta sería la idea general del caso particular que ya vimos en el inciso b): las emociones son irracionales si su objeto intencional está conformado o distorsionado por una emoción.

⁴³ A veces estas hipótesis conducen en la misma dirección, como es el caso de cuando se adaptan los deseos para desear lo que no se tiene, pero también pueden conducir a conclusiones opuestas, como es el caso de la adaptación de los deseos a lo que es posible en tanto mecanismo no-autónomo para producir la mayor felicidad posible.

f) La última hipótesis es que las emociones racionales serían las que producen felicidad y las irracionales las que causan inútiles sufrimientos. Esta hipótesis, difiere de las anteriores porque se aplica a las predisposiciones emotivas y no a las emociones en acto. Aquí, la idea de racionalidad es una guía de acción y no sólo un criterio de valoración. En comparación con la capacidad de controlar las emociones en acto, nuestra capacidad de cambiar las predisposiciones emotivas es mayor. Sin embargo, enfatiza Elster, el hecho de poder cambiar no implica que sea racional hacerlo: primero podría ser más útil evitar las situaciones en las que surgen las emociones que se quieren evitar; los costos de la planificación del carácter pueden ser más elevados que la desdicha de esas emociones; asimismo, puede haber un acoplamiento de emociones por lo que el cambio de una que lleva a la desdicha también impide el gozo de otra.⁴⁴ Además, suponiendo que sí es racional, es decir, que el sufrimiento neto puede reducirse, las consideraciones sobre la eliminación de las emociones que se vinculan a sufrimientos innecesarios, debe contemplar su diversidad y especificidad: no es lo mismo tratar de eliminar la culpa que la vergüenza, el miedo o la envidia, así como tampoco toda culpa, vergüenza, miedo o envidia son iguales.⁴⁵ Finalmente, si se define la racionalidad como felicidad y bienestar, no sólo tenemos el problema de la no coincidencia entre el punto de vista colectivo y el individual, sino también el hecho de que los hombres no son simplemente máquinas para producir felicidad: sin duda hay una gran diferencia entre la infelicidad de Sócrates y la alegría de un cerdo.

⁴⁴ En este último caso, la racionalidad de cambiar las emociones dependería del resultado global del *flujo* de experiencias, es decir, si es positivo o negativo. Obviamente poder calcular esto sería muy complicado ya que, dentro de ciertos límites, no sólo se pueden manipular las inclinaciones emocionales sino también el tipo de experiencias al cual uno se expone.

⁴⁵ Como vimos, con la culpa se puede tratar de reparar el daño, con la vergüenza en cambio el sujeto busca esconderse e incluso, cuando el sufrimiento es insoportable, puede llegar a culpar a otro de los resultados negativos de sus propios actos. Por otra parte, por ejemplo la eliminación de la vergüenza puede llevar a rechazar ciertas prácticas sociales que llevan al sufrimiento, pero también puede llevar a un concepto más débil de responsabilidad individual.

De esta manera, en relación al significado que puede tener la idea de que una emoción es racional, tenemos primeramente las nociones derivadas, es decir, las emociones son racionales si tienen una base racional. Sin embargo, Elster pone especial énfasis en las nociones no derivadas, entre las cuales destaca tres ideas básicas: la racionalidad como factor de felicidad, como coherencia y como apropiación. Como hemos visto, la primera si bien puede ser útil es incompleta (nuestra vida no se limita a ser felices), la segunda recoge una parte (aunque no muy significativa) de nuestras intuiciones, y la tercera sería la más interesante ya que plantea cómo *las emociones pueden interferir entre sí*, volviéndose inoportunas con respecto a la situación que las originó.

ii) Deseos racionales

Pero el posible significado de la idea de que las emociones son racionales, es sólo la primera parte del problema que nos ocupa. Pasemos ahora, a la segunda parte: ¿cuál puede ser el posible significado de la idea de que un deseo es racional dadas ciertas emociones?⁴⁶ Para una noción cognitiva (ascética) de las emociones, una creencia puede hacer que un deseo sea racional, sin que las emociones intervengan en acto, incluso aunque haya una excitación no es necesario involucrarla para racionalizar el deseo:

⁴⁶ Elster no es el único que se ha preguntado por la racionalidad de los deseos más allá de la simple consistencia. Entre los autores que se han hecho esta pregunta destaca el planteamiento de Thomas Nagel: la motivación por razones no siempre depende de deseos previos, a veces un deseo puede nacer del reconocimiento de que hay razón para hacer o querer algo (por ejemplo, el altruismo procedería del reconocimiento del deseo e intereses de los otros y no requeriría que el agente tenga ningún deseo, salvo el que motiva el reconocimiento mismo. (cf. *The Possibility of Altruism*). Esta propuesta de que un deseo puede ser motivado por razones que justifican el fin de la acción, es retomado por Luis Villoro quien enfatiza que no por ello se pueden relegar los deseos: las razones son condiciones necesarias pero no suficientes de la acción, se necesita la presencia de un deseo que motiva la decisión de actuar. Por otra parte, además de los deseos posteriores a la razón, debe haber un deseo inicial para encontrar razones que guiarán nuestro comportamiento. (Cf. *El poder y el valor. Fundamentos de una ética política*).

pensemos en el caso del miedo a que nos roben el automóvil si no le ponemos ningún sistema de alarma o seguridad. La creencia en que puede ocurrir el robo y el deseo de que no ocurra, vuelve racional el deseo de hacer algo para impedirlo, es decir, instalar algún sistema de alarma o seguridad en el automóvil. En este caso, los deseos no tienen su origen en alguna emoción, sino más bien se trata de una cadena de deseos relacionados entre sí para alcanzar un fin: la instalación del sistema de alarma o seguridad en el automóvil es un medio para que no me lo roben.

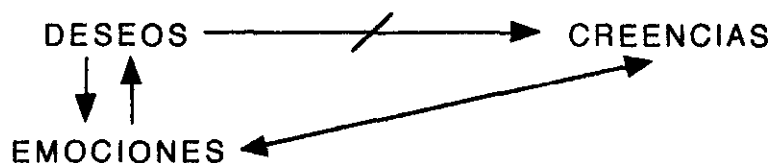
Por otra parte, una emoción puede producir un comportamiento que es instrumentalmente irracional en relación al deseo, el cual finalmente se ve frustrado. Por ejemplo, los celos pueden llevar a alguien a tratar de destruir al adversario, pero justo con ello se aleja aún más del ser amado. Aunque obviamente, la tendencia a una acción que acompaña a una emoción no es necesariamente irracional como instrumento para realizar el deseo cuya frustración o satisfacción estaba en su origen: por ejemplo, alguien que se siente herido tenderá naturalmente a atacar al objeto que le produjo el daño. Sin embargo, lo que es natural tampoco es necesariamente racional: el miedo puede llevar a responder rápidamente en un caso de emergencia, pero también puede llevar a la parálisis. Por ello, frente a los que consideran que las *tendencias naturales* deben servir a algún propósito, sostiene Elster,

Soy muy escéptico en relación a los análisis funcionales de las emociones, ya sea que se basen en sus consecuencias socialmente útiles, o en su efecto de fomentar la adaptación [...] es necesario estar atentos a no confundir la noción objetiva de adaptación con aquella subjetiva de racionalidad.⁴⁷

Con estas observaciones, podemos ahora retomar el segundo esquema de la TER e interpretar la flecha bidireccional que relaciona los deseos con las emociones. Por un lado, las emociones pueden ser racionalizadas por los

deseos en el sentido de que pueden ser reacciones apropiadas a la frustración o satisfacción de un deseo. Por otra parte, un deseo puede ser racionalizado por una emoción si ésta lleva a mantener las condiciones que satisfagan los deseos que la han hecho surgir o también si lleva a eliminar las condiciones que lo impiden.⁴⁸

Sin embargo, queda todavía un hilo suelto. Comúnmente se supone que la relación entre emociones y actividad cognitiva es análoga a la relación que hay entre ésta y los deseos. Como ya hemos señalado, el mejor modo para favorecer la realización de los deseos es no involucrarlos en el proceso de formación de una creencia, como se indica con la flecha cruzada en la siguiente figura:



Elster se pregunta si también hay que cruzar la flecha que va de las emociones a las creencias. Si bien las emociones, al igual que los deseos, pueden deformar nuestra valoración cognitiva de las situaciones, en cuanto a sus particularidades difieren en algunos puntos importantes. Cuando un deseo tiene un efecto causal sobre la formación de una creencia, normalmente produce la creencia de que la realidad es como queremos que sea. En cambio, cuando una emoción tiene un efecto causal sobre la formación de una creencia, produce la creencia de que las cosas están de tal modo que se justifica la emoción. En este sentido, al igual que en el caso de la acción, también para las emociones se puede llegar a descubrir que

⁴⁷ Jon Elster, *Piú tristi ma piú saggi?*, p. 79.

la razón por la que la incertidumbre representa para nosotros un tormento es que está acompañada de la creencia que nuestra emoción puede estar privada de justificación.⁴⁸

A pesar de esta similitud con los deseos, las emociones pueden también desempeñar un papel positivo en la formación de una creencia. Podemos usar nuestras reacciones emotivas como indicios de una "evaluación inconsciente" de una determinada situación: muchas valoraciones cognitivas que están en la base de una emoción son ellas mismas plasmadas por una emoción. De ahí la pregunta por considerar a la presencia de una emoción como una indicación atendible de la presencia de una creencia atendible, aunque aquí también no estamos en posibilidad de decidir cuáles sean estos casos.

Pero la relación de las emociones y la formación de creencias es peculiar. Por el lado de las emociones, si se toma la felicidad y la tristeza como emociones extremas, al centro se ubicaría un estado base con tonalidad emotiva neutra. Por el lado de las creencias, el estado base correspondería a la caracterización de su formación racional, es decir, a la no influencia de prejuicios calientes o de carácter motivacional. Al respecto, señala Elster, uno de los descubrimientos más sorprendentes de los últimos quince años en el campo de psicología fue que el estado base desde el punto de vista emotivo no coincide con el estado base cognitivo. Al contrario, las personas capaces de mirar al mundo sin prejuicios son depresivas, son *más tristes pero más sabios*. Por ello, parecería que

las distorsiones cognitivas son el precio que debemos pagar para tener la motivación necesaria para proceder en la vida. Si tuviéramos una idea más adecuada de lo que podemos hacer,

⁴⁸ Aunque, como hemos visto, en la medida en que una emoción (en cuanto excitación) está involucrada causalmente en la producción de los deseos, no hay en general ninguna razón para esperar que éstos sean racionales.

⁴⁹ Jon Elster, *Piú tristi ma piú saggi?*, p. 82.

emprenderíamos menos y realizaríamos menos [...] debemos tal vez aceptar el hecho de que la racionalidad cognitiva puede ser alcanzada sólo al precio de renunciar a todo *en relación a lo cual* queremos ser racionales. Aceptar este dato de hecho significa ser más tristes pero más sabios.⁵⁰

Esta posición, a pesar de sus diferencias, coincidiría también con Robert Nozick quien concluye su libro *La naturaleza de la racionalidad* con el siguiente párrafo:

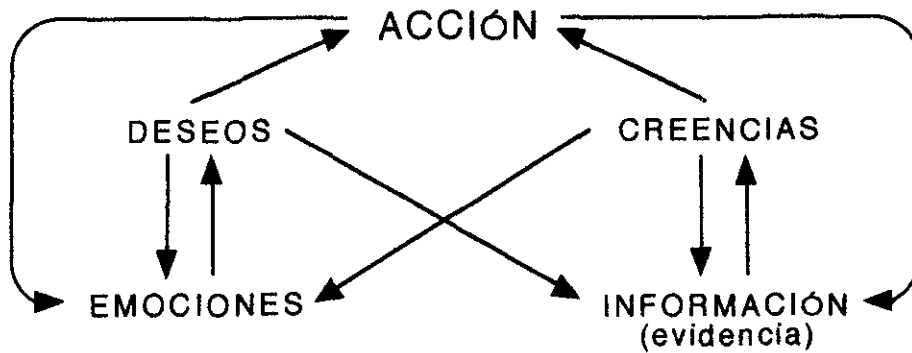
Nuestros principios fijan aquello por lo que vale nuestra vida, nuestros propósitos crean la luz en la que nuestra vida se baña, y nuestra racionalidad, así la individual como la coordinada, define y simboliza la distancia que hemos recorrido desde la mera animalidad. Por esos medios nuestras vidas llegan a significar más que lo que instrumentalmente dan. Y al significar más, nuestras vidas dan más.⁵¹

En este capítulo vimos dos aspectos de la racionalidad amplia de los deseos que también socavan las pretensiones iniciales de la TER. Sin pretender desechar esta teoría, los planteamientos que vimos remiten a la necesidad de incorporar otros aspectos para la explicación de una acción racional, tomando distancia de las concepciones ortodoxas de la TER. Con ello, las visiones economicistas del agente racional van quedando atrás y la racionalidad se presenta mucho más rica y compleja que un simple cálculo oportunista. Frente a la simple consistencia del conjunto de creencias y deseos, se incorporaron las nociones de juicio y autonomía, además de las emociones como posibilidad de racionalización de los deseos a semejanza del papel que juega la información en la racionalización de las creencias. Estos elementos se deben ubicar de manera dinámica ya que se interrelacionan de diversas maneras, conformando un complejo equilibrio que, podemos

⁵⁰ Jon Elster, *Ibid.*, p. 86.

⁵¹ Robert Nozick, *La naturaleza de la racionalidad*, pp. 241-242.

aventurar, no es tampoco demasiado estable. De esta manera, podríamos finalmente agregar al modelo modificado de la TER, un vínculo más entre la propia acción y los elementos básicos de racionalidad, cerrando así el círculo de un sistema complejo:



A MODO DE CONCLUSIONES

Para concluir retomemos algunos de los principales elementos que presentamos en los capítulos anteriores. La intención es encaminar las posibles características de una redefinición del concepto de razón. Como dijimos al inicio, la exigencia de esta redefinición no es exclusiva de la filosofía sino que se está dando en diversas disciplinas. Aquí nos centramos en la TER no sólo por ser una de las teorías contemporáneas en donde se defiende un concepto fuerte de razón sino también por el impacto que ha tenido en la toma de decisiones. En especial recuperamos los planteamientos de uno de sus más grandes defensores que es a la vez uno de sus más incisivos críticos internos: Jon Elster.

I

La TER se puede concebir de diversas maneras. La más elemental sería la que comprende a todas las conductas como un medio en el que se optimizan esfuerzos o costos para alcanzar de la mejor manera ciertos fines. Esta caracterización, que se queda en el nivel de la mera descripción, llevaría a plantear incluso la racionalidad del comportamiento animal. Por ejemplo, aun cuando no sepamos cómo es que hace sus cálculos, se puede observar un *comportamiento optimizador* cuando una leona busca atrapar a su presa. De acuerdo con esto, parecería que todo comportamiento finalmente es racional y el problema radica más bien en la descripción que hagamos de él. Una posición menos ortodoxa, plantea la necesidad de restringirse a las acciones humanas intencionales e incorporar la racionalidad de las propias razones de la acción. En este caso, los deseos y las creencias del agente, en tanto que razones de la acción, deben ser a su vez racionales. Como vimos, esta petición puede ser

simplemente a un nivel formal: el conjunto de creencias y de deseos deben ser consistentes.

Frente a esta *teoría estricta de la racionalidad*, Elster se ha preocupado especialmente por señalar la pertinencia de incorporar una *teoría amplia de la racionalidad*. Con ello se va dando un giro hacia la consideración de ciertos aspectos cualitativos que normalmente han sido relegados en la TER. El caso de las creencias es el más aceptado, al grado de ser incorporado en las condiciones de la estructura de la explicación de elección racional: las creencias además de ser consistentes deben estar fundadas *correctamente* en la evidencia disponible, la cual a su vez debe ser *adecuada* a la situación de elección. Más allá de los problemas propiamente epistemológicos que conlleva esta condición, remite a una de las principales nociones que atraviesan la historia de la filosofía: el *juicio*. En esta capacidad para evaluar y sintetizar la información pertinente, ya se hace presente un concepto de razón más débil, en tanto ésta debe enfrentarse a cada caso y situación alejándose de las grandes pretensiones de universalidad y univocidad.¹

Pero en especial, por lo que se refiere a los deseos, la *teoría amplia de la racionalidad* marca una gran distancia con las visiones más ortodoxas de la TER. Como vimos, aquí la noción clave es la *autonomía*. Si bien ésta queda delimitada de manera negativa en contraposición a la formación de deseos distorsionados por pulsiones o defectos cognoscitivos, la noción de *autonomía* remite a un ámbito problemático que se aleja del agente tradicional de la TER. Nos encontramos con que los deseos no pueden tomarse simplemente como dados, y lejos de pretenderse legislar sobre sus contenidos se enfatiza la formación consciente de los mismos.² De hecho, el sentido y relevancia de la

¹ Como ha señalado Carlos Pereda, -al diferenciar la razón *austera* de la *enfática*-, el ejercicio de la razón no tiene porque reducirse al cálculo exacto y a conclusiones necesarias, a criterios precisos, fijos y generales: "Encontramos racionalidad en donde encontramos argumentación, y ésta se dice de muchas maneras." (Carlos Pereda, *Razón e incertidumbre*, p. 9.)

² Antoni Domenèch por ejemplo distingue la racionalidad que denomina *erótica* en tanto tiene fuerza motivacional propia, y la racionalidad *inerte* que, en contraposición, no tiene dicha energía. En especial, frente a los deseos dados e inescrutables ligados a la naturaleza humana, el preguntarnos por la racionalidad de los deseos sería la base de una cultura moral que tendría

racionalidad instrumental se juega en esos deseos que se buscan realizar. Finalmente, elegir los deseos es elegir un modo de vida. En este sentido, podríamos plantear que la autonomía ligada a la elección de los deseos es la capacidad para elegir un modo de vida. Vida que es mucho más compleja que una simple jerarquización de preferencias guiadas por el egoísmo. Por ello, frente a las concepciones ortodoxas de la TER, autores como Amartya Sen han señalado:

el hombre *puramente* económico es casi un retrasado mental desde el punto de vista social. La teoría económica se ha ocupado mucho de este tonto racional arrellanado en la comodidad de su ordenamiento *único* de preferencias para todos los propósitos.³

Además de la noción de autonomía, Elster recurre también a la incorporación de las emociones en el modelo de la TER. Más allá de los problemas puntuales que puede implicar dicha incorporación en donde las emociones racionalizan los deseos, esta propuesta se puede presentar como una alternativa al planteamiento clásico del *interés*. Con base en el trabajo de Albert O. Hirschman, vimos que la irrupción moderna del término *interés* jugó un peculiar papel en el enfrentamiento entre razón y pasiones. Recordemos cómo en dicho término se conjugaba originalmente lo *mejor* de estos dos ámbitos enfrentados:

Se veía al interés participando de la mejor naturaleza de cada una de aquellas categorías, como la pasión del amor a sí mismo elevada y contenida por la razón, y como la razón dotada de dirección y fuerza por esa pasión. La forma híbrida de la acción

como piedra de toque la libertad de elección de los deseos. (Cf. Antoni Domenèch, *De la ética a la política*)

³ Amartya K. Sen, "Los tontos racionales: una crítica de los fundamentos conductistas de la teoría económica" en *Filosofía y teoría económica*, comps. Frank Hann y Martin Hollis, p. 202. El autor propone una estructura más compleja para incorporar la diversidad de aspectos que puede comprender el comportamiento humano: el ordenamiento de ordenamientos de preferencias.

humana resultante se consideraba libre de la naturaleza destructiva de la pasión y de la ineficacia de la razón.⁴

Pero el término *interés* se fue restringiendo a un sentido más estrecho como fue el amor al dinero y el deseo de ganancia, lo que permitió concebir cierto equilibrio estable en las relaciones interdependientes de los seres humanos. Esta concepción acabó por limitarse al ámbito económico, la cual, al menos desde el siglo XIX, fue también fuertemente criticada. Sin embargo, la vida humana es mucho más compleja y rica que el estrecho ámbito económico. Finalmente, a pesar de su indiscutible relevancia, el hombre racional no se circunscribe al hombre económico. Por ello, la incorporación de las emociones para racionalizar los deseos abre un nuevo horizonte de interpretación de la TER.

Dejando atrás la noción de *interés*, con la incorporación de las emociones se recupera nuevamente el señalamiento de que las pasiones y la razón no necesariamente conforman una dicotomía, aunque se puedan enfrentar también se pueden acompañar, yuxtaponer y recomponer sin disolver sus diferencias. Finalmente, frente a la concepción generalizada de una razón que se opone a las pasiones, encontramos que dentro de la propia TER las emociones pueden dejar de ser esas fuerzas oscuras que cartesianamente hay que dominar. Por el contrario, más cercanos a Hume, éstas pueden ser la base sobre la que adquiere sentido la elección racional. Por ello, aunque sea desde otra perspectiva, coincidiríamos con Remo Bodei cuando afirma,

La pasión aparecería de este modo como la sombra de la razón misma, como una construcción de sentido y una actitud ya íntimamente revestida de una inteligencia y cultura propias, fruto de elaboraciones milenarias, en tanto que la razón se revelaría, a su vez, como "apasionada", selectiva y parcial, cómplice de esas mismas pasiones que dice combatir.⁵

⁴ Albert O. Hirschman, *Las pasiones y los intereses*, p. 50.

⁵ Remo Bodei, *Una geometría de las pasiones*, p. 13.

Lejos de las visiones enfrentadas en donde por un lado está la razón, el orden, la universalidad y la necesidad, por el otro, las pasiones, el caos, la diversidad y la contingencia, vimos que las pasiones pueden tener cierta racionalidad o "lógica", así como también las grandes pretensiones de la razón pueden volverse problemáticas.

II

Retomemos algunas de las fisuras del lado de la razón. La TER parte de una concepción instrumental de la razón que caracteriza a la racionalidad como una racionalidad *optimizadora*. Si bien se trata de un principio mínimo para la acción e interacción de los individuos, esta racionalidad tiene sus alcances y límites. La racionalidad *optimizadora* se vuelve *problemática* cuando la información que dispone el agente no es completa o cuando debe tomarse la decisión dentro de juegos con soluciones subóptimas o sin estrategias dominantes. Como vimos, en estos últimos el caso paradigmático son los juegos sin solución (*la gallina*) ya que en ellos el agente no puede decidir racionalmente qué hacer. Sin embargo, los juegos con estrategias dominantes que llevan a soluciones subóptimas (*el dilema del prisionero*) son tal vez los más desconcertantes al refutar el optimismo smithiano: la racionalidad individual lleva al peor resultado para todos.

Nosotros en especial nos interesamos por dos casos extremos en donde la razón se enfrenta a una situación peculiar: lo más racional es abdicar de la propia razón. Con esta *abdicación racional de la razón*, encontramos que además de la racionalidad *amplia* y la racionalidad *problemática*, la racionalidad también puede ser *imperfecta e indeterminada*. La primera obedece al problema de la debilidad de la voluntad: frente a la debilidad de la voluntad, la racionalidad puede recurrir a la estrategia de *atarse a sí misma* tipificada en el caso de Ulises. Como vimos, con el mecanismo de *atarse a sí mismo*, el agente imperfectamente racional modifica las condiciones externas de su situación

(primer filtro) de tal manera que en un momento posterior no tenga ninguna opción que elegir (segundo filtro). Por otra parte, el caso de la racionalidad *indeterminada* obedece a la incertidumbre: en la toma de decisiones bajo incertidumbre la razón puede reconocer sus límites y aceptar que no hay razones que fundamenten la elección. Como vimos, el recurrir por ejemplo a loterías para la toma de decisiones bajo incertidumbre, es una manera racional de enfrentar los límites de la razón. En este sentido, el hiperracionalismo (el pretender que en toda situación siempre se puede elegir con base en razones) se presenta como una forma de irracionalidad.

A pesar de estos lúcidos planteamientos, aquí hay que hacer algunas observaciones. Si bien se puede reconocer el papel básico que juega en la vida del ser humano la racionalidad instrumental, al grado de que cualquier concepción de racionalidad debería poderla incorporar, la preeminencia de la TER sobre otras alternativas teóricas no es tan clara. Retomemos los señalamientos finales del capítulo dos.

El caso de la racionalidad *satisfecha o limitada* no puede reducirse simplemente a una forma de conducta maximizadora, porque justamente es a partir de una visión evolucionista de la racionalidad que se plantea el problema: a pesar de los indudables avances que hemos desarrollado (y que podremos desarrollar) a lo largo de la historia,

Estos avances aún nos dejan muy cortos en cuanto a nuestra facultad para manejar todas las complejidades del mundo. Pero el mundo –incluso el contemporáneo, afortunadamente– está vacío en su mayor parte, con la mayoría de las cosas sólo débilmente relacionadas entre sí, y es únicamente con este mundo con el que tiene que vérselas la razón humana.⁶

De esta manera, no tanto en contraposición a la TER, sino más bien delimitando sus pretensiones, la propuesta de una racionalidad *satisfecha o limitada* enfatiza el carácter provisional con el que tiene que trabajar la razón.

Esta razón eminentemente *situacional* tampoco se contrapone de manera necesaria a la caricatura del estructuralismo que presentó Elster.⁷ Obviamente, si las restricciones iniciales reducen el conjunto factible al grado de no dejar espacio para la elección, estas teorías se enfrentarían. Sin embargo, no necesariamente es el caso. A pesar del nivel de abstracción en que se mueve el modelo de la TER, finalmente remite a los seres humanos concretos, cuyos actos, deseos, emociones, creencias e información son de este mundo. En este sentido, hay ciertas determinaciones o condicionantes que desde una gran diversidad de perspectivas pueden ayudar a que esta racionalidad instrumental toque piso: cuestiones biológicas, físicas, sociales, históricas, culturas, ideológicas, políticas, morales, económicas, etc.⁸ Incluso, el reconocimiento explícito de Elster a la teoría de las normas sociales se puede interpretar apuntando hacia esa dirección.

Con esto, sin embargo, no pretendemos sugerir la posibilidad de una gran teoría general que abarcaría a todas las posibles perspectivas y alternativas teóricas. Lejos de un saber omniabarcante, intuimos más bien que estas parcelas teóricas se pueden recomponer de diversas formas dependiendo del problema a abordar. Seguramente, de acuerdo a las características del caso, algunas serán más pertinentes que otras.

En esta misma tónica, podemos ubicar la insistencia del propio Elster en la necesidad de explicar la formación de los deseos y las creencias del agente,

⁶ Herbert A. Simon, *Naturaleza y límites de la razón humana*, p. 136.

⁷ Con el fin de no cometer una injusticia, habría que aclarar que a lo largo de la obra de Elster también podemos apreciar una evolución, por lo que algunas afirmaciones tajantes dejan de ser para él mismo adecuadas. De hecho, podemos observar cómo el autor pasa de una posición inicial altamente optimista en cuanto al poder de la razón y la defensa de la TER, a una posición más moderada que no se cierra a las posibles aportaciones de otras perspectivas teóricas.

⁸ En este sentido, simpatizamos con las conclusiones a las que llegan Donald P. Green e Ian Shapiro en su libro *Pathologies of Rational Choice Theory*. Entre ellas podemos destacar especialmente las siguientes observaciones: 1) sin duda el cálculo estratégico puede ser una variable explicatoria pero no es la única (pensemos por ejemplo en las conductas tradicionales, las normas, la cultura que diferencia capacidades, las contingencias de la circunstancia histórica); 2) en lugar de construir modelos que no soportan su aplicación empírica, hay que recuperar la complejidad del ámbito social; y 3) en lugar de plantearse la disyuntiva entre TER o no, la pregunta es más bien por cómo la racionalidad interactúa con otros aspectos de la naturaleza y organización humanas que producen los fenómenos que se pretenden explicar.

así como los resultados (a veces imprevistos) de las interacciones individuales. Por ello, llega a plantear la necesidad (aunque estemos *años luz* de alcanzarla) de una teoría social que comprenda al mismo tiempo una explicación causal de la formación de deseos y creencias individuales, una explicación intencional de las acciones individuales y una explicación causal de los macroestados como resultado de dichas acciones.

III

A lo largo de la obra de Jon Elster hemos rastreado algunas de sus principales reflexiones sobre la TER. Sin pretender ser exhaustivos, ya que esta obra además de extensa abarca una gran diversidad temática,⁹ recuperamos algunos de los elementos que consideramos especialmente significativos para un desencantamiento de la de razón.

Frente a las visiones y los discursos que nos presentan a una razón fuerte que es capaz no sólo de dominar al mundo sino también de marcar el camino para la transformación de la sociedad y de los propios hombres, trabajos como los de Elster permiten tomar una sana y necesaria distancia. A partir de las críticas internas a la TER, podemos destacar los siguientes elementos para una razón desencantada:

1. Lo racional no necesariamente está ligado a lo universal, lo necesario, lo óptimo y lo unívoco.
2. Lo racional no necesariamente se contrapone a lo pasional, lo particular, lo contingente y lo diverso.
3. Lo racional tiene límites más allá de los cuales el pretender recurrir a razones se vuelve irracional.

⁹ En la página de Jon Elster en internet se registran hasta 1998: 54 libros y 141 artículos publicados. (Cf. <http://home.sol.no/hansom/elster.htm>)

Asimismo, frente a las pretensiones de elaborar una teoría general sobre las acciones racionales, trabajos como los de Elster nos advierten que sus objetivos (establecer proposiciones generales e invariables) es un *sueño ilusorio*.¹⁰ Sin embargo, la alternativa a ello no es un método meramente descriptivo. Entre estos dos extremos, el autor destaca la pertinencia de incluir el estudio de los *mecanismos*: modelos o cadenas causales específicos que en cada caso permiten explicar *a posteriori* las acciones e interacciones individuales. Como vimos, ejemplos de dos mecanismos, uno racional y otro irracional, son el *atarse a sí mismo* y las *uvas verdes*.

De hecho, como individuos estamos insertos y nos constituimos en un medio social altamente complejo y nosotros mismos somos mucho más complejos que el abstracto agente inicial de la TER. Las razones de la acción racional, es decir, los deseos, las emociones, las creencias y la información, no se vinculan de manera lineal y mecánica, sino que se interrelacionan y retroalimentan de diversas maneras. No somos ángeles ni demonios. No siempre que podemos somos racionales y no siempre podemos ser racionales. Algunas veces somos egoístas, y otras altruistas. En ocasiones nos dejamos llevar por las "pasiones", y a veces somos especialmente "reflexivos", además de que nuestras gamas de deseos y creencias cubren distintos tonos, niveles y órdenes.

Sin embargo, hay que enfatizar, la razón desencantada le sigue apostando a la razón. Una razón que lejos de las arrogancias¹¹ y el desprecio por el otro y lo diverso, trata de comprenderlo no imputando irracionalidad a la ligera. Como bien ha señalado Elster,

[...] aun si la racionalidad a menudo no tiene importancia, está ausente o es inestable, sin embargo hay un meollo de casos importantes en que es indispensable el modelo de la elección racional. Afirmando, además, que este modelo es lógicamente anterior

¹⁰ Cf. Jon Elster, "Introduction: Why Political Psychology?" en *Political Psychology*, pp. 1-34.

¹¹ Tomamos la sugerente expresión de otro texto de Carlos Pereda: *Crítica de la razón arrogante*.

a las alternativas, en el sentido de que el científico social siempre debe ser guiado por un postulado de racionalidad, aun si puede terminar descubriéndolo violado en muchos casos particulares. Esta suposición es un "principio de caridad" similar al tan a menudo empleado en la interpretación textual. Nunca deben tomarse en su valor aparente las contradicciones textuales, sino que se debe considerar si el contexto no nos dará una clave para la congruencia. De manera similar, siempre hay que ver muy minuciosamente la conducta aparentemente irracional para ver si, después de todo, surge allí alguna pauta. Huelga decirlo: no hay que extender demasiado la caridad, y puede llegar un punto en que el observador simplemente tenga que afirmar que, hasta donde él lo comprende, la conducta no tiene un motivo racional.¹²

De esta manera, hay que insistir, un acto por sí mismo no puede ser considerado racional ni irracional, sólo se le puede atribuir el primer o segundo adjetivo dependiendo en cada caso de sus complejas relaciones con los deseos, emociones, creencias e información del agente. A nivel teórico, el reto de la razón desencantada sigue siendo descubrir las posibles conexiones causales, pero ya sin la pretensión de una gran teoría general y sin caer tampoco en meras descripciones. A nivel individual el reto no es más sencillo, se trata de una razón que está ligada a la noción de *autonomía*, en última instancia al pensamiento crítico que es capaz de tomar distancia de sí mismo para reconocer la propia diversidad y construir creencias y deseos de segundo orden. En este sentido, la apuesta de la razón desencantada nos seguiría remitiendo al viejo Aristóteles:

[...] todo aquel que sea capaz de vivir según su propia voluntad, debe proponerse algún fin del vivir bien, como el honor, la gloria, la riqueza o la cultura, y fijando en él sus ojos ejecutará todos sus actos, toda vez que es signo de una gran demencia el no ordenar uno su vida en relación con un fin. De la mayor importancia es el que cada uno determine dentro de sí mismo, sin precipitaciones ni negligencia, en qué cosa de las que nos rodean consiste el bien

¹² Jon Elster, *Ulysses and the Sirens*, pp. 116-117.

vivir y cuáles son las condiciones indispensables para que esto sea accesible a los hombres.¹³

Más allá de las pretensiones aristotélicas, la razón desencantada está abierta a la diversidad y a la provisionalidad (tanto del otro como de uno mismo), en cuyos fines sin embargo, y esto no hay que olvidarlo, se juega la elección de un modo de vida. Un cuestionado *bien vivir* que no puede (o no debería) desentenderse de las condiciones indispensables para que efectivamente pueda ser accesible a todos los hombres.

¹³ Aristóteles, *Ética eudemia*, 1214b, p. 3.

REFERENCIAS BIBLIOGRÁFICAS

- Aguilar Rivero, Mariflor (coord.), *Límites de la subjetividad*, México, UNAM/Fontamara, 1999.
- Aguilar Villanueva, Luis F., *Weber: la idea de ciencia social*, 2 vol., México, UNAM/Porrúa, 1988.
- Anderson, Perry, *El Estado Absolutista*, 3ª ed., tr. de S. Juliá, México, Siglo XXI, 1982.
- Aristóteles, *Ética eudemia*, tr., introd. y notas de A. Gómez Robledo, México, UNAM, 1994.
- Aristóteles, *Ética Nicomaquea*, en *Obras*, tr., est. prel, preámbulos y notas de F. de P. Samaranch, 2a. ed., Madrid, Aguilar, 1977.
- Arrow, Kenneth J., *Social Choice and Individual Values*, Cowles Foundation of Yale University, 1970.
- Axelrod, Robert, *La evolución de la cooperación. El dilema del prisionero y la teoría de juegos*, tr. de L. Bou, Madrid, Alianza, Alianza Universidad No. 474, 1986.
- Bacon, Francis, "De la verdad" en *Ensayos sobre moral y política*, introd. de T. Brachet, tr. de A. Roda Rivas, México, UNAM, Nuestros Clásicos No. 42, 1974.
- Berlin, Isaiah, *Contra la corriente. Ensayos sobre historia de las ideas*, tr. de H. Rodríguez Toro, 2a. reimp., Madrid, FCE, 1992.
- Biblia de Jerusalén*, nueva ed., Bilbao, Desclée de Brouwer, 1975.
- Bobbio, Norberto, "El modelo iusnaturalista" en *Sociedad y Estado en la filosofía política moderna*, tr. J. Fernández Santillán, México, Fondo de Cultura Económica, Col. Popular No. 330, 1986.
- Bobbio, Norberto, "La grande dicotomía" en *Dalla Struttura alla funzione. Nuovi studi di teoria del diritto*, 2a. ed., Milano, Comunità, 1984.

- Bobbio, Norberto, "La regola della maggioranza e i suoi limiti" en *Soggetti e potere. Un dibattito su società civile e crisi della politica*, ed. V. Dini, Napoli, Bibliopolis, Saggi Bibliopolis No. 13, 1983.
- Bodei, Remo, *Una geometría de las pasiones. Miedo, esperanza y felicidad: filosofía y uso político*, tr. de J. R. Monreal, Barcelona, Muchnik, 1995.
- Churchalan, Paul. M., *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*, tr. de M. N. Mizraji, Barcelona, Gedisa, 1992.
- Código Civil para el Distrito Federal*, pról., rev. y comentarios de G. Trejo, México, SISTA, 1994.
- Cruz, Manuel (coord.), *Acción humana*, Barcelona, Ariel, 1997.
- Davidson, Donald, *Essays on Actions and Events*, Oxford, Clarendon Press, 1980.
- Descartes, Rene, *Discurso del método*, tr., pról y notas de M. García Morente, 14a. ed., México, Espasa Calpe, Austral No. 6, 1978.
- Descartes, René, *Las pasiones del alma*, est. prel., notas y tr. de J. A. Martínez Martínez, Madrid, Tecnos, Clásicos del Pensamiento 126, 1997.
- Descartes, René, *Reglas para la dirección del espíritu en Dos opúsculos*, 2ª ed., introd. de L. Villoro, México, UNAM, Nuestros Clásicos 10, 1972.
- Dieterlen, Paulette, "El individualismo metodológico" en *Sociológica* No. 14, septiembre-diciembre 1990.
- Dieterlen, Paulette, *Marxismo analítico: explicaciones funcionales e intenciones*, México, UNAM, Facultad de Filosofía y Letras/Coordinación Gral. de Estudios de Posgrado, Seminarios, 1995.
- Domènech, Antoni *De la ética a la política. De la razón erótica a la razón inerte*, Barcelona, Crítica, Crítica/Filosofía 8, 1989.
- Echeverría, Bolívar, *Las ilusiones de la modernidad*, México, UNAM/El Equilibrista, 1995.
- Edgington, Dorothy, "The Logic of Uncertainty" en *Crítica, Revista Hispanoamericana de Filosofía*, vol. XXVII, No. 81, diciembre de 1995.

- Elster, Jon, *Argomentare e negoziare*, tr. de G. Rigamonti, Milano, Anabasi, 1993.
- Elster, Jon, *Domar la suerte. La aleatoriedad en decisiones individuales y sociales*, introd. de Antoni Domènech, tr. de C. Castells, Barcelona, Paidós/ICE de la Universidad Autónoma de Barcelona, 1991.
- Elster, Jon, *Economics. Análisis de la interacción entre racionalidad, emoción, preferencias y normas sociales en la economía de la acción individual y sus desviaciones*, tr. de I. Cudich, Barcelona, Gedisa, 1997.
- Elster, Jon, *El cambio tecnológico. Investigaciones sobre la racionalidad y la transformación social*, 2ª ed., tr. de M. Mizraji, Barcelona, Gedisa, 1992.
- Elster, Jon, "Introduction" a *Rational Choice*, ed. J. Elster. Oxford, Basil Blackwell, 1986.
- Elster, Jon, *Piú tristi ma piú saggi? Razionalità ed emozioni*, tr. de M. Spranzi, Milano, Anabasi, 1994.
- Elster, Jon, *Political Psychology*, Cambridge, Cambridge University Press, 1993.
- Elster, Jon, "Sadder but wiser? Rationality and the emotions" en *Social Science Information*, No. 24, 1985.
- Elster, Jon, *Sour Grapes. Studies in the subversion of rationality*, Cambridge, Cambridge University Press/Editions de la Maison des Sciences de L'Homme, Paperback Library, 1985.
- Elster, Jon, "Sour grapes – utilitarianism and the genesis of wants" en *Utilitarianism and Beyond*, eds. Amartya Sen y Bernard Williams, Cambridge, Cambridge University Press/Editions de la Maison des Sciences de L'Homme, 1982.
- Elster, Jon, *Solomonic Judgements. Studies in the Limitations of Rationality*, New York, Cambridge University Press/Editions de la Maison des Sciences de L'Homme, 1989.
- Elster, Jon, *The Cement of Society. A Study of Social Order*, New York, Cambridge University Press, 1989.

- Elster, Jon, "The Nature and Scope of Rational-Choice Explanation", en *Actions and Events, Perspectives on the Philosophy of Donald Davidson*, editado por Ernest LePore y Brian P. McLaughlin, Oxford, Basil Blackwell, 1985.
- Elster, Jon, *Tuercas y tornillos. Una introducción a los conceptos básicos de las ciencias sociales*, 2ª ed., tr. de A. Bonanno, Barcelona, Gedisa, 1991.
- Elster, Jon, *Ulysses and the Sirens. Studies in rationality and irrationality*, ed. revisada, Cambridge, Cambridge University Press/Editions de la Maison des Sciences de L'Homme, Paperback Library, 1984.
- Elster, Jon, *Una introducción a Karl Marx*, 2ª ed., tr. de M. García Aldonate, México, Siglo XXI, 1992.
- Eurípides, *Hipólito en Tragedias*, pról. de C. García Gual, tr. de J. Alemany y Bolufer, Madrid, Edaf, Biblioteca Edaf 149, 1983.
- Fischer, H. R., Retzer, A. y Schweitzer, J. (comps.) *El final de los grandes proyectos*, tr. de J. Legris, Barcelona, Gedisa, 1997.
- Freud, Sigmund, "Esquema de psicoanálisis" en *Obras completas*, vol. XXIII, ed. de J. Strachey, tr. de J. L. Etcheverry, Buenos Aires, Amorrortu, 1976.
- Freud, Sigmund, "Más allá del principio de placer" en *Obras completas*, vol. XVIII, ed. de J. Strachey, tr. de J. L. Etcheverry, Buenos Aires, Amorrortu, 1976.
- Fukuyama, Francis, *El fin de la historia*, tr. de P. Alexandre, mimeo, 1989.
- Gargani, Aldo (ed.), *Crisi della ragione. Nuovi modelli nel rapporto tra sapere e attività umane*, Torino, Einaudi, Einaudi Paperbacks 106, 1979.
- Green, Donald P. y Shapiro, Ian, *Pathologies of Rational Choice Theory. A Critique of Applications in Political Science*, New York, Yale University Press, 1994.
- Habermas, Jürgen, *Teoría de la acción comunicativa*, 2 vol., tr. de M. JiménezRedondo, Madrid, Taurus, Ensayistas 278, 1987.
- Hacking, Ian, *El surgimiento de la probabilidad. Un estudio filosófico de las ideas tempranas acerca de la probabilidad, la inducción y la inferencia estadística*, tr. de J. A. Alvarez, Barcelona, Gedisa, 1995.

- Hahn, Frank y Hollis, Martin (comps.), *Filosofía y teoría económica*, tr. de E. L. Suárez, México, Fondo de Cultura Económica, Breviarios 398, 1986.
- Hansberg, Olbeth, *La diversidad de las emociones*, México, Fondo de Cultura Económica, 1996.
- Hargreaves, Shaun P. y Varoufakis, Yanis, *Games Theory: A Critical Introduction*, London, Routledge, 1995.
- Harrison, Ross (ed.), *Rational Action. Studies in Philosophy and Social Science*, Cambridge, CUP, 1979.
- Harsanyi, John, "Morality and the theory of rational behaviour" en *Utilitarianism and Beyond*, eds. Amartya Sen y Bernard Williams, Cambridge, Cambridge University Press/Editions de la Maison des Sciences de L'Homme, 1982.
- Hegel, G. W. F., *Fenomenología del espíritu*, 4ª reimp., tr. de W. Roces, México, Fondo de Cultura Económica, 1981.
- Hegel, G. W. F., *Filosofía del derecho*, pról. de J. Garzón Bates, México, UNAM, Nuestro Clásicos 51, 1975.
- Hirschman, Albert O., *Las pasiones y los intereses. Argumentos políticos en favor del capitalismo antes de su triunfo*, tr. de E. L. Suárez, México, FCE, Sección de Obras de Economía, 1978.
- Hobbes, Thomas, *Leviatán o la materia, forma y poder de una república eclesiástica y civil*, 2ª reimp., tr. y prefacio de M. Sánchez Sarto, México, Fondo de Cultura Económica, 1984.
- Hollis, Martin, "Rational Man and Social Science" en *Rational Action. Studies in Philosophy and Social Science*, ed. Ross Harrison, Cambridge, CUP, 1979.
- Homero, *Odisea*, int. y notas de J. Alsina, tr. de F. Gutiérrez, Barcelona, RBA, 1995.
- Hume, David, *Investigación sobre el entendimiento humano*, tr. de F. Ramos González, México, Gernika, 1994.

- Hume, David, *Tratado de la naturaleza humana*, 2 vol., ed. F. Duque, Madrid, Editora Nacional, Biblioteca de Literatura y el Pensamiento Universales 22, 1981.
- Hurwicz, Leonid, "The Theory of Economic Behavior", en *The World of Mathematics*, vol. II, ed. J. R. Newman, New York, Simon and Schuster, 1956.
- Josephus, F., *The Jewish War*, ed. gral. G. Cornfeld, Israel, Zondervan Publishing House, 1982
- Kant, Immanuel, *Crítica de la razón pura*, 4ª ed., pról., tr. y notas de P. Ribas, Madrid, Alfaguara, 1985.
- Kant, Immanuel, "¿Qué es la Ilustración?" en *Filosofía de la historia*, 3ª reimp., pról. y tr. de E. Ímaz, México, Fondo de Cultura Económica, Colección Popular 147, 1987.
- Kaplan, Abraham, "Sociology Learns the Language of Mathematics" en *The World of Mathematics*, vol. II, ed. J. R. Newman, New York, Simon and Schuster, 1956.
- Koyré, Alexandre, *Del mundo cerrado al universo infinito*, tr. de C. Solís Santos, 5a. ed., México, Siglo XXI, 1986.
- La Fontaine, Jean de, *et. al.*, *Las mejores fábulas*, Madrid, Edimat, 1998.
- LePore, Ernest y McLaughlin, Brian P. (eds.), *Actions and Events, Perspectives on the Philosophy of Donald Davidson*, Oxford, Basil Blackwell, 1985.
- Lukes, Steven, *El individualismo*, tr. de J. L. Alvarez, Barcelona, Península, Homo Sociologicus No. 8, 1975.
- Lyotard, Jean-François, *La condición postmoderna. Informe sobre el saber*, tr. de M. Antolín Rato, México, Red Editorial Iberoamerica, 1990.
- Lyotard, Jean-François, *La posmodernidad (explicada a los niños)*, 1ª reimp., tr. de E. Lynch, México, Gedisa, 1989.
- MacIntyre, Alasdair, *Historia de la ética*, 4ª reimp., tr. de R. J. Walton, Barcelona, Paidós, Paidós Básica 2, 1991.

- Maquiavelo, Nicolas, *El príncipe* en *Obras Completas*, ed. P. Alvarez Tabío, La Habana, Ed. de Ciencias Sociales, Instituto Cubano del Libro, 1971.
- Mari, G. (ed.) *Moderno-postmoderno. Soggetto, tempo, sapere nella società attuale*, Milano, Feltrinelli, Saggi, 1987
- Marramao, Giacomo, *L'ordine disincantato*, 1ª reimp., Roma, Editori Riuniti, Biblioteca Minima, 1986.
- Marramao, Giacomo, *Poder y secularización*, pról. de S. Giner, tr. de J. R. Capella, Barcelona, Península, Homo Sociologicus 46, 1989.
- Morgenstern, Oskar, "Teoría de los juegos" en *Metodología y crítica económica*, selecc. de Camilo Dagum, tr. de E. L. Suárez, México, FCE, El Trimestre Económico, Lecturas 26, 1978.
- Morgenstern, Oskar y von Neumann, J., *Theory of Games and Economic Behaviour*, Princeton, N. J., Princeton University Press, 1953.
- Muguerza, Javier, *La razón sin esperanza (Siete trabajos y un problema de ética)*, 2ª ed., Madrid, Taurus, Ensayistas 148, 1986.
- Nagel, Thomas, *The Possibility of Altruism*, Princeton, New Jersey, Princeton University Press, 1978.
- Newman, J. R. (ed.), *The World of Mathematics*, New York, Simon and Schuster, 1956.
- Nozick, Robert, *La naturaleza de la racionalidad*, tr. de Antoni Domènech, Barcelona, Paidós, Paidós Básica 74, 1995.
- Pascal, Blaise, "A la muy ilustre Academia de Ciencias Parisina 1654" en *Obras matemáticas. Selección de textos*, introd. de J. R. Martínez, tr. de S. Ramírez Castañeda, México, UNAM, MATHEMA, 1995.
- Pascal, Blaise, *Pensieri*, 1a. reimp., tr., introd. y notas de P. Serini, Milano, Arnoldo Mondadori, 1985.
- Pears, D. F., "La akrasia y el razonamiento práctico" en *Segundo Simposio Internacional de Filosofía*, vol 2., tr. de S. Bello, México, UNAM, 1987.
- Pereda, Carlos, *Crítica de la razón arrogante. Cuatro panfletos civiles*, México, Taurus, 1999.

- Pereda, Carlos, *Razón e incertidumbre*, México, Siglo XXI/UNAM, 1994.
- Pico della Mirandola, Giovanni , *Discurso sobre la dignidad del hombre*, tr., est. prel. y notas de A. Ruiz Díaz, Buenos Aires, Goncourt, 1978.
- Prigogine, Ilya, *El fin de las certidumbres*, 5ª ed., tr. de P. Jacomet, Santiago de Chile, Andrés Bello, 1997.
- Rabossi, Eduardo, "La filosofía de la acción y la filosofía de la mente" en *Acción humana*, coord. Manuel Cruz, Barcelona, Ariel, 1997.
- Rescher, Nicholas, "Choice Without Preference. A Study of the History and of the Logic of the Problem of 'Buridan's Ass'", en *Kant-Studien*, vol. 51, No. 2, 1959-1960.
- Rescher, Nicholas, *La racionalidad. Una indagación filosófica sobre la naturaleza y la justificación de la razón*, tr. de S. Nuccetelli, Madrid, Tecnos, 1993.
- Rescher, Nicholas, *La suerte. Aventuras y desventuras de la vida cotidiana*, tr. de c. Gardini, Barcelona, Andres Bello, 1997.
- Resnik, Michael D., *Elecciones. Una introducción a la teoría de la decisión*, tr. de S. Villarrea y B. Rodríguez, Barcelona. Gedisa, 1998.
- Rossi, Paolo, "*Idola della modernità*" en *Moderno-postmoderno. Soggetto, tempo, sapere nella società attuale*, ed. de G. Mari, Milano, Feltrinelli, Saggi, 1987.
- Saab, Salma, "El problema del auto-engaño y el modelo de niveles de creencias" en *Límites de la subjetividad*, coord. Mariflor Aguilar Rivero, México, UNAM/Fontamara, 1999.
- Savage, L. J., *The Foundations of Statistics*, Nueva York, 1954.
- Sen, Amartya K., "Los tontos racionales: una crítica de los fundamentos conductistas de la teoría económica", en *Filosofía y teoría económica*, comps. Frank Hahn y Martin Hollis, tr. de E. L. Suárez, México, FCE, Breviarios No. 398, 1986.

- Sen, Amartya y Williams, Bernard (eds.), *Utilitarianism and Beyond*, Cambridge, Cambridge University Press/Editions de la Maison des Sciences de L'Homme, 1982.
- Silva, Carmen, "El naturalismo humeano y el problema mente-cuerpo" publicado en *El problema de la relación mente-cuerpo*, comps. Laura Benítez y José A. Robles, México, UNAM, 1993.
- Simon, Herbert A., "De la racionalidad sustantiva a la procesal" en *Filosofía y teoría económica*, comps. Frank Hahn y Martin Hollis, tr. de E. L. Suárez, México, Fondo de Cultura Económica, Breviarios 398, 1986.
- Simon, Herbert A., *Naturaleza y límites de la razón humana*, tr. de E. Guerrero Tapia, México, FCE, Breviarios 455, 1989.
- Smith, Adam, *Investigación sobre la naturaleza y causas de la riqueza de las naciones*, 4ª reimp., ed. E. Cannan, introd. de M. Lerner, tr. y est. prel. de G. Franco, México, Fondo de Cultura Económica, 1984.
- Stroud, Barry, *Hume*, 2ª ed., tr. de A. Ziri6n, México, UNAM, 1995.
- Suetonio, Cayo, *Los doce C6sares*, vol. 12, tr. de J. Arnal, Madrid, Sarpe, Biblioteca de la Historia, 1985.
- Tarski, Alfred, "Truth and Proof" en *Scientific American*, junio de 1969.
- Vajda, S., "Theory of Games" en *The World of Mathematics*, vol. II, ed. J. R. Newman, New York, Simon and Schuster, 1956.
- Villoro, Luis, *Creer, saber, conocer*, 4ª ed., México, Siglo XXI, 1987.
- Villoro, Luis, *El pensamiento moderno. Filosofía del Renacimiento*, México, El Colegio Nacional-FCE, Cuadernos de La Gaceta No. 82, 1992.
- Villoro, Luis, *El poder y el valor. Fundamentos de una 6tica pol6tica*, México, El Colegio Nacional-FCE, 1997.
- Wang, Hao, *A Logical Journey. From G6del to Philosophy*, Cambridge, MIT Press, 1996.
- Weber, Max, "La ciencia como vocaci6n" en *El pol6tico y el cient6fico*, 1ª reimp., introd. de R. Aron, tr. de F. Rubio Llorente, México, Alianza, Libro de Bolsillo 71, 1989.

Weber, Max, *La ética protestante y el espíritu del capitalismo*, 8ª ed., Madrid, Alianza, 1984.

Welsch, Wolfgang, "Topoi de la posmodernidad" en *El final de los grandes proyectos*, comps. Hans Rudi Fischer, Arnold Retzer y Jochen Schweitzer, tr. de J. Legris, Barcelona, Gedisa, 1997.

Williams, Bernard, *Problemas del yo*, tr. de J. M. G. Holguera, México, UNAM, 1986.