

## Cuando el abogado es culpable y el acusado inocente: por qué los argumentos habituales de defensa de la teoría de la elección racional no resultan convincentes para un realista científico \*

*Jaume López*

A la teoría de la elección racional se debe en gran medida el importante desarrollo teórico que ha experimentado la ciencia política (y otras ciencias sociales) en las dos últimas décadas. Sólo por este hecho es lógico que cuente con numerosos defensores. Asimismo, no obstante, también existe un número importante de detractores. El argumento principal de éstos para criticar la teoría de la elección racional es la falta de verosimilitud de algunos de los supuestos en los que se basa la teoría y el uso que sus defensores hacen de modelos muy simples con intenciones explicativas. En ellos se da una imagen muy simplificada de la realidad que resulta, por ello, en muchas ocasiones implausible. La reacción de los defensores ante estas críticas suele remitir hacia argumentos de carácter instrumentalista (y en gran parte de inspiración popperiana) que apelan a criterios como la parsimonia o la capacidad de falsación para resaltar la validez de la teoría y los productos que se derivan de ella. En realidad, es difícil que defensores y críticos se pongan de acuerdo en estas discusiones porque, sencillamente, tienen concepciones de la ciencia y de lo que es una explicación científica correcta completamente distintas. Se podría decir que se enfrentan desde “esquinas” opuestas del “campo epistemológico”. Desde una de

---

\* Este artículo recoge y resume partes de mi tesis doctoral titulada *Realismo científico y teoría de la elección racional. Cómo y por qué un realista debería defender la teoría de la elección racional*, codirigida por el Dr. Ignacio Sánchez-Cuenca (IESA Juan March, Madrid) y la Dra. Rosa Virós (UPF, Barcelona), y que fue defendida en mayo de 2002. Así mismo, reproduce, en gran parte, la comunicación presentada en el *V Congreso español de Ciencia Política y de la Administración* (La Laguna-Tenerife, 26, 27 y 28 de septiembre de 2001); este artículo obtuvo el premio AECPA 2002 a la mejor ponencia.

estas "esquinas", el realismo científico, la mayoría de los argumentos utilizados para defender el enfoque racional-económico no resultan convincentes.

*Palabras clave: teoría de la elección racional, racionalidad, epistemología, criterios de validez, explicación.*

---

## 1. TEORÍA DE LA ELECCIÓN RACIONAL: PARADA Y FONDA

La *teoría de la elección racional* está en la base de muchos de los avances teóricos que han producido las ciencias sociales en la segunda mitad del siglo xx. Podría decirse, sin resultar demasiado atrevido, que probablemente a ella se deba la principal revolución teórica o cambio de paradigma que se ha dado en estas ciencias en los últimos cincuenta años. En la ciencia política, por ejemplo, la aplicación de la teoría de la elección racional ha generado todo un nuevo marco teórico, una nueva forma de ver la política, que ha recibido distintos nombres: *nueva economía política, teoría política positiva, teoría política formal, aproximación económica en política, aproximación racional a la política...* Bajo estas distintas etiquetas se agrupan toda una serie de desarrollos teóricos y resultados que rompen con las explicaciones más tradicionales de fenómenos muy diversos, desde el voto a la competencia electoral entre partidos, desde el desarrollo de la burocracia a los fracasos en la organización de acciones colectivas.

No es éste el lugar para hacer un repaso sistemático de todos los resultados obtenidos en el marco del enfoque racional o racional-económico (para una buena panorámica véase, por ejemplo, Alt y Shepsle, 1990; Dunleavy, 1991; Elster, 1986; Heap *et al.*, 1992; Hindess, 1988; Lewin, 1991; McLean, 1987; Renwick Monroe, 1991; Self, 1993; Shepsle y Bonchek, 1997). En todo caso, hay que subrayar el hecho de que su aplicación ha cosechado importantes éxitos empíricos y teóricos. *Teóricos*, porque ha permitido desarrollar un buen número de nuevas hipótesis, modelos, teorías y líneas de investigación. *Empíricos*, en tanto que las consecuencias empíricas que se derivan de muchos de ellos se corresponden con la realidad <sup>1</sup>.

Ahora bien, como todo nuevo paradigma o enfoque sus resultados no sólo resultan novedosos, sino que, en muchos casos, han provocado una multitud de debates, a veces enconados, entre sus defensores y sus críticos. En el núcleo de estos debates se entrecruzan toda una serie de argumentos y contraargumentos acerca de la validez de los mismos. *Para sus defensores, los éxitos predictivos, la simplicidad de sus axiomas y modelos, la universalidad de su ámbito de aplicación, etc., son pruebas entre otras que avalan a*

---

1. Lo que no quiere decir por ello que nos resulten útiles para realizar predicciones. En la mayoría de los casos no es así, puesto que para su aplicación con fines predictivos se requiere, igualmente, de un conocimiento de las condiciones iniciales muy preciso del que no siempre disponemos.

*este paradigma como un instrumento útil, quizás el más útil desarrollado nunca en las ciencias sociales, para realizar explicaciones. Para sus detractores, en cambio, existen múltiples razones para rechazar este paradigma como inadecuado e inútil para proporcionarnos explicaciones realmente válidas de lo social.* La implausibilidad de la mayoría de los supuestos en los que se basan, empezando por la definición del *homo economicus*, es la piedra angular sobre la que descansan gran parte de estas críticas en contra de la explosión de resultados que este paradigma ha producido y que pretenden pasar por genuinas explicaciones sociales.

Suele ocurrir en todas las ciencias: cuando se produce una novedad teórica que permite una fácil aplicación a diversos campos de la disciplina, la comunidad científica se vuelca en la exploración de sus posibles virtualidades generándose un gran número de resultados en forma de artículos, libros, etc. Queda para un segundo momento otro tipo de actividad intelectual que no busca la aplicación empírica de la teoría, sino una reflexión sobre sus fundamentos. Es una tarea que la mayoría de las veces pasa a un segundo plano porque sus resultados nunca son espectaculares y en el peor de los casos pueden implicar haber de tirar por la borda algunos de esos esfuerzos anteriores. No obstante, la necesidad de hacer una “parada y fonda” tras la “gran explosión” para analizar los elementos teóricos sobre los que se sustentan todos esos resultados es evidente <sup>2</sup>. La simple posibilidad de aplicación y de generación de hipótesis desde un determinado marco teórico, por muy fructífero que pueda resultar, no garantiza *per se* la solvencia del mismo. En otras palabras, no nos garantiza que aquello que nos ofrece resulten ser genuinas explicaciones científicas <sup>3</sup>.

En el presente artículo pretendo analizar los argumentos que más habitualmente se han utilizado para defender que la teoría de la elección racional y, en general, el enfoque racional-económico son, efectivamente, instrumentos válidos para elaborar explicaciones científicas. Para ello será necesario, primero, establecer con claridad las premisas teóricas de este enfoque y sus principales interpretaciones (sección 2), así como conocer los términos en los que se ha establecido el debate entre detractores y defensores (sección 3). Una vez hecho esto será necesario, igualmente, distinguir las distintas concepciones de la ciencia que, de forma implícita, se barajan en estas discusiones al tiempo

2. Jesús Mosterín, refiriéndose a la ciencia en general, hace esta misma reflexión (*La Vanguardia*, Libros, 27 de abril del 2001: 11): «La ciencia actual ha progresado tanto que su transmisión y desarrollo serían inconcebibles sin una extremada división del trabajo intelectual. El científico especializado sabe cada vez más sobre cada vez menos, hasta que llega a saberlo casi todo sobre casi nada. Eso es lo que le permite seguir avanzando. Sin embargo, la imagen fragmentada que nos ofrece el espejo roto de la ciencia ha de ser recompuesta en una visión global coherente y fiable».

3. Esta tarea suele reservarse a los filósofos de la ciencia. En el caso que nos ocupa, a los filósofos de la ciencias sociales. Me gustaría reivindicar desde aquí este espacio de reflexión en las ciencias sociales, y por consiguiente también en la ciencia política, cuya importancia parece pasar todavía hoy desapercibida en nuestro país a tenor de su poca presencia (o su inexistencia) en los departamentos universitarios.

que se constata que no es posible juzgar los resultados científicos (ni las metodologías utilizadas para conseguirlos) sin vincularlos a una determinada posición epistemológica en el seno de la cual se definen cuáles deben ser los objetivos de la ciencia y qué hay que entender por explicación. Básicamente conviene distinguir entre posturas antirrealistas (positivismo e instrumentalismo) y posiciones realistas (secciones 4 y 5).

Con todos estos instrumentos en la mano llegamos a la tesis central que pretendo defender aquí: la mayoría de los argumentos esgrimidos por los *rational choicers* pueden encajarse perfectamente en concepciones de la ciencia y de la explicación antirrealistas, pero no pueden resultar convincentes para un realista científico (sección 6). Ahora bien, conviene distinguir muy claramente entre los argumentos utilizados para justificar la validez de un resultado y el resultado mismo. Una cosa es afirmar que las razones que se han dado para defenderlos no le resultan convincentes a un realista por apelar a criterios o valores epistemológicos ajenos a su concepción de la ciencia y la explicación, y otra cosa muy distinta es que los resultados conseguidos con la aplicación del enfoque racional-económico no puedan ser, igualmente, defendidos por un realista, aunque por razones y con argumentos distintos. A esta distinción alude el título del artículo en el que —quiero subrayarlo una vez más— no se analizan los resultados conseguidos en el seno de este paradigma, sino que sólo se pasa revista a los argumentos aducidos habitualmente para defender su validez y que, desde una posición realista, no pueden ser más que criticados. Reservo para una próxima ocasión la tarea “más positiva” de abordar cómo podría elaborarse una nueva defensa del paradigma racional-económico más acorde con los criterios del realismo y qué tipo de interpretación de sus resultados requeriría.

## II. EL NÚCLEO DE LA TEORÍA Y DOS INTERPRETACIONES

Las premisas básicas de la teoría de la elección racional son muy simples y constituyen lo que podríamos denominar como *el núcleo de la teoría*. Todas las explicaciones elaboradas desde el enfoque racional-económico comparten estas premisas y, por tanto, el esquema argumental mínimo que las mismas proponen:

- i) Lo social se explica a través de lo individual. Es decir, sus explicaciones se fundamentan sobre el *individualismo metodológico*, que concibe los fenómenos sociales como resultado de las acciones de los individuos [Fenómeno colectivo → Acción individual]<sup>4</sup>.
- ii) Cualquier acción de un individuo explicable desde la teoría de la elección racional es fruto de una elección [Acción individual → Elección].

4. La flecha significa “remite a”, o “se explica por”.

- iii) Es decir, de una decisión consciente [Elección → Decisión].
- iv) Y racional. Lo que en este contexto quiere decir que se trata de una elección coherente y orientada a un fin. A este tipo de racionalidad se la suele conocer como *racionalidad instrumental* [Decisión → Racional].
- v) La elección es fruto de un cálculo de los costes y beneficios asociados a los distintos curso de acción posible [Racional → Cálculo].

En general, se considera que con ese cálculo el individuo busca la *optimización*. Ésta es una característica que suele vincularse al núcleo de la teoría, para algunos autores, incluso, éste es uno de los rasgos más característicos de la teoría. Sin embargo, como veremos, esta premisa resulta un tanto más problemática que las cinco anteriores y para que las diversas formas de interpretar la teoría de la elección racional puedan coincidir en un mismo —aunque mínimo— núcleo teórico es necesario distinguirla y no incluirla dentro de las cinco anteriores. ¿Qué se entiende por *optimizar*? Elegir el curso de acción que proporciona un *mayor beneficio neto* al individuo: éste compara los beneficios esperados de los distintos cursos de acción, así como los costes que puede asociar a cada uno de ellos, y elige aquel curso de acción en el que la diferencia es máxima, es decir, el beneficio neto es máximo. Igualmente, también puede significar, si los objetivos del individuo ya están prefijados, establecer una relación *eficiente* entre medios y fines, es decir, conseguir lo que uno se propone de la mejor manera posible (en este caso se busca el mayor beneficio neto rebajando al máximo los costes de la acción). Cuando introducimos funciones de utilidad la optimización se convierte en la *maximización* de una función.

¿Cómo se explica, pues, a tenor de estas premisas un fenómeno social? Siempre como resultado (previsto o no) de acciones individuales que se interpretan como la mejor acción posible para el individuo. Para que éste sea capaz de actuar de esta manera hay que suponer que:

1. Elige constantemente. O al menos, que cada vez que actúa elige. Aunque también puede interpretarse en sentido contrario: sólo es posible explicar desde la teoría de la elección racional si hay elección <sup>5</sup>.
2. El individuo dispone de un conjunto de creencias sobre cómo es el mundo y las consecuencias posibles de sus acciones, y un conjunto de deseos u objetivos. A partir de ellos es capaz de ordenar los distintos resultados posibles que pueden derivarse de

5. Esta última interpretación choca con una visión universalista de la teoría muchas veces defendida, según la cual la teoría es aplicable para explicar cualquier fenómeno social. No obstante, en contraposición, también puede afirmarse que —como hacen muchos *rational choicers*— las actuaciones de los individuos son fruto de una decisión en muchas más ocasiones de lo que habitualmente se piensa, y de aquí el carácter universal que puede tener la misma.

sus acciones construyendo así un orden de preferencias. Aunque esta operación puede resultar aparentemente sencilla y aproblemática hay que remarcar dos aspectos que a menudo pasan desapercibidos y que esta operación implica: para comparar diversos resultados es necesario que el individuo sea capaz de situarlos en un mismo plano valorativo, lo que no siempre resulta fácil; y en segundo lugar, es necesario no sólo que tenga unos deseos, sino unas creencias bien fundadas sobre los mismos; que conozca, por ejemplo, cuán estables son estos deseos especialmente cuando se relacionan con objetivos que se alejan en el tiempo (de lo contrario, es difícil que sus preferencias puedan ser consistentes; véase el siguiente punto) <sup>6</sup>.

3. Sobre el contenido de estas preferencias hay distintas posiciones, como veremos a continuación, pero todas ellas coinciden en exigir que éstas sean *consistentes*, es decir, deben tener unos mínimos requisitos de coherencia. De lo contrario, podría darse el caso de que un individuo actuase en un momento dado en contradicción con sus propias preferencias. Igualmente, la teoría misma estaría sujeta a posibles inconsistencias lógicas. Para que esto no ocurra es necesario, por ejemplo, que las preferencias sean transitivas: si alguien prefiere un helado de chocolate a uno de fresa, y uno de fresa a uno de vainilla, tendrá que preferir —no puede ser de otro modo— un helado de chocolate a uno de vainilla. Lo que parece ciertamente razonable. El requisito de transitividad es seguramente el más repetidamente señalado por los teóricos de la elección racional a la hora de describir las exigencias formales de la teoría. No obstante, acompañando a este requisito, hay que citar otras condiciones que presentan sobre todo relevancia desde un punto de vista matemático, pues se hacen necesarias para poder elaborar funciones de utilidad (que deben cumplir, al menos, con las propiedades matemáticas que definen a una función). Así, junto a la condición de transitividad hay que citar a las de reflexividad, complitud y continuidad <sup>7</sup> (a éstas hay que añadir algunas más en caso de decisiones bajo incertidumbre; véase, por ejemplo, Heap *et al.*, 1992).

Todos los defensores de la teoría de la elección racional coinciden en exigir estos mínimos requisitos para poder considerar que una elección —y, por extensión, una

---

6. Sobre estos temas, véase el artículo introductorio de James G. March [(1978) 1986: 165] y su previsión —fallida— final: «En los últimos veinte años, la contradicción que se produce entre la investigación centrada en la búsqueda de sentido en el comportamiento y la mejora del comportamiento se ha manifestado en el estudio prioritario de nuestras interpretaciones sobre el modo en que la información sobre las consecuencias futuras es recogida y procesada. Este esfuerzo se debe considerablemente a la idea de racionalidad limitada y a una concepción de las decisiones humanas limitada por las capacidades cognitivas de los seres humanos. En los próximos veinte años, sospecho que la contradicción derivará cada vez más a que prestemos atención a cómo nuestras creencias sobre las futuras preferencias se generan y utilizan».

7. *Reflexividad*: cualquier resultado es tan bueno como sí mismo. *Complitud*: cualesquiera dos resultados pueden ser siempre comparados y ordenados. *Transitividad*: si el resultado A es preferido a B, y B es preferido a C, entonces A tiene que ser preferido a C. *Continuidad*: siempre es posible hacer indiferente al individuo ante dos resultados modificando las proporciones de lo conseguido en cada uno (lo que excluye cualquier ordenación lexicográfica).

acción— es *racional*. En otras palabras, ¿qué es comportarse racionalmente? Es, como *mínimo*, hacerlo de acuerdo con los requisitos antes mencionados. Pero a partir de aquí existen dos posiciones distintas entre los teóricos de la elección racional.

a) *Una definición estricta de racionalidad*

Para algunos, para que la definición de *racionalidad* resulte completa hay que añadir dos requisitos más. El primero:

- vi) El individuo dispone de perfecta información y/o perfecta capacidad de cálculo.

De otro modo —se preguntan aquellos que defienden esta concepción de la racionalidad—, ¿cómo podríamos calificar de *racional* la acción de alguien que no dispone de toda la información relevante para tomar una decisión, o que no sabe cómo tratarla? Tras esta cuestión se encuentra, igualmente, una reflexión de carácter más metodológico (y, seguramente, de mayor relevancia en las discusiones entre las dos concepciones de la racionalidad): los defensores de esta sexta premisa señalan que *sin ella es imposible, o cuanto menos muy difícil, derivar predicciones precisas de las cinco premisas anteriores, aun conociendo las preferencias de los individuos*. Si no suponemos, por ejemplo, que los individuos disponen de perfecta capacidad de cálculo, la teoría de la utilidad esperada, uno de los productos teóricos que se derivan de la teoría de la elección racional, pierde gran parte de su potencial predictivo.

La segunda condición que suelen poner en su definición de *racionalidad* los que defienden la anterior es que:

- vii) El sujeto es egoísta. Es decir, sólo mira para sí, o en otras palabras: el bienestar de los demás no entra a conformar las propias preferencias; al sujeto sólo le interesa el bienestar propio.

Aunque esta definición a veces es defendida con argumentos que tratan de demostrar la naturaleza intrínsecamente egoísta de la naturaleza humana, en la mayoría de los casos se justifica con el mismo argumento metodológico que la anterior, a saber: sin esta condición las predicciones se hacen muy difíciles, si no imposibles. A veces, incluso, se argumenta que sin estas dos últimas condiciones la teoría de la elección racional adolece de una estructura tautológica.

A la definición de la racionalidad constituida por todas estas premisas, es decir, desde la i) a la vii), se la suele denominar *definición estricta de la racionalidad* porque es muy exigente (quizás, también podría denominarse *restrictiva*). Sólo los individuos que cumplen con todas estas características se les puede considerar racionales, es decir, con-

seguirán pasar por este “estrecho paso hacia la racionalidad”. La imagen del individuo que surge tras la aplicación de todas estas premisas (supuestos o condiciones) es la de un individuo que es un perfecto calculador de los costes y beneficios que implican sus acciones, que actúa para maximizar su utilidad. Para ello, parte de una escala de preferencias en las que están ordenados distintos resultados elaborada bajo un principio egoísta. A esta particular definición del individuo se la suele conocer en la literatura científica como *homo economicus*. Se puede decir, por tanto, que *un homo economicus es un homo rationalis de acuerdo con la definición estricta de racionalidad*.

b) *Una definición amplia de racionalidad*

No todo el mundo está de acuerdo con esta definición de racionalidad. Otros autores, defensores igualmente de la aproximación racional en las ciencias sociales —y coincidiendo en las primeras cinco premisas—, consideran, sin embargo, que estas dos últimas son demasiado exigentes. Por eso amplían la definición de racionalidad para que en ella quepan un número mayor de individuos. Para empezar, afirman que:

vi') No es necesario que un individuo disponga ni de perfecta información ni de perfecta capacidad de cálculo.

Consideran que los individuos no siempre disponen de la información relevante para hacer determinados cálculos o que la forma de calcular y de elegir no siempre coincide con la que señalan los defensores de la visión estrecha. Se producen errores de cálculo o simplificaciones. A veces el individuo actúa siguiendo unas reglas de actuación (*rules of thumb*) que le llevan a abandonar el ideal de optimización. Puesto que no siempre le es posible calcular qué curso de acción es el mejor en todos los sentidos, ya sea por no disponer de toda la información relevante o porque es incapaz de determinar cuál es la elección adecuada, el individuo puede optar por elegir, sencillamente, un buen curso de acción, atendiendo a distintas consideraciones, sin que esto signifique elegir el mejor, optimizar. Pero esto no quiere decir que la elección no haya sido racional; lo es porque el individuo elige, ciertamente, una buena opción aunque no sea la mejor. Sobre esta base —la idea de las limitaciones computacionales de los individuos y la sustitución del objetivo optimizador por un objetivo satisfactor— se erige la llamada *teoría de la racionalidad limitada* (cuyo principal referente son los trabajos de Herbert A. Simon; véase por ejemplo, 1983, 1985, 1990).

Para algunos, sin embargo, este punto no justifica la distinción entre dos tipos de racionalidad. Afirman que “satisfacer” unos objetivos puede traducirse como «optimizar teniendo en cuenta los costes de conseguir la información relevante para la decisión», lo que hace que la acción observada no sea en sí misma la más eficiente, sino que

la eficiencia deba encontrarse en el conjunto del proceso decisional. Igualmente, otros consideran que el seguimiento de reglas o métodos para actuar no es más que una forma de evitar costes de información. William Riker (1990: 172) es uno de estos autores. Afirma que reconocer que los individuos no disponen de perfecta información ni de perfecta capacidad de cálculo «elimina la innecesaria distinción entre racionalidad estricta y racionalidad limitada». Asimismo, Jackman (1993: 282), siguiendo a Riker, expresa esta idea con rotundidad: «*Satisfacer*, en la terminología de Simon, no es más que una forma concisa de decir “optimización bajo condiciones de información limitada e incertidumbre”».

La segunda condición que relajan quienes sustentan esta otra visión de la racionalidad es la del egoísmo y la sustituyen por la idea de autorreferencialidad:

vii') Un individuo racional no tiene por qué ser egoísta, pero sí ha de ser autorreferencial.

Conviene distinguir claramente ambos conceptos. Por *autorreferencial*<sup>8</sup> se entiende que la conformación de las preferencias depende *exclusivamente* de la satisfacción que las distintas opciones genera en el *propio* sujeto. Es decir, sencillamente, que la acción de los individuos depende de la *propia* función de utilidad (y no de la de ningún otro). Se podría decir que el individuo no está alienado, sabe lo que quiere, y actúa según sus preferencias y las de nadie más. Pero nada se dice acerca de los elementos que influyen en la elaboración de éstas. Si en la conformación de las propias preferencias influye el bienestar de los demás el individuo, aunque autocentrado, difícilmente será estrictamente egoísta. Para los defensores de esta concepción de la racionalidad, tan racional puede ser un individuo egoísta como una altruista. Así, por ejemplo, el altruismo de un monje que sólo piensa en ayudar a los demás no se distingue para la teoría de la elección racional de cualquier otra acción, y se explica, por tanto, de igual forma: el monje actúa buscando maximizar su utilidad; es decir, de forma autointeresada, si bien, en este caso, sólo contribuyendo al bienestar de los demás conseguirá su objetivo.

El tema del altruismo y el egoísmo ha generado muchas confusiones<sup>9</sup>; a veces el problema estriba en no distinguir entre *acciones* y *preferencias*. En el enfoque racional, donde tiene realmente sentido hablar de altruismo y egoísmo es en relación a las accio-

8. En la literatura anglosajona se utilizan indistintamente los términos *self-regarding*, *self-centered*, *self-oriented*, o incluso *non-altruistic*, para hacer referencia a la idea de autorreferencialidad de las preferencias.

9. Esta confusión ha producido, entre otras cosas, que muchas veces críticos y defensores de la teoría de la elección racional no se entiendan. Así, por ejemplo, cuando se ha señalado que los individuos no son siempre egoístas, y se ha planteado esto como una objeción al enfoque racional-económico, sus defensores han respondido invariablemente que los sujetos son siempre egoístas aunque parezcan altruistas porque sus acciones siempre buscan satisfacer sus propias preferencias. Respuesta que nunca ha satisfecho las voces críticas que, a su vez, han contestado que incluir las acciones altruistas en el ámbito de la teoría de la elección racional supone estirar los límites de la teoría hasta extremos que la convierten en tautológica: resulta que todo es explicable (con el mismo marco teórico podemos explicar dos resultados contrarios: egoísta y altruista). El

nes, no en las preferencias. Las acciones de los individuos pueden ser altruistas (si buscan satisfacer a los demás) o egoístas. Pero el orden de preferencias, desde el punto de vista de la teoría de la elección racional, invariablemente se construye a partir de comparar la distinta satisfacción que al sujeto le proporcionan los resultados de diferentes acciones. Podría decirse, estirando un poco el concepto, que las preferencias son siempre “egoístas” porque el mecanismo de ordenación siempre reside en el propio sujeto, y en su propia capacidad —y de nadie más— de disfrutar los distintos resultados, aunque éstos afecten a los demás, positiva o negativamente. En todo caso, también puede distinguirse entre *individuos* que tienen en cuenta el bienestar de los demás a la hora de formar sus preferencias (individuos altruistas) e individuos que no (como es el caso del *homo economicus*)<sup>10</sup>. Desde esta visión ambos podrían resultar igualmente racionales<sup>11</sup>. Está claro que esta segunda concepción de la racionalidad es menos exigente que la primera. Por eso podemos calificarla como *visión amplia de la racionalidad*. Desde este punto de vista hay muchos más sujetos racionales. En otras palabras, el rasero de definición es mucho menos estrecho.

Estas dos visiones o interpretaciones de la definición de racionalidad suponen sendas formas de entender y aplicar la teoría de la elección racional. No se puede decir, no obstante, que ambas gocen de unos éxitos y una expansión comparables. La definición estricta es, sin duda, la que va a la cabeza en este sentido. De ella se derivan las principales consecuencias que se asocian al enfoque racional-económico y que tan popular lo han hecho en algunos ámbitos académicos. Tanto la teoría de la utilidad esperada como la teoría de juegos, por ejemplo, son productos derivados de la interpretación estricta. Asimismo, la mayoría de los modelos y explicaciones producidos por este enfo-

---

problema es que, en la polémica, unos (los críticos) se están refiriendo a acciones altruistas, y naturalmente (como afirman sus defensores) que la teoría de la elección racional puede explicar acciones altruistas! Mientras que los otros (los defensores) están entendiendo *preferencias* altruistas (o no autorreferenciales), y naturalmente que no es posible (como señalan sus detractores) una preferencia “altruista” (o no autorreferencial), al menos desde la óptica de la elección racional (pues esto sólo podría deberse a una alienación del individuo)!

10. Sen (1985) defiende la necesidad de distinguir entre tres supuestos independientes de la teoría económica que conforman la visión tradicional del *homo economicus*: 1. El *bienestar* de un individuo depende exclusivamente del consumo propio, no le afecta el bienestar de nadie más (bienestar autocentrado); 2. El único *objetivo* de un individuo es maximizar su función de utilidad, no tiene en cuenta ningún otro objetivo (como, por ejemplo, justicia social) (objetivo autocentrado); 3. Toda *elección* de un individuo se encamina a conseguir los objetivos propios, no sigue ninguna pauta o norma de comportamiento (elección autocentrada).

11. Como afirma Riker (1990: 173), la teoría de la elección racional «no especifica ningún fin en particular». Se supone que todo el mundo es autointeresado y elige lo que le proporciona mayor satisfacción, pero el contenido del autointerés no se especifica. Si este autointerés se manifiesta de forma altruista o egoísta, es «una cuestión de investigación empírica, no un juicio *a priori*» (Riker, 1990: 242). Esto ha hecho que, de nuevo, algunos consideren que la visión estrecha no es más que un caso particular de la visión amplia. A mí, sin embargo, me parece que intentar subsumir una definición en la otra plantea más inconvenientes que ventajas. Para empezar, no nos permite entender correctamente porque las explicaciones que nos ofrece la teoría de la elección racional se han acompañado de argumentos tan diversos para su justificación.

que que más interés han suscitado tienen como uno de sus supuestos de partida al *homo economicus*. No es extraño así que, en general, la teoría de la elección racional se indentifique sobre todo con una interpretación estricta mucho más que con una amplia. Su “ventaja” con respecto a esta última no es difícil de descubrir: los supuestos vi) y vii) dotan de mayor capacidad informativa y predictiva a la teoría; con ellos resulta más fácil generar modelos e hipótesis explicativos. Por su parte, una interpretación amplia quizás resulte, en principio, más verosímil, pero lo que, en todo caso, es seguro es que su capacidad informativa y predictiva resulta muy escasa, y podría decirse que, interpretada así, la teoría de la elección racional no parece distinguirse demasiado de cualquier otro enfoque intencional en el que, simplemente, se afirme que todos los individuos actúan de acuerdo con unos deseos y unas creencias.

### III. ARGUMENTOS Y CONTRAARGUMENTOS EN TORNO A LA VALIDEZ DE LA TEORÍA DE LA ELECCIÓN RACIONAL

#### 1. *Críticas*

Las críticas más habituales que se han lanzado en contra de la teoría de la elección racional pueden agruparse en tres grandes grupos a tenor de sus argumentos o del aspecto negativo que destacan, a saber: *críticas metodológicas*, *críticas psicológicas* y *críticas sociológicas*.

#### a) *Críticas metodológicas: entre la tautología y la incapacidad predictiva*

Algunos críticos señalan que muchas explicaciones del enfoque racional-económico son tautológicas <sup>12</sup>, lo que favorece que con un mismo esquema teórico puedan explicarse

12. A veces, no obstante, como señala Opp (1999), se hace un uso no adecuado del término “tautología” para llevar a cabo algunas de estas críticas. Por “tautología” ha de entenderse estrictamente una afirmación *analíticamente* verdadera, es decir, una afirmación cuya verdad puede determinarse analizando *exclusivamente* el significado de sus términos. (Por ejemplo, «todos los solteros no están casados».) Muchas veces, cuando se acusa a una explicación de tautológica, se quiere decir, en rigor, que (1) es *circular* porque se apoya sobre hipótesis auxiliares que no disponen de un soporte empírico independiente a la teoría que pretenden apoyar. Por ejemplo, cuando se afirma que alguien actúa porque los beneficios que le genera esa acción exceden sus costes. Dado que no hay forma independiente de conocer esos beneficios, la afirmación resulta circular: actúa porque los beneficios superan los costes, pero la única manera de saber que esto es así es observar la actuación del individuo. Otras veces, se quiere decir que la explicación *no es falsable* (2). En este sentido, hay que notar que efectivamente ninguna tautología es falsable, pero, como veremos más adelante, existen muchos enunciados que tampoco son falsables y que no pueden considerarse, en absoluto, tautológicos (por ejemplo, «la lucha de clases es el motor del progreso histórico»: el famoso libro de Karl Popper *La miseria del historicismo* trata de esta cuestión). Finalmente, a veces, se usa el calificativo “tautológico” para hacer

fenómenos completamente opuestos. Quienes apuntan estas críticas señalan que las explicaciones que proporciona la teoría de la elección racional «no son más que la explicación de una definición» (Tilly, 1995: 1595), el desarrollo de un ideal —particular— de racionalidad del que se hacen derivar toda una serie de consecuencias prácticas, dejando en un segundo plano explicativo las afirmaciones de contenido empírico.

Otro tipo de críticas metodológicas hacen hincapié en las dificultades de la teoría para realizar predicciones novedosas y no caer en movimientos *ad hoc*. Por ejemplo, Green y Shapiro (1994)<sup>13</sup> se centran en el análisis de cuatro áreas centrales de las aplicaciones de la teoría de la elección racional en las cuales las promesas de elaborar predicciones novedosas han cosechado amplios fracasos. Por ejemplo, en el ámbito del comportamiento político, la participación electoral no parece verse afectada por —ni reflejar— a menudo los modelos crecidos al amparo del enfoque racional-económico, que sugieren escasas participaciones. Lo mismo puede decirse, en el ámbito de la acción colectiva, de las dificultades para dar cuenta de la contribución desinteresada de la gente en la consecución de beneficios colectivos o, en la negociación colectiva, de la estabilidad de las decisiones políticas que los modelos parecen descartar. Green y Shapiro también hacen referencia a los modelos espaciales de competencia electoral y la diferencia entre sus previsiones de convergencia en el espacio ideológico y la realidad política de muchos países.

En esta dirección también se destaca que, precisamente, para dar cuenta de estas “anomalías” que presenta la aplicación de la teoría se realizan constantemente movimientos *ad hoc* que trasladan la carga explicativa (y de la prueba) a premisas auxiliares ajenas a la teoría, lo que la convierte fácilmente en infalsificable.

Ésta constituye otra crítica habitual a la teoría de la elección racional: a menudo sus explicaciones se construyen siguiendo un esquema que las convierte en infalsificables. Esto ocurre sobre todo cuando no existen pruebas independientes de la propia acción del individuo para justificar las afirmaciones que se hacen sobre sus preferencias. Es decir, que se explica la acción a partir de unas preferencias que se infieren de la acción misma. Cuando esto es así no resulta difícil “demostrar” que la acción del individuo maximiza sus intereses (que hemos descubierto a partir de la acción misma) y explicar la acción como resultado de un comportamiento racional. Cualquier acción parece encajar en una definición de racionalidad, pero esto no nos informa de nada, puesto que las preferencias ya han sido inferidas “racionalmente” por el propio investigador. Cuan-

---

referencia, en realidad, a *un enunciado trivial* (3), con un muy bajo nivel informativo. Por ejemplo, cuando se afirma que «la teoría de la elección racional es psicología popular formalizada» (Rosenberg, 1995: 78).

13. La publicación del libro de Green y Shapiro *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science* desencadenó un gran número de debates, réplicas y contrarréplicas, algunas de las cuales están recogidas en Friedman, J. (ed., 1996): *The Rational Choice Controversy. Economic Models of Politics Reconsidered*.

do no es posible desvincular la explicación de la acción de la explicación de las preferencias que la han causado, afirmar que la acción es racional puede resultar una afirmación vacía y las explicaciones construidas bajo este esquema adquieren un aire de funcionalismo injustificado (Elster, 2000). Pese a que todas estas críticas están muy relacionadas, encuadrándose todas ellas en una perspectiva metodológica, nótese la diferencia entre ellas: las que subrayan el aspecto tautológico critican sobre todo que en las explicaciones de la teoría de la elección racional todo parezca, precisamente, racional; las que destacan el carácter infalsificable advierten que el mecanismo explicativo siempre se mantiene inalterable y que en ellas sólo varían las hipótesis auxiliares, sobre las cuales recae todo el peso explicativo y la carga de la prueba.

b) Críticas psicológicas: ¿quién ha dicho que seamos así de racionales?

A diferencia de las anteriores, las críticas psicológicas y sociológicas no se centran en la *aplicación* que se ha hecho de la teoría en la producción de explicaciones, sino que dirigen sus ataques a la teoría misma remarcando la implausibilidad de sus premisas. Como resultado, lo que remarcan estas críticas no es el carácter tautológico, vacío o infalsable de las explicaciones elaboradas bajo la aproximación económico-racional, sino, más rotundamente, su falsedad por proporcionarnos una imagen de los individuos y, en definitiva, del mundo que contradice numerosas observaciones empíricas sobre cuestiones de gran relevancia para la teoría. En la medida en que el eje central de la teoría es una determinada definición de la racionalidad (en la que, pese a las variaciones, todos sus defensores coinciden), estas críticas conllevan un cuestionamiento de la misma y vienen a decirnos que o los seres humanos son irracionales porque no actúan como la teoría postula, o que afirmar que son racionales quiere decir necesariamente otra cosa.

Todas y cada una de las premisas básicas de la teoría son atacadas con argumentos esencialmente de carácter científico que intentan demostrar su implausibilidad (a excepción de la premisa del individualismo metodológico que por su naturaleza sólo puede recibir críticas de carácter filosófico-epistemológicas). Así, estas críticas señalan que no todas las acciones del individuo son fruto de una elección [contra la premisa ii)], no siempre actuamos de acuerdo a una racionalidad instrumental [contra la premisa iv)], ni efectuamos un cálculo de costes y beneficios [contra la premisa v)], ni nos mueve un mecanismo maximizador. Tampoco es cierto que nuestras preferencias sean siempre transitivas. Por contra, estas críticas remarcan la importancia de cuestiones como las emociones, las reglas heurísticas y los fallos cognitivos, los comportamientos satisfactores, las normas sociales, etc.

Pueden señalarse distintas líneas de investigación en las que se agrupan estas críticas. Una de las más destacadas hace referencia al concepto de *racionalidad limitada* y la lógica satisfactoria en la actuación de los individuos. Existe una amplia literatura científica

que coincide en señalar que los individuos no maximizan sus preferencias al actuar, sino que, *simplemente, satisfacen algunos de sus objetivos*. Esta literatura se nutre de autores de diversos campos. Por ejemplo, sobre todo desde la economía, diversos autores han destacado la complejidad de la información requerida para hacer frente a una estrategia optimizadora y han señalado el uso de reglas por parte de los individuos como medio para hacer frente a los retos que plantea el entorno y, de este modo, actuar. Desde la psicología, en cambio, se ha hecho más énfasis en las características y límites de los mecanismos cognitivos con los que cuentan los individuos para procesar la información (los llamados fallos o errores cognitivos). En la primera línea de investigación sobresalen los trabajos de autores como Herbert A. Simon (al cual ya he aludido) o Ronald A. Heiner (1983, 1985, 1989). En lo que respecta a la investigación de los fallos o ilusiones cognitivas, así como a las reglas heurísticas para hacerlos frente, se destacan los trabajos de Daniel Kahneman y Amos Tversky (1979, *et al.* 1982, 1984, 1986) especialmente en el estudio de decisiones tomadas en situaciones de incertidumbre<sup>14</sup>.

Unos autores hacen más hincapié en los problemas derivados de la obtención de información que imposibilitan que ésta sea perfecta o suficiente para optimizar recursos (el sujeto no dispone de la información necesaria para ello); otros destacan los problemas “biológico-computacionales” existentes para organizar y tratar esa información en vistas a desarrollar una estrategia optimizadora (el sujeto no dispone de los medios necesarios para tratar esa información). Así, mientras unos se fijan en la calidad del *input*, otros lo hacen en el proceso de cálculo, pero en ambos casos se destaca la imposibilidad de conseguir un resultado óptimo y un comportamiento racional como el que postula la teoría.

### c) Críticas sociológicas: ¿acaso somos átomos flotando en el vacío?

Las críticas sociológicas, al igual que las críticas psicológicas, pretenden mostrar que las explicaciones que se elaboran con la teoría de la elección racional se fundamentan sobre premisas falsas. Fundamentalmente, las críticas sociológicas son de dos tipos. Las que se dirigen en contra de una definición de racionalidad exclusivamente instrumental, y las que se centran en la crítica al individualismo metodológico omnipresente en el enfoque racional-económico.

Las primeras destacan la importancia de las normas sociales, las convenciones, las rutinas, los roles, etc., en el origen de las acciones de los individuos. Unas normas que nos vienen dadas, que se adquieren de una manera no sistemática ni del todo consciente

---

14. El libro de Massimo Piattelli Palmarini (1993) es una interesante y muy divulgativa introducción a estos temas (a destacar la amplia bibliografía temática).

(por socialización, imitación, etc.) y que condicionan las decisiones suficientemente como para que ocupen un lugar muy destacado en las explicaciones del comportamiento. Aunque el sujeto decida, la identificación de los costes y beneficios pertinentes va a ser completamente mediada por una serie de normas sociales; de ahí su enorme relevancia. La visión de las normas que proporcionan estas críticas no es la de unas simples reglas del juego que los individuos utilizan para conseguir sus objetivos, o sea, en su propio provecho (tal y como son vistas en el contexto del paradigma racional-económico). Se trata, más bien, de unas normas *modeladoras* de los objetivos (de las preferencias) y de los recursos (sociales, cognitivos) disponibles (que el sujeto aprecia como disponibles). El individuo no se erige ante las normas con independencia de ellas, no las acata en un acto consciente (como si fueran un texto legal), sino que estas normas constituyen al sujeto (contribuyen a dotarlo de una determinada identidad) y explican, así, su proceder.

Junto a este aspecto también se remarca la importancia de las acciones que no constituyen un medio para la consecución de un fin, sino que su razón de ser y su explicación ha de circunscribirse a la acción misma y a los efectos inmediatos que produce en el sujeto. Existe una amplia literatura sobre la importancia de las acciones calificadas —entre otras denominaciones— como *expresivas* (Abelson, 1995; Taylor, 1996). Desde una perspectiva racional amplia, no obstante, se ha considerado que estas acciones también podían ser racionales, aunque orientadas hacia el proceso y no el resultado de la acción.

La sociedad, entendida como un todo holístico que trasciende las acciones particulares de los individuos, resulta una pieza fundamental que, para los críticos sociológicos, el paradigma racional-económico ha subestimado como elemento explicativo. La sociedad no es una simple estructura de costes y beneficios, o un conjunto de arreglos institucionales explícitos de los que se aprovecha el individuo más capaz para elaborar sus estrategias. La sociedad ha de entenderse, afirman estos autores, como el producto de una evolución cultural de la que han ido surgiendo unas reglas y pautas de comportamiento compartidas que proporcionan condiciones suficientes para elaborar funciones de elección, y criterios para reconocer sobre qué principio de racionalidad éstas se han de fundamentar, dado el carácter de la acción a efectuar. Las críticas sociológicas proponen un modelo alternativo de individuo, el llamado *homo sociologicus*<sup>15</sup>.

---

15. La oposición entre *homo economicus* y *homo sociologicus* ha servido tradicionalmente a diversos autores —entre los que destacan Brian Barry (1970) y Karl Popper (1945)— para realizar una amalgama entre ciencia política/economía-libertad-elección-racional-individualismo, por un lado, y sociología-determinismo-conductismo-irracional-holismo, por otro, que les permite hacer una defensa del *homo economicus* a partir de motivos éticos que nada tiene que ver con razonamientos sobre su valor explicativo. De este modo, se ha recalado que mientras que el *homo economicus* es un hombre libre que elige según sus preferencias, el *homo sociologicus* es un esclavo *determinado* por roles y normas sociales. Por lo que, en la medida en que nos consideremos libres, sólo el *homo economicus* puede explicar nuestras acciones, sólo éste constituye una buena descripción de un individuo, y sólo unas teorías basadas en ella podrán resultar exitosas. No obstante, la asociación entre *homo economicus* y libertad, por un lado, y *homo sociologicus* y determinismo, por otro, es una falacia. En

Siguiendo por esta línea, otro tipo de críticas que se suelen enmarcar dentro de las llamadas críticas sociológicas —y que proceden, principalmente, de sociólogos e historiadores— son aquellas que remarcan lo innecesario que es muchas veces el referirse a los comportamientos estrictamente individuales para llevar a cabo buenas explicaciones de fenómenos sociales (Bhargava, 1993). ¿Qué nos aporta el individualismo metodológico tan íntimamente ligado a la teoría de la elección racional? Muchas veces, afirman estos autores, nada. Ahora bien, no hay que confundir el individualismo *ontológico* con el *individualismo metodológico*. Posiblemente muy pocos, si alguno, de estos autores no coincidirá en afirmar que en la base de todo fenómeno social están los comportamientos individuales. Los “átomos” que forman la realidad social son, sin lugar a dudas, los individuos. Pero una cosa bien distinta es afirmar que, exclusivamente haciendo referencia a estos individuos y a sus comportamientos individuales, es posible proporcionar una explicación de un fenómeno social, como afirman a menudo los defensores del enfoque racional-económico.

## 2. *Defensas habituales*

Ante estas críticas, los argumentos que suelen aparecer en defensa de la teoría de la elección racional se concentran en destacar *la validez del método* seguido en la elaboración de explicaciones. Dejan de lado, completamente, cualquier debate como los que hemos visto sobre la verosimilitud de las premisas de la teoría, descartándolo como si, desde un punto de vista metodológico, resultase una discusión menor. El tipo de argumentos que se esgrimen son los siguientes:

— *Los supuestos no importan.*

Una primera línea de defensa consiste, precisamente, en afirmar que todas las críticas que se centran en presentar evidencia empírica para desacreditar las premisas de la teoría de la elección racional son irrelevantes porque *nadie pretende que esos supuestos se correspondan con la realidad*. El objetivo que se busca con la aplicación de la teoría

---

este sentido, quizás sea suficiente con señalar que el determinismo es un rasgo de cualquier explicación, puesto que ésta requiere de un supuesto de necesidad para su elaboración. Explicamos señalando condiciones que asociamos a un resultado. Tan determinado en este sentido se encuentra un *homo economicus* cuya acción es el resultado de la función que maximiza su utilidad, como el *homo sociologicus* cuya acción es el resultado de conformarse a un rol o normal social dada. Por otra parte, tan *propias* son las preferencias a partir de las cuales actúa el *homo economicus*, como lo es la aceptación de unas reglas que, en último término, depende de la voluntad del *homo sociologicus* (la posibilidad de romperlas en cualquier momento se mantiene siempre abierta).

es otro. Como muestra de este tipo de argumentos pueden servirnos las afirmaciones que se hacen al inicio de un manual de teoría de juegos (Morrow, 1994: 20):

«Nosotros no asumimos que el proceso de decisión sea una serie de cálculos literales (...) La teoría de la utilidad no es un intento de explicar los procesos cognitivos de los individuos (...) Nuestro objetivo no es explicar lo cognitivo, sino entender actos políticos».

En este mismo sentido (Morrow, 1994: 7, énfasis añadido):

«El criterio adecuado para juzgar el “realismo” de un argumento es la exactitud de sus conclusiones».

Se destaca, en cambio, el papel que ha de tener la elaboración de predicciones como medio con el que poder comprobar la validez de las hipótesis explicativas propuestas. En este sentido, suele afirmarse que:

— *Lo más importante es que se puedan realizar predicciones y que éstas encajen con lo observado. Más aún, es suficiente y necesario que la predicción sea exitosa para poder aceptar como válida una explicación.*

Esta idea se repite empleando distintas fórmulas. Como, por ejemplo, en Crain y Tollison (eds., 1990):

«La utilidad de la aproximación de la elección pública radica en su habilidad para generar hipótesis contrastables y predicciones sobre el comportamiento político».

— *No hay que hacer referencia a aquello que no se puede observar.*

Los argumentos anteriores se suelen conectar con la idea de que, en la medida en que la mayoría de las variables de carácter psicológico —e, incluso, sociológico— a las que aluden las críticas son de difícil observación (si es que ésta es posible), no deberían entrar a formar parte de las explicaciones que se ofrecen. Olson (1965: 61, énfasis añadido), por ejemplo, pone de manifiesto esta idea al descartar el uso de motivaciones morales en sus explicaciones:

«No es posible conseguir una prueba empírica de la motivación que hay tras la acción de una persona; no es posible decir definitivamente si un determinado individuo actuó por razones morales o por otras razones en un caso particular».

Por eso, de acuerdo con estos autores, no se ha de hacer referencia a este tipo de variables que «podría convertir a la teoría en incontrastable» (p. 61) y sujeta a posibles

movimientos *ad hoc* (véase el próximo argumento defensivo). También Becker (1976; con Stigler, 1977) se refiere a que los gustos no se pueden observar para defender a la teoría de la elección racional y su particular visión acerca de la estabilidad de las preferencias. Desde su perspectiva sólo han de importarnos las constricciones del entorno a las que ha de hacer frente el sujeto, suponiendo unos objetivos o preferencias iguales y universales para todos los individuos.

— *Hay que evitar cualquier tipo de movimiento ad hoc.*

Uno de los mayores peligros de hacer referencia a variables de difícil observación es, a juicio de muchos de los defensores del enfoque racional-económico, sucumbir a movimientos *ad hoc* o *post hoc*. Estos movimientos implican hacer encajar la explicación en los datos recogidos tras oportunas modificaciones realizadas *a posteriori* de la hipótesis inicial. El problema es especialmente importante cuando la valoración de una explicación descansa exclusivamente sobre su capacidad de generar hipótesis falsificables, como ocurre en el marco de una metodología falsacionista.

Es muy usual que los defensores de la teoría de la elección racional señalen que, precisamente, uno de los rasgos destacados que distingue a esta teoría de las de tipo psicológico o sociológico desde donde se lanzan las críticas es que su formulación evita estos movimientos mientras que esto no es posible en estas otras teorías. Como una muestra de estas preocupaciones véase Becker [(1976) 1997: 49, énfasis añadido]:

«El supuesto de preferencias estables *sirve* como fundamento estable *para generar predicciones* sobre las respuestas a varios cambios y evita que el analista sucumba a la simple tentación de recurrir al socorrido cambio de preferencias para “explicar” cualquier clara contradicción de sus predicciones»<sup>16</sup>.

También en Olson (1982: 24-29) encontramos continuas referencias a este problema y a la necesidad de superarlo.

— *Hay que conseguir modelos/explicaciones elegantes, generales (de gran alcance) y sencillas. Más aún, si es posible, las teorías han de ser de ámbito universal. En otras palabras, la parsimonia de la teoría es un objetivo metodológico prioritario.*

Los defensores de la teoría de la elección racional han hecho referencia muchas veces a estas características para dotar a sus explicaciones de valor frente a las críticas.

16. En las traducciones al castellano de Becker no aparece el término *ad hoc* para calificar estos cambio de argumentos y ha sido sustituido por adjetivos como “circunstanciales”, “específicos”, “amañados”, etc. Sólo en la traducción de 1980 se mantiene el uso de esta partícula tal y como figura en el original.

Quizás un extremo de estas defensas —por otra parte, muy habituales— es la visión “universalizadora” del premio Nobel de economía Gary Becker [(1976) 1997: 57, énfasis en el original] para quien:

«el enfoque económico suministra un valioso y unificado marco para comprender *todo* el comportamiento humano».

O, igualmente (más adelante en la misma página, énfasis añadido):

«Hay que ver a todo el comportamiento humano como integrador de unos participantes que maximizan su utilidad desde un conjunto estable de preferencias y que acumulan una cantidad óptima de información y otros insumos en una variedad de mercados (...) Si esta argumentación es correcta, el enfoque económico nos brinda *un marco unificado* para la comprensión del comportamiento».

Ni las teorías psicológicas ni las teorías sociológicas permiten, al menos en su desarrollo actual, cumplir con estos objetivos metodológicos a los que sí, en cambio, puede hacer frente supuestamente la teoría de la elección racional.

La idea de que el objetivo de una teoría ha de ser tener un alcance universal y de que su estructura explicativa ha de ser lo más parsimoniosa posible (explicar el máximo de fenómenos con el mínimo de premisas) es una referencia constante cuando se destacan las virtudes de la teoría de la elección racional. Así, por ejemplo, Ordeshook (1993: 72, énfasis añadido) subraya que la teoría de la elección racional resulta «una teoría *coherente, parsimoniosa y deductiva*», rasgos que la distinguen entre las teorías sociales y la dotan de valor. En esa misma obra, un poco más adelante, Ordeshook (1993: 76, énfasis añadido) destaca que, de todos los logros de la teoría de la elección racional, «ninguno es tan importante como el de haber conducido hacia una *reintegración de la política y la economía bajo un paradigma común y una misma estructura deductiva*».

Compartiendo esta visión se pueden encontrar múltiples ejemplos. Monroe (1991: 2, énfasis añadido) afirma que la teoría de la elección racional es «uno de los paradigmas dominantes de la ciencia política y social, ofreciendo explicaciones penetrantes, rigurosas y *parsimoniosas*». La búsqueda de la universalidad y la parsimonia por encima de todo en las explicaciones elaboradas desde el enfoque racional-económico también se pone de manifiesto en comentarios como éstos (Goetze y Galderisi, 1989: 38, énfasis añadido):

«Si las funciones de utilidad y las percepciones difieren ampliamente [entre la gente], y si la gente tiene combinaciones muy diferentes de motivos altruistas y autointeresados, entonces la construcción de modelos explicativos podría verse frustrada. Formas de comportamiento *universal* podrían no ser descubribles».

En este mismo sentido (Enelow, 1981: 1077-1078):

«[El hecho de que] nos veamos forzados a decir que dos modelos diferentes —el modelo de voto sincero y el de voto sofisticado— deben usarse simultáneamente para explicar lo

que observamos, es una decisión que, como mínimo, *es contraria a la parsimonia científica* y reclama una explicación de por qué el comportamiento de algunos votantes debe explicarse con un modelo y el comportamiento de otros votantes con otro».

Estos razonamientos y similares son utilizados para destacar las ventajas de la teoría de la elección racional y de las explicaciones elaboradas con ella.

— *Si no hay una teoría mejor, entonces la que mejor funciona, pese a sus limitaciones, es válida.*

Este argumento lakatosiano se suele citar como el argumento definitivo en contra de todas las críticas anteriores. Se suele desacreditar a los críticos del enfoque racional-económico por no proponer teorías mejores, o con el mismo grado de parsimonia y elegancia. El conocido argumento de Lakatos (1970) es, en resumen, que una teoría debería ser sólo rechazada cuando aparece otra teoría que da cuenta de todo lo que la anterior teoría explicaba, y además realiza nuevas predicciones, produciéndose de este modo un progreso empírico. En palabras de Kenneth Shepsle (1995: 217): «no dejes ir nada hasta que tengas alguna otra cosa de la que asirte» (citado en Cox, 1999, y Bates *et al.*, 2000: 699). Dennis Chong también suscribe esta visión y afirma: «Una teoría no puede ser rechazada debido a hechos disconformes con ella. (...) Sólo puede ser sustituida por una teoría superior» (citado en Cohn, 1999: 30).

En Elster (1986: 26-27, énfasis añadido) puede observarse este argumento combinado con una llamada a la simplicidad:

«La economía neoclásica [en la que Elster incluye el enfoque racional-económico] será destronada si y cuando la teoría que postula la racionalidad limitada (*satisficing theory*) y la psicología aúnen fuerzas para producir una explicación *simple* y robusta de los distintos niveles de satisfacción, o la teoría sociológica aparezca con una teoría *simple* y robusta sobre la relación entre normas sociales y racionalidad instrumental. Hasta que esto ocurra, el dominio continuo de la teoría neoclásica está asegurado *dado que uno no puede derrotar a algo con nada*».

La discusión entre críticos y defensores de la teoría de la elección racional no ha pasado nunca del lanzamiento cruzado de los argumentos que he ido exponiendo. Un primer elemento que llama la atención al compararlos es que, a excepción de las críticas metodológicas, los críticos insisten en que las premisas de la teoría no son *verosímiles* y, por tanto, la teoría no es válida; mientras que los defensores les contestan que *el método seguido y los objetivos conseguidos son correctos* y, por tanto, la teoría es válida, como si se tratase de alguna versión del juego de los disparates («por aquí me han dicho... y por aquí me han contestado...»). Seguramente, más que de un juego de los disparates se trate de un diálogo de sordos en el que *cada uno de los interlocutores*

*tiene en mente cosas distintas cuando piensa acerca de qué buscamos conseguir con una explicación científica, o a qué llamamos conocimiento científico.*

#### IV. ¿VÁLIDO? ¿PARA QUÉ? LAS TRES “ESQUINAS” DEL “CAMPO EPISTEMOLÓGICO”

Pero, ¿cuándo decimos que una explicación científica es *válida*? Los científicos no suelen entrar en este tipo de cuestiones; sencillamente afirman que sus explicaciones o teorías son válidas cuando les parecen correctas. Pero, como ocurre en otros ámbitos, no son muy conscientes de que “lo correcto” no es único y que depende de lo que estemos persiguiendo, incluso si no somos del todo conscientes de ello. Todo conocimiento científico es hipotético, es decir, no definitivo. Esto es lógico porque con él afirmamos cosas que no son evidentes, que no se nos aparecen directamente ante nuestros ojos, y aunque así fuera, a estas alturas, ya estamos suficientemente prevenidos sobre lo relativo del sentido común y de aquello que “resulta evidente”. Por todo ello, todo conocimiento “de calidad” requiere pasar por unos “filtros” que garanticen, cuanto menos, que hay razones para poder considerarlo parte del acervo del conocimiento “de calidad” del que disponemos en un momento dado, es decir, del conocimiento científico<sup>17</sup>. A estos filtros podemos denominarlos “criterios de validez”.

En la historia de la ciencia pueden distinguirse dos formas muy distintas de entender qué es el conocimiento científico; cada una de ellas implica unos criterios y metodologías divergentes. La base de esta distinción resulta difícil de definir; posiblemente descansa sobre una diferente actitud, más o menos optimista, en relación a las habilidades y capacidades cognitivas de los seres humanos. ¿Hasta qué punto podemos conocer el mundo?, ¿hasta qué punto podemos fiarnos de nuestras habilidades? *Los optimistas* piensan que los seres humanos están dotados de un aparato cognitivo que puede permitirles conocer el mundo tal y como es; el desarrollo de métodos e instrumentos para perfeccionar este aparato puede proporcionarnos un progresivo perfeccionamiento del conocimiento que tenemos del mundo, el cual constituye, por tanto, una imagen del mismo que se va puliendo y precisando. *Los pesimistas*, en cambio, desconfían de las posibilidades de nuestro aparato cognitivo. Así, consideran totalmente infundada la idea de que lo que nuestros sentidos y reflexiones nos proporcionan es un conocimiento directo del mundo. A diferencia de los primeros, no creen que seamos capaces de tener nunca

---

17. Como afirma Jorge Wagensberg (*El País*, 28 de junio del 2001: 12) un científico es siempre un “creedor” racional. Nunca puede estar seguro del todo de su conocimiento —y lo sabe—, pero —por eso— exige todas las garantías que la realidad puede ofrecer en un momento y lugar antes de aceptarlo. Su actitud crítica lo distancia totalmente del creyente ingenuo (que exige muy pocas garantías) o del crédulo (que no exige ninguna, como suele ocurrir, por ejemplo, con los creyentes de una religión).

una imagen de cómo es verdaderamente el mundo y por ello adoptan una actitud más defensiva y modesta en relación a lo que los seres humanos conocen o pueden llegar a conocer. La conclusión de estos planteamientos se materializa en un sencillo razonamiento: si no somos capaces de conocer el mundo tal y como es, entonces, nuestro conocimiento ha de centrarse en buscar —o fundamentarse en— otro tipo de objetivos, éstos sí al alcance de los seres humanos. (El “¿cómo es la realidad?” que mueve a los realistas pasa a ser sustituida por una más modesta y prioritaria interrogación sobre “¿cómo es la ciencia?": qué características ha de tener lo que puedo afirmar sobre el mundo para poder ser considerado como conocimiento *científico*.)

Esta división de actitudes ante lo que nos es posible conocer posiblemente esté en la base de la distinción epistemológica primordial que existe entre una epistemología o concepción de la ciencia *realista*, y una epistemología o concepción de la ciencia *antirrealista*. *Para los realistas la ciencia ha de proporcionarnos una imagen del mundo tal y como es, más allá de cómo lo percibimos. Para los antirrealistas, en cambio, éste no es el objetivo de la ciencia. En su lugar sitúan otros fines de pretensiones más modestas como, por ejemplo, establecer relaciones útiles entre observables.* Desde la concepción antirrealista, una teoría científica no pretende proporcionarnos una descripción —en algún sentido— de la realidad, sino una “red conceptual” que nos permite “recoger o capturar fenómenos” relacionándolos entre sí a través de su subsunción en leyes generales<sup>18</sup>.

La distinción entre realismo y antirrealismo es seguramente la más recurrente cuando se trata de realizar una clasificación epistemológica. Es una distinción que se apoya ante todo en la diferente concepción acerca de *los objetivos de la ciencia* que tienen las distintas epistemologías. Pero puede irse un poco más allá a la hora de distinguir entre distintas posiciones epistemológicas que implican sendas concepciones de la ciencia. Para ello resulta útil examinar las principales discusiones al calor de las cuales se ha ido fraguando la oposición entre realistas y antirrealistas. En ellas pueden destacarse dos debates fundamentales; a saber: el que se ha dado en torno al uso de los términos teóricos (sin un referente inmediatamente observable) en las teorías científicas, y el que se refiere a la concepción de la explicación. El examen de estas dos discusiones básicas puede permitirnos elaborar una clasificación epistemológica un poco más completa, articulada en dos ejes que se corresponden con cada uno de estos debates.

18. En la literatura anglosajona se suele hacer referencia a esta idea con la expresión *scientific theories save the phenomena*, que supone un *leitmotiv* en torno al cual se pueden reunir las posturas antirrealistas. La oposición entre este objetivo y el ideal explicativo se pone de manifiesto, por ejemplo, en un clásico de la filosofía de la ciencia como es el físico y filósofo francés Pierre Duhem (1906: 19) cuando afirma que una teoría científica «no es una explicación. Es un sistema de proposiciones matemáticas deducido de un número reducido de principios, que tiene por objeto representar un conjunto de leyes experimentales de la forma más simple, completa y exacta posible».

a) *Eje 1: El debate sobre los términos teóricos*

Para algunos científicos, tanto en las ciencias naturales como en las sociales, la ciencia ha de intentar desembarazarse de cualquier referente no observable en sus explicaciones. Es decir, ha de evitar el uso de términos teóricos en sus explicaciones y teorías. Por *término teórico* se entiende en la filosofía de la ciencia aquel que no tiene un referente inmediatamente observable (al menos no en un momento dado del desarrollo del conocimiento científico, ni siquiera con los aparatos de que disponemos para aumentar la potencia de nuestros sentidos) y, por tanto, cuya definición no puede basarse exclusiva y directamente en la referencia a un hecho o suceso observado, sino sólo indirectamente. Como ejemplos de términos que han sido o son teóricos, en las ciencias naturales pueden citarse los de “átomo”, “electrón”, “quark”, “supercuerda”, “virus”, “gen”, etc.<sup>19</sup>. En las ciencias sociales también encontramos este tipo de términos, por ejemplo, al hablar de “curvas de utilidad”, “actitud”, “disonancia cognitiva”, “sentimiento de comunidad”, “capital social”, etc.<sup>20</sup>.

Los que critican el uso de estos términos afirman que, precisamente, una de las notas más características que distinguen al conocimiento científico de cualquier otra forma de conocimiento es su dependencia de juicios estrictamente empíricos. En las ciencias naturales, especialmente en la física, esta epistemología cobró importancia en el siglo xx (aunque dispone de múltiples precedentes) a raíz de la expansión de las ideas del llamado Círculo de Viena. Las ciencias sociales no fueron ajenas a su impacto, como lo demuestra la relevancia que durante décadas tuvieron los planteamientos conductistas en ellas.

En la filosofía de la ciencia de las últimas décadas se ha producido un intenso y fértil debate en torno a *la interpretación de las teorías científicas*, muy especialmente en el ámbito de la física a raíz de las teorías cuántica y de la relatividad, que ha enfrentado a realistas y antirrealistas y ha permitido perfilar ambas posturas. *Este debate se ha centrado en la cuestión de cómo interpretar los últimos avances teóricos que nos proporciona la física en los que partes muy esenciales de la teoría son inobservables*. Este hecho, que resulta una característica primordial de las dos teorías que conforman la visión contemporánea del universo —la teoría cuántica y la teoría de la relatividad—, ha originado un encendido debate en la filosofía de la ciencia contemporánea entre *aquellos para los que las teorías nos proporcionan una imagen o descripción de cómo es el mundo y*

19. Se empezó a hablar de virus, gen o átomo antes de que pudiesen observarse. (El hecho de que su condición de inobservables haya sido temporal es, precisamente, uno de los argumentos esgrimidos por aquellos que defienden su inclusión en las teorías.)

20. El problema de los términos teóricos puede relacionarse en las ciencias sociales con el de las *categorías naturales*, puesto que en estas ciencias con ellos se hace referencia más a mecanismos o propiedades que a entidades. Véase sobre el problema de las categorías naturales en ciencias sociales a Rosenberg (1988).

*sus términos teóricos hacen referencia a entidades o procesos que se dan en el mundo (los realistas) y aquellos para los que la descripción de la realidad no es el objetivo de la ciencia y los términos teóricos, en todo caso, han de interpretarse como metáforas útiles —creaciones conceptuales— sin contenido ontológico (antirrealistas).*

Este tipo de discusiones no es —aunque pueda parecerlo— extraño a las ciencias sociales. Sin ir más lejos, y en relación al tema que nos interesa, se pueden leer comentarios como el siguiente. Bates *et al.* (2000: 699) responden a las críticas de Elster (2000), que les achaca tener una concepción muy estrecha de la teoría de la elección racional, afirmando que Elster falla el blanco porque:

«Utilizamos modelos para explicar, no para describir: nuestros modelos, por tanto, no necesitan incorporar cada una de las características de la vida humana».

Pero, ¿qué es explicar?, ¿no es acaso describir en un cierto sentido?

b) *Eje 2: El debate sobre la concepción de la explicación y la causalidad*

Éste es otro debate clave de la filosofía de la ciencia. Para algunos científicos, una explicación científica comparte con una predicción un mismo esquema lógico. En realidad, podría decirse que explicación y predicción son una misma cosa con consecuencias distintas. Una hace referencia a algo que ha de ocurrir; la otra, a algo que ya ha ocurrido. Pero es lo mismo. Suelen tener una definición de la explicación que se califica de *nomológico-deductiva*. De acuerdo con ella, explicar un hecho, o un fenómeno, consiste en subsumirlo en una estructura deductiva formada por una ley natural y unas condiciones iniciales de las cuales sólo se puede deducir el hecho en cuestión. Por otra parte, la idea de causalidad a la que suelen hacer referencia —y que se refleja en la forma que adquiere esta ley natural— resulta completamente empirista. Estos autores utilizan una definición de ley natural basada en una concepción humana de causa, entendida como sucesión regular de acontecimientos sin ningún contenido ontológico. Así, una ley no nos dice más que “siempre que se da A ocurre B”, sin que ello nos permita entrar en consideraciones sobre “la estructura real del mundo” (a la cual, dado su carácter inobservable en último término, no hace falta hacer referencia). Por ello, a los que defienden esta visión se les puede calificar de antirrealistas.

Oponiéndose a ellos, otros autores afirman que explicación y predicción no resultan estructuras simétricas, que se pueden dar genuinas explicaciones científicas que no impliquen predicciones, y que explicar es, ante todo, describir mecanismos causales realmente existentes. Es decir, su concepción de causa también es muy distinta. No la entienden como una sucesión de acontecimientos observados, sino como un poder generativo que puede darse incluso si no hay regularidad (si se da en un acontecimiento único, o en

una configuración de sucesos que no permite la identificación de leyes). A estos autores se los puede calificar como realistas.

Si cruzamos las distintas posiciones que se dan en estos dos debates nos aparecen tres posturas epistemológicas básicas (en realidad, cuatro, pero una de ellas puede rechazarse por su irrelevancia en la historia de la filosofía de la ciencia) que suponen sendas concepciones de la ciencia, tal y como se expresa en el gráfico siguiente:

LAS "TRES ESQUINAS" DEL "CAMPO EPISTEMOLÓGICO"

	<i>Aceptan términos teóricos</i>	<i>No aceptan términos teóricos</i>
<i>Simetría explicación/predicción</i>	<b>Instrumentalismo</b>	<b>Positivismo</b>
<i>No simetría explicación/predicción</i>	<b>Realismo científico</b>	(Realismo ingenuo)

De este modo, aunque pueden darse muchos criterios de validez distintos con los que juzgar el conocimiento científico, es posible defender que éstos pueden agruparse en tres grandes grupos o epistemologías que conforman tres maneras distintas de entender el papel de la ciencia y de la explicación científica. A estas tres posturas epistemológicas se las puede calificar como *realismo científico*, *instrumentalismo* y *positivismo* siguiendo la tradición principal en la que se mueven sus respectivos defensores, aunque la definición que hago de ellas intenta ser lo más analítica posible. Es decir, no intenta basarse sobre anteriores definiciones históricas —tan diversas a fin de cuentas como autores hay en ellas—, sino sobre el análisis del núcleo común más coherente de cada una de estas tres posturas que permite realizar una clasificación de concepciones de la ciencia y los criterios de validez asociados a ellas lo más completa y disyunta posible. Así, un *positivista* puede definirse como un científico que no acepta el uso de términos teóricos y utiliza una concepción de la explicación nomológico-deductiva. Un *instrumentalista* utiliza esta misma definición de explicación, pero no pone reparos a introducir términos teóricos en sus teorías y explicaciones siempre y cuando éstos resulten útiles para realizar predicciones<sup>21</sup>. Finalmente, un *realista* considera que toda explicación es una descripción que no necesariamente ha de ir vinculada a una regularidad legaliforme y en ella tienen cabida los términos teóricos en tanto que creamos que describen aquello que no podemos observar.

21. Aunque Popper se autocalificaba como realista ("crítico"), su metodología falsacionista puede encuadrarse en una concepción de la ciencia instrumentalista (simetría explicación/predicción, aceptación de términos teóricos).

Conviene darse cuenta de lo irreconciliable e inconmensurable de estas posturas con concepciones de la ciencia radicalmente opuestas. Las prioridades y criterios de validez que surgirán de cada una de ellas para juzgar los resultados científicos serán, igualmente, distintos y, más aún, igualmente, un tanto irreconciliables. Se fundan sobre supuestos y objetivos diferentes; entienden de manera diferente lo que es una buena explicación científica. Por ello no es de extrañar que la mayoría de las discusiones metodológicas suscitadas por los resultados conseguidos con el enfoque racional-económico estén condenadas a ser un “diálogo de sordos”. Lo están porque en ellas defensores y críticos entrecruzan argumentos que parten de perspectivas epistemológicas opuestas sin que parezcan estar muy al corriente de ello. Buscan cosas distintas y, por tanto, aceptarán como válidas cosas distintas. Una segunda conclusión que puede extraerse es que, para juzgar los argumentos de defensa del enfoque racional-económico, o de cualquier producto científico, es estrictamente necesario posicionarse claramente en una postura epistemológica determinada. De lo contrario, es fácil caer en mezclas incoherentes de argumentos generados bajo concepciones de la ciencia distintas.

## V. CRITERIOS DE VALIDEZ Y REALISMO CIENTÍFICO

¿Qué se va a exigir a una teoría o una explicación desde una concepción realista? Ante todo que se presente como —que pretenda ser— una imagen de la estructura causal del mundo (en la acepción realista del término “causa”), y que se acompañe de argumentos para creer<sup>22</sup> que así es. En otras palabras, la descripción del mundo que nos proporcionan sus supuestos y premisas será tan importante como el grado de encaje empírico que consigan sus conclusiones. Requiere que todos sus contenidos nos resulten verosímiles o plausibles, aunque pueda aceptarse perfectamente su carácter simplificador (que lo que la teoría señala como relevante para la explicación sólo sea una porción muy estrecha de la realidad). No es suficiente con que las consecuencias empíricas que puedan derivarse de la teoría se correspondan con la realidad; la teoría misma tiene una referencia ontológica que está en juego cuando juzgamos la validez de sus propuestas. Todo esto le suena a chino a un antirrealista. Aplicando esto a las explicaciones que nos ofrece el enfoque racional-económico, el realista requiere, por ejemplo, que cuando hablamos de decisiones o preferencias esto no se considere una metáfora —una simple fórmula para un cálculo teórico—, sino que nos estemos refiriendo a unas decisiones y unas preferencias *reales* que han realizado los individuos, quizás no utilizando el mismo lenguaje que el propuesto por la teoría pero con los que, definitivamente, se pueda establecer una correspondencia.

---

22. La ciencia es, en último término, siempre hipotética, pero nos da argumentos racionales para defender sus postulados.

*Ante la búsqueda de verosimilitud (de correspondencia entre esta imagen global que ofrece la teoría y el mundo), para el realista todo el resto queda en un segundo plano.* En la filosofía de la ciencia contemporánea clásica —lo que en el mundo anglosajón se ha venido a llamar *the received view*— se suelen citar una serie de características que deberían tener las teorías científicas y que, por tanto, pueden servir tanto para orientar la tarea investigadora de los científicos como para constituir criterios con los que juzgar y elegir entre teorías rivales (Kuhn, 1977: 322; Kincaid, 1996: 50). Estos criterios son:

- *Simplicidad*: cuanto más simple sea una teoría, mejor.
- *Precisión empírica*: cuanto mejor encaje con los datos de los que disponemos, mejor.
- *Éxito predictivo*: cuanto mejores sean sus predicciones en cantidad y calidad, mejor.
- *Coherencia interna (o consistencia lógica)*: una teoría no puede presentar contradicciones internas.
- *Coherencia externa (o unidad)*: las informaciones que nos proporciona una teoría han de ser coherentes con las que nos proporcionan otras teorías científicas; cuantas más interconexiones se puedan establecer entre teorías diversas, mejor.
- *Universalidad (o amplio alcance)*: cuanto mayor sea el alcance de una teoría, cuanto más amplio sea el ámbito de aplicación de la teoría, cuanto mayor sea la variedad de fenómenos de los que puede dar cuenta, mejor.
- *Fecundidad*: cuanto mayor sea la capacidad de la teoría para generar nuevas teorías subsidiarias o elaboraciones conceptuales diversas, mejor.

A veces también se añaden a esta lista otros criterios que, en algunos casos, no dejan de ser reformulaciones de los anteriores. Así, por ejemplo, se citan: *poder unificador*, *falta de formulaciones ad hoc*, *capacidad de generar predicciones novedosas*. Como sustituto de los criterios de simplicidad y universalidad se suele invocar otro criterio que tiene una presencia destacada en numerosos manuales de metodología. Se trata del criterio de *parsimonia*: explicar el máximo con el mínimo.

Esta lista de criterios constituye un clásico de la metodología. Y uno diría que *a priori* suenan todos ellos como bastante razonables. Y seguramente lo son. Lo que pasa es que no todos son igualmente indicados para conseguir determinados objetivos, es decir, para aproximarnos al ideal de la ciencia que una epistemología particular establece. Aunque muchas veces se han defendido y propuesto como valores en sí mismos, *no hay que olvidar su conexión con las diversas perspectivas epistemológicas*.

¿Hasta qué punto estos criterios encajan con una visión realista de la ciencia? Desde una concepción realista, estos criterios han de estar completamente supeditados al objetivo de conseguir imágenes que se correspondan cada vez mejor con lo que ocurre en

el mundo. En la medida en que esto no sea así, su valor como criterios de validez queda cuanto menos limitado o en un segundo plano. Éste es, sin duda, un problema que no atañe a los antirrealistas, para los cuales estos criterios tienen un valor epistemológico por sí mismos —no subsidiario— pues garantizan una mayor calidad conceptual de la teoría (independientemente de su contenido ontológico). *Para un antirrealista, y de acuerdo con su concepción de los objetivos de la ciencia, cuanto más se ajuste una explicación a estos criterios, más correcta, o válida, va a resultar.*

*Para un realista, en cambio, es estrictamente necesario plantearse la conexión que existe entre estos criterios y sus objetivos explicativos. En qué medida la búsqueda de estos ideales nos garantiza una mayor correspondencia entre realidad y explicación (que se entiende como una descripción peculiar —interesada primordialmente en la estructura causal del mundo— de la misma). Aparecen como primordiales cuestiones tales como ¿una teoría o explicación más simple, o parsimoniosa, tiene más posibilidades de ser plausible que una que no lo es? Si la respuesta es que sí, la simplicidad puede resultar un criterio útil para un realista; si la respuesta es que no, su relevancia en la elección entre teorías o explicaciones no puede sino pasar a un segundo plano.*

Permítaseme un ejemplo para aclarar este punto. Es como si nos dijese que las manzanas mejor que sean rojas, brillantes, sin manchas, con un pequeño rabo, simétricas respecto de un eje vertical. Sin duda, todos estos criterios nos pueden ayudar a elegir una manzana entre todas las de la caja. En vez de hacerlo al azar aplicamos estos criterios y podemos elegir a una manzana entre las demás. Pero si lo que buscamos es una manzana sabrosa, en la medida en que no tengamos una forma de establecer una relación constante entre estas características y el sabor de la manzana, estos criterios que pueden ayudarnos a distinguir entre manzanas no nos sirven de cara a nuestro objetivo. Incluso podría darse el caso de que nos desviasen de él, si resultase que las manzanas más sabrosas son, en realidad, las más deformes, manchadas y las menos brillantes. Los criterios antes mencionados, sin duda nos pueden servir para elegir entre manzanas. Lo que no tiene sentido es que apelemos a ellos para defender que la manzana elegida es la más sabrosa mientras no dispongamos de razones para relacionar estas características con la que a nosotros nos parece principal.

Esto es lo que, a menudo, ocurre en los debates científicos. Se apela a una serie de criterios o rasgos para elegir entre teorías, o para defenderlas o atacarlas. En principio, esas características suenan razonables. Y se han utilizado indistintamente, sin vincularlas a ninguna epistemología en concreto, como si fueran buenas (epistemológicamente hablando) por sí mismas. Y, sin duda, es cierto que nos pueden permitir distinguir entre teorías, pero eso no nos sirve de nada si esa distinción resulta incoherente con la concepción de la ciencia que uno defiende.

Para un antirrealista, todos los criterios mencionados resultan útiles; si es un positivista habrá que añadir su exigencia a evitar referencias a inobservables. Para un realista, su utilidad no está tan clara. En realidad, como ya he afirmado, no deja de ser subsidiaria.

De todos los criterios mencionados, sólo dos se destacan por tener una relación clara con los objetivos que se plantea el realista. Por un lado, como en cualquier otra epistemología, éste requerirá la *adecuación empírica* de la teoría: ésta debería encajar lo mejor posible con los datos de los que disponemos (y, en el mejor de los casos, debería ser capaz de producir consecuencias empíricas directamente contrastables). Pero, igualmente, para él se hará necesaria lo que podríamos denominar *adecuación metafísica* (que abarca también lo no estrictamente observable o lo que se da por supuesto), que se produce cuando existe una coherencia entre el contenido de la teoría y la imagen que nos proporciona del mundo y la imagen ontológica general de la que disponemos a raíz del conjunto de conocimientos que nos ofrecen las distintas ciencias.

Esta doble adecuación es lógica y resulta totalmente justificada en una epistemología realista. En el realismo —a diferencia del positivismo— no se pone ninguna traba a construir teorías que incluyen referencias a inobservables que son, por tanto, difíciles de corroborar empíricamente. En el criterio de demarcación del realismo, la observabilidad de los términos incluidos en una teoría no es una exigencia. En consecuencia, la adecuación empírica resulta pobre en sí misma y, sin duda, requiere ir acompañada de la adecuación metafísica u ontológica. Los dos criterios deben cruzar sus puntos de mira sobre la teoría en juicio. Por otra parte, y en esa misma dirección, hay que subrayar que en el realismo —a diferencia del instrumentalismo— el objetivo que se persigue con una teoría (explicativa) no es nunca puramente instrumental, sino que se pretende dotarnos de una imagen ontológica (de lo real). Todos sus términos refieren a entidades o mecanismos existentes, exigencia que no encontramos en el criterio de demarcación falsacionista generalmente adoptado por los instrumentalistas. Ambas cuestiones, fundamentales en la definición de una postura realista, hacen necesaria la búsqueda de *unidad* entre todas las teorías científicas; se podría decir que puesto que la realidad es única (no por cuestiones pragmáticas).

Esta doble exigencia, adecuación empírica pero también metafísica, presenta sus problemas (a los que son ajenos otras epistemologías). Por ejemplo, a la hora de plantearse qué puede garantizarnos mayor plausibilidad. ¿Qué es mejor, una teoría cuyas predicciones resultan hasta un cierto punto exitosas pero que presenta unas premisas que entran en contradicción con lo que otras ciencias o teorías aceptadas afirman, u otra que encaja con nuestra imagen global del mundo pero que no nos permite elaborar predicciones novedosas?

La capacidad predictiva de una teoría no es, para el realista, el valor epistemológico supremo, sino uno más de los indicadores que tenemos para juzgar su plausibilidad. Una teoría o un modelo pueden resultar útiles y proporcionarnos buenas predicciones y, sin embargo, no ser explicativos de acuerdo con la concepción de la explicación que defiende el realista, que la entiende como una descripción de mecanismos causales. Y, por otra parte, hay buenas explicaciones para un realista que no permiten llevar

a cabo predicciones (naturalmente, un antirrealista se mostraría en desacuerdo con esta afirmación).

## V. DESDE LA PERSPECTIVA DE UN REALISTA CIENTÍFICO: EL ACUSADO ES INOCENTE, EL ABOGADO CULPABLE

Posiblemente el realismo pueda calificarse como “la epistemología más popular”, pues es la que la mayoría de la gente sostiene, a menos que uno sea un filósofo o un científico aficionado a la filosofía. Nos resulta la más natural. Todos (a excepción de esos filósofos y científicos) creemos que la realidad está allí afuera, que es independiente de nosotros y que podemos conocerla, y cuando eso hacemos somos capaces de realizar afirmaciones verdaderas acerca de su naturaleza. Éstas son, más o menos, las premisas básicas sobre las que se asienta el realismo científico para el que cualquier teoría científica intenta ante todo describir la realidad. Pero, precisamente porque nos resulta tan natural, el realismo, en tanto que concepción de la ciencia, tradicionalmente se ha visto desprovisto del entramado argumental con el que se presentan y justifican sus competidoras epistemológicas. Las posturas antirrealistas (positivistas e instrumentalistas), al resultar menos intuitivas, necesariamente se han presentado en el marco de programas sistemáticos de definición y defensa perfectamente estructurados en los que el supuesto “sentido común” del realismo es contrarrestado por profundas argumentaciones filosóficas. De todas maneras, lo cierto es que el realismo también resulta la epistemología natural para muchos científicos, ya sea con o sin el soporte de un programa filosófico detrás.

Sin embargo, por lo que respecta a las ciencias sociales, carecer de ese apoyo ha restado credibilidad y prestigio académico a determinados posicionamientos metodológicos y a los productos derivados de ellos. En algunas ciencias, más que en otras, el adoptar programas epistemológicos antirrealistas ha resultado durante muchos años un signo de calidad que pretendía —inadecuadamente— situar a las ciencias sociales al mismo nivel de prestigio que las ciencias naturales, alejándolas de la reflexión social del ciudadano común y dotándolas, al mismo tiempo, de “cientificidad”. En la economía, por ejemplo, que suele presentarse como el paradigma a seguir por el resto de ciencias sociales, las argumentaciones de carácter positivista o instrumentalista, entre las que hay que destacar por su amplia presencia la metodología falsacionista popperiana (Blaug, 1992; Hands, 1993; Backhouse, ed., 1994), acompañan frecuentemente las presentaciones de teorías, modelos y explicaciones. Esto mismo ocurre con la teoría de la elección racional.

Si observamos los criterios de defensa que más arriba hemos ido repasando, nos damos cuenta de que resultan completamente coherentes con una concepción de la ciencia antirrealista, pero no pueden resultar convincentes para un realista. Recordé-

moslos: los supuestos no importan; lo más importante es que se puedan realizar predicciones y que éstas encajen con lo observado; más aún, es suficiente y necesario que la predicción sea exitosa para poder aceptar como válida una explicación; no hay que hacer referencia a aquello que no se puede observar; hay que conseguir modelos/explicaciones elegantes, generales (de gran alcance) y sencillas; si es posible las teorías han de ser de ámbito universal; la parsimonia de la teoría es un objetivo metodológico prioritario.

Junto a estos criterios se encuentran dos que merecen algún comentario adicional. En primer lugar, el que hace referencia a que hay que evitar cualquier movimiento *ad hoc*. Naturalmente, un movimiento *ad hoc* no presenta ninguna virtud en sí mismo, tampoco para un realista. Pero su relevancia y necesidad de evitarlos es menor si la validez de una teoría o una explicación no se hacen depender exclusivamente de su capacidad predictiva y su posibilidad de falsación respectivamente. Cuando esto no es así, como ocurre en una metodología realista, el diálogo entre empírea y teoría se puede establecer de manera menos estricta, estableciéndose vías muy distintas de dotar de plausibilidad y validez a una hipótesis que se escapen del examen de sus consecuencias empíricas.

En segundo lugar, recuperemos el argumento lakatosiano: si no hay una teoría mejor, entonces la que mejor funciona, pese a sus limitaciones, es válida. En relación a este argumento hay que decir que, aunque la plausibilidad de una teoría puede depender de lo que se afirma en otras teorías, en el caso de que ésta resulte implausible (por las razones que sean), este juicio no puede quedar suspendido por el hecho de que no haya ninguna teoría alternativa disponible. Para ilustrar este punto, nada mejor que el comentario, en tono jocoso, de Green y Shapiro (1996: 258: este ejemplo se lo deben a Brian Barry):

«Antes del descubrimiento de la clorofila, los científicos no sabían por qué la hierba es verde, pero las mentes científicas podían tener, sin embargo, buenas razones para rechazar una teoría que diese cuenta de ello basándose en que la luna está hecha de queso verde».

*El hecho de que no hubiese una teoría alternativa dispuesta a dar cuenta de este hecho no reducía la implausibilidad de la misma. Y siguen con un ejemplo menos anodino (Green y Shapiro, 1996: 258):*

«De igual modo, si alguien avanza una teoría que busca explicar los resultados relativamente pobres de los negros en ciertos campos por referencia a su alegada inferioridad genética, uno no tiene por qué tener en su posesión una teoría alternativa con la que dar cuenta de esos pobres resultados para poder concluir que no hay ni una pizca de evidencia creíble con la que apoyar la visión de la inferioridad genética. Naturalmente uno ha de saber *algo* para rechazar estas teorías como implausibles, pero no tiene que ser mucho,

y ciertamente uno no tiene por qué tener una explicación que dé cuenta del fenómeno en cuestión».

La carga de la prueba está en quien presenta la teoría. En este caso, de nada sirve el que no exista una teoría alternativa en las ciencias sociales con las mismas pretensiones “para derrotar a la teoría de la elección racional”. Ésta tiene que ser plausible por méritos propios, y de ellos ha de depender su validez, y no del hecho de que “no hay nada mejor”.

Examinados los argumentos que se suelen invocar para afirmar la validez de la teoría de la elección racional y de las explicaciones que se derivan de ella puede afirmarse rotundamente que, desde una perspectiva realista, éstos resultan inadecuados. Ante ellos, un realista no puede sino quedarse indiferente. “¿Que su teoría es parsimoniosa?”, “¡y qué!” —exclamará el realista—, “¡dígame si es verosímil!”. El comentario podría reformularse de forma más precisa: “Dígame si su aducida parsimonia puede ser una prueba de la plausibilidad del mecanismo causal apuntado”, que es lo que, en definitiva, le interesa a un realista. Sólo en la medida en que sea posible elaborar este tipo de conexiones y demostrar que las virtudes a las que apelan los defensores del enfoque racional-económico pueden resultar coherentes con los objetivos realistas sus argumentos resultarán de interés para un realista. Y estas conexiones no son tan fáciles de hacer.

Ahora bien, una cosa es que los argumentos habituales utilizados para su defensa no resulten convincentes para un realista, y otra cosa distinta es que un realista no pueda defender con otros argumentos un buen número de los resultados conseguidos con la aplicación del enfoque racional-económico. *Quizás los abogados de la aproximación económico-racional no hayan sido los mejores, al menos no para un realista, pero hay numerosos resultados conseguidos con la aplicación de este enfoque teórico que un realista, con sus propios argumentos y por otras razones distintas a las habituales, puede considerar totalmente válidos.* Probablemente para llevar a cabo esta tarea sea necesario reinterpretar parte de los éxitos teóricos y empíricos conseguidos con la teoría, reelaborando el argumento explicativo que se suele asociar a los mismos (es decir, reinterpretando los éxitos empíricos conseguidos y reajustando la explicación como respuesta plausible a una pregunta muy concreta y, en ocasiones, alejada de un supuesto interés por la acción de los individuos). No es ésta una tarea nueva en la ciencia (o exclusiva de las ciencias sociales). No hay más que recordar, en este sentido, por ejemplo, la famosa polémica interpretativa que sobre los resultados conseguidos con la aplicación de la teoría cuántica iniciaron Bohr y Einstein, y que aún hoy colea.

Desde el realismo científico puede defenderse también la teoría de la elección racional y algunos de sus resultados; lo único que he constatado aquí es lo nefastos que pueden resultar sus abogados para aquellos que defienden una concepción de la ciencia realista. Lo que puede explicar que muchos científicos sociales que de una manera implícita se adhieren a esta posición epistemológica se hayan alejado de ella. Un análisis

más atento, sin embargo —que dejo para otra ocasión—, nos tiene que servir para descubrir por qué esto es un error y por qué hay razones para que algunos éxitos de la teoría de la elección racional puedan ser compartidos y defendidos también desde una perspectiva realista.

### Referencias

- Abelson, Robert P. 1995. «The Secret Existence of Expressive Behavior», en Jeffrey Friedman (ed.). 1996.
- Alt, James E., y Kenneth A. Shepsle (eds.). 1990. *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press.
- Backhouse, Roger E. (ed.). 1994. *New Directions in Economic Methodology*. Londres: Routledge.
- Barnett, William A., Melvin J. Hinich y Norman J. Schofield (eds.). 1993. *Political Economy: Institutions, Competition, and Representation*. Cambridge: Cambridge University Press.
- Barry, Brian M. 1970. *Sociólogos, economistas y democracia*. Buenos Aires: Amorrurtu.
- Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal y Barry R. Weingast. 2000. «The Analytic Narrative Project», *American Political Science Review*, 94, 3 (september).
- Becker, Gary. 1976. «El Enfoque económico del comportamiento humano» (primer capítulo de su libro *The Economic Approach to Human Behavior*). Traducido en *Información Comercial Española*, 557, enero 1980, y en Febrero y Schwartz, 1995.
- Becker, Gary, y George J. Stigler. 1977. «De Gustibus Non Est Disputandum», *American Economic Review*, 67: 76-90. Trad. en Febrero y Schwartz, 1995.
- Bhargava, Rajev. 1992. *Individualism in Social Science. Forms and limits of a methodology*. Oxford: Clarendon.
- Blaug, Mark. 1992. *The Methodology of Economics: How Economists Explain*. Cambridge: Cambridge University Press.
- Cohn, Jonathan. 1999. «When did Political Science Forget about Politics?», *The New Republic*, october.
- Cox, Gary W. 1999. «The Empirical Content of Rational Choice Theory. A Reply to Green and Shapiro», *Journal of Theoretical Politics*, 11 (2): 147-169.
- Crain, W. Mark, y Robert D. Tollison (eds.). 1990. *Predicting Politics. Essays in Empirical Public Choice*. Ann Arbor: The University of Michigan Press.
- Crotty, William (ed.). 1991. *Political Science: Looking to the Future*. Evanston (Ill.): Northwestern University Press.
- Duhem, Pierre. 1906. *The Aim and Structure of Physical Theory*, traducción de la segunda edición (1914) en Princeton University Press, 1954.

- Dunleavy, Patrick. 1991. *Democracy, Bureaucracy & Public Choice. Economic Explanations in Political Science*. Hemel Hempstead: Harvester-Wheatsheaf.
- Elster, Jon. 1986. «Introduction», en Jon Elster (ed.). 1986.
- Elster, Jon (ed.). 1986. *Rational Choice*. Oxford: Basil Blackwell.
- Elster, Jon. 2000. «Rational Choice History: A Case of Excessive Ambition», *American Political Science Review*, 94, 3 (september).
- Enelow, James. 1981. «Saving Amendments, Killer Amendments, and an Expected Utility Theory of Sophisticated Voting», *Journal of Politics*, 43.
- Friedman, Jeffrey (ed.). 1996. *The Rational Choice Controversy. Economic Models of Politics Reconsidered*. New Haven: Yale University Press.
- Goetze, David, y Peter Galderisi. 1989. «Explaining Collective Action with Rational Models», *Public Choice*, 62: 25-39.
- Green, Donald P., e Ian Shapiro. 1994. *Pathologies of Rational Choice Theory: a Critique of Applications in Political Science*. New Haven: Yale University Press.
- Green, Donald P., e Ian Shapiro. 1996. «Pathologies Revisited: Reflections on Our Critics», en Friedman (ed.). 1996.
- Hands, D. Wade. 1993. «Popper and Lakatos in Economic Methodology», en Mäki, Gustafsson y Knudsen (eds.). 1993.
- Heap, S. H., M. Hollis, B. Lyons, R. Sugden y A. Weale. 1992. *The Theory of Choice. A Critical Guide*. Oxford: Blackwell.
- Heiner, Ronald A. 1983. «The Origin of Predictable Behavior», *American Economic Review*, 4: 560-95.
- Heiner, Ronald A. 1985. «Origin of Predictable Behavior: Further Modeling and Applications», *American Economic Review*, 75: 391-402.
- Heiner, Ronald A. 1989. «Imperfect Choice and Self-Stabilizing Rules», *Economics and Philosophy*, 5: 19-32.
- Hindess, Barry. 1988. *Choice, Rationality and Social Theory*. Londres: Unwin Hyman.
- Hogarth, Robin y Melvin Reder (eds.). 1987. *Rational Choice. The Contrast Between Economics and Psychology*. Chicago: The Chicago University Press.
- Jackman, Robert W. 1993. «Rationality and Political Participation», *American Journal of Political Science*, 37, 1: 279-290.
- Kahneman, Daniel, y Amos, Tversky. 1979. «Prospect Theory: An Analysis of Decision Under Risk», *Econometrica*, 47: 263-291.
- Kahneman, Daniel, y Amos, Tversky. 1984. «Choices, Values and Frames», *American Psychologist*, 39: 341-350.
- Kahneman, D., P. Slovic y A. Tversky (eds.). 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kincaid, Harold. 1996. *Philosophical Foundations of the Social Sciences. Analyzing Controversies in Social Research*. Cambridge: CUP.

- Kuhn, T. S. 1977. «Objectivity, Value Judgement and Theory Choice», en *The Essential Tension*. Chicago: The University of Chicago Press: 320-351.
- Lakatos, Imre. 1970. «Falsification and the Methodology of Scientific Research Programmes», en Lakatos y Musgrave (eds.). 1970.
- Lakatos, Imre, y Alan Musgrave (eds.). 1970. *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.
- Lewin, Leif. 1991. *Self-Interest and Public Interest in Western Politics*. Oxford: Oxford University Press.
- Mäki, Uskali, Bo Gustafsson y Christian Knudsen (eds.). 1993. *Rationality, Institutions and Economic Methodology*. Londres: Routledge.
- March, James G. 1986. «Bounded Rationality, Ambiguity, and the Engineering of Choice», en Elster (ed.). 1986.
- McLean, Iain. 1987. *Public Choice. An Introduction*. Oxford: Basil Blackwell.
- Monroe, Kristen R. 1991. «The Theory of Rational Action: What Is It? How Useful Is It for Political Science?», en Crotty (ed.). 1991.
- Morrow, J. D. 1994. *Game Theory for Political Scientists*. Princeton: Princeton University Press.
- Olson, Mancur. 1965. *The Logic of Collective Action. Public Goods and the Theory of Groups*. Cambridge (Mass.): Harvard University Press, segunda edición (con apéndice) de 1971. (Hay traducción en castellano: *La lógica de la acción colectiva*. México: Limusa-Noriega, 1992.)
- Olson, Mancur. 1982. *Auge y decadencia de las naciones*. Barcelona: Ariel, 1986.
- Opp, Karl-Dieter. 1999. «Contending Conceptions of the Theory of Rational Action», *Journal of Theoretical Politics*, 11 (2): 171-202.
- Ordeshook, Peter C. 1993. «The development of contemporary political theory», en Barnett *et al.* (ed.). 1993.
- Piattelli Palmerini, Massimo. 1993. *Los túneles de la mente. ¿Qué se esconde tras nuestros errores?* Barcelona: Crítica, 1995.
- Popper, Karl R. 1945. *La sociedad abierta y sus enemigos*. Barcelona: Paidós, 1994.
- Renwick Monroe, Kristen (ed.). 1991. *The Economic Approach to Politics. A Critical Reassessment of the Theory of Rational Action*. Nueva York: Harper Collins.
- Riker, William H. 1990. «Political Science and Rational Choice», en Alt y Shesple (eds.). 1990.
- Rosenberg, Alexander. 1988/1995. *Philosophy of Social Science*. Oxford: Clarendon Press. [Segunda edición en Boulder (Co): Westview, 1995].
- Self, Peter. 1993. *Government by the Market? The Politics of Public Choice*. Londres: Macmillan.
- Sen, Amartya. 1985. «Goals, Commitment, and Identity», *Journal of Law, Economics, and Organization*, 1, 2: 341-355.

- Shepsle, Kenneth. 1995. «Statistical Political Philosophy and Positive Political Theory», *Critical Review*, 9: 213-222.
- Shepsle, Kenneth A., y Mark S. Bonchek. 1997. *Analyzing Politics. Rationality, Behavior, and Institutions*. Nueva York-Londres: W. W. Norton.
- Simon, Herbert A. 1983. *Reason in Human Affairs*. Oxford: Basil Blackwell.
- Simon, Herbert A. 1985. «Human Nature in Politics: The Dialogue of Psychology with Political Science», *American Political Science Review*, 79.
- Simon, Herbert A. 1990. «Invariants of Human Behavior», *Annu. Review of Psychology*, 41: 1-19.
- Taylor, Michael. 1996. «When Rationality Fails», en J. Friedman (ed.). 1996.
- Tilly, Charles. 1995. «To Explain Political Processes», *American Journal of Sociology*, 100 (6): 1594-1610.
- Tversky, A., y Daniel Kahneman. 1986. «Rational Choice and the Framing of Decisions», en Hogarth y Reder (eds.). 1986.

JAUME LÓPEZ

E-mail: jaume.lopez@upf.edu

Es profesor ayudante en la Universitat Pompeu Fabra y profesor-consultor en la Universitat Oberta de Catalunya. Doctor en Ciencia Política por la Universitat Pompeu Fabra y Master en Filosofía de las Ciencias Sociales por la London School of Economics. Sus áreas de especialización son la teoría política, la teoría de la acción colectiva y la metodología de las ciencias sociales. Su tesis se tituló *Realismo científico y teoría de la elección racional. Cómo y por qué un realista debería defender la teoría de la elección racional*.