

DOCUMENTOS DEL CIECE

**Análisis filosófico–
epistemológico de Prospect
Theory y la Neuro-Economía**

Eduardo Scarano y Gustavo Marqués
(compiladores)

Staff

Director
Gustavo Marqués

Secretaría
Hernán Ruggeri
Laura Rosenberg

Editor Responsable

Centro de Investigación en Epistemología de las Ciencias
Económicas

Facultad de Ciencias Económicas
Universidad de Buenos Aires

Av. Córdoba 2122 1º p. Aula 111
(1120) Ciudad Autónoma de Buenos Aires
Argentina
Tel. (54-11) 4370-6152
Correo electrónico: ciece@econ.uba.ar

ISSN: 1851-0922

Índice

<u>Prólogo</u>	5
<u>El consecuencialismo de la teoría de la decisión, por Eduardo Scarano</u>	7
<u>Prospect Theory: ¿Una teoría de la decisión <i>neoclásica</i>?, por Diego Weisman</u>	27
<u>El principio epistémico de confianza en tanto condición de posibilidad de la decisión racional, por Pablo García y Sandra Maceri</u>	43
<u>El desarrollo de Behavioral Economics y su relación con la teoría de la elección racional, por Leonardo Ivarola</u>	61
<u>Un realismo empírico epistemológicamente adecuado para Behavioral Economics, por Gustavo Marqués</u>	81
<u>Tipos neuroeconómicos como tipos institucionales. Un esquema de interpretación, por Alejandro Miroli</u>	95

Prólogo

Los trabajos que se presentan en este volumen han sido realizados en el marco del proyecto UBACyT E029, “Evaluación metodológica de la teoría económica estándar de la decisión racional y sus relaciones con las disciplinas colindantes, la psicología y la neurobiología”, desarrollado en el CIECE durante 2008 – 2010. El proyecto ha examinado algunas de las principales alternativas a la Teoría de la Utilidad Esperada que es la teoría estándar de decisión racional en condiciones de riesgo e incertidumbre. Todas las alternativas referidas explotan algún aspecto de la vasta concepción conocida bajo el rótulo de Racionalidad Acotada, que supone en los agente capacidad limitada de registro y procesamiento de la información, y concibe la racionalidad como satisfacción o la adquisición de pre-compromisos. El volumen ofrece una evaluación metodológica de las nuevas perspectivas, examina su impacto en el corpus de la teoría económica tradicional, su status como ciencia separada de la psicología y de las ciencias naturales, y compara los méritos respectivos de EUT y sus alternativas en cuanto a sus capacidades para explicar y predecir.

Se ha hecho especial hincapié en Prospect Theory, expuesta originalmente por Daniel Kahneman y Amos Tversky en 1979, que es una teoría meramente descriptiva de la elección en condiciones de riesgo y tiene el mérito de eliminar muchas de las anomalías a que se enfrenta la Teoría de la Utilidad Esperada. De Prospect Theory se ha examinado su versión más “clásica”, que da origen a la mayor parte de la literatura agrupada bajo el rótulo de “Paternalismo Libertario” y aquella versión hedonística y menos ortodoxa que opone a la noción ex – ante estándar de “decision utility” el concepto de “experience utility”, que es de naturaleza ex – post. Esta última perspectiva profundiza la relación entre economía y psicología y promete interesantes aplicaciones a la teoría económica del bienestar. El proyecto ha prestado también atención a la llamada neuro-economía, que es una rama de la economía comportamental y experimental, que en su versión más audaz procura fundar la microeconomía en el funcionamiento del cerebro.

Eduardo Scarano y Gustavo Marqués.

EL CONSECUENCIALISMO DE LA TEORÍA DE LA DECISIÓN

Eduardo R. Scarano (CIECE – FCE - Universidad de Buenos Aires)

Resumen

La formulación estándar de la teoría de la decisión es de tipo consecuencialista. Es una caracterización usual asociada íntimamente a una concepción tradicional de la racionalidad enfrentada a otra perspectiva también tradicional de tipo deontológica. Sin embargo, los resultados de la contrastación de la teoría de la decisión, especialmente la experimental, han obligado a reformularla introduciendo otros factores explicativos no cognitivos que la convierten en una teoría no consecuencialista sin, por esto, volverla una teoría deontológica. El factor más importante introducido es la emoción.

Este nuevo tipo de teoría no consecuencialista tiene implicaciones de largo alcance, tanto el campo económico y de otras disciplinas como en el metodológico y en el estrictamente filosófico. El centro del debate es la racionalidad.

Introducción

La teoría de la decisión, denominada usualmente teoría de la racionalidad, constituye una teoría básica de la microeconomía y, por lo tanto, de la economía estándar. Puede considerársela una teoría de la acción pues explica por qué un agente económico realiza una acción, que generalizada adecuadamente a otros ámbitos también explica acciones no económicas.

En este trabajo no deseamos examinar la teoría de la decisión por sí misma, sino solamente una de sus características, el consecuencialismo, que implica elegir una acción por las consecuencias a las que conduce. Una teoría rival de la anterior es la teoría deontológica que selecciona las acciones por los valores de cada acción independientemente de sus consecuencias. La dificultad mayor de este último tipo de teorías es la poca claridad cuando debemos especificar cómo argumentar y cómo elegir entre valores.

Ambas se disputan las explicaciones y aparecen como teorías dicotómicas, o se es consecuencialista o se es deontologista y no ambas

cosas a la vez. Uno de los resultados extraordinarios de teorías de la decisión alternativas que presentaremos es que no se subsumen en ninguna de esas dos.

Los resultados de la contrastación de la teoría de la decisión, especialmente experimentales, condujo a distintas reformulaciones para absorber las anomalías, los contraejemplos que aparecieron. Uno de los principales factores que se introdujeron con ese fin fue la emoción y los sentimientos en el proceso de toma de decisiones. Estos factores usualmente aparecen opuestos a o perturbadores de la racionalidad. En las nuevas teorías que los introducen la primera manifestación obvia es la imposibilidad de mantener la dimensión consecuencialista de la decisión; ahora las emociones o sentimientos intervienen y se superponen a la racionalidad tradicional. Otra manifestación inmediata es la dificultad de restringir las emociones o sentimientos a una caracterización puramente cognitiva. Estas nuevas teorías cuestionan tanto el consecuencialismo como el cognitivismo.

Las implicaciones de estos debates son de largo alcance, tanto el campo económico y de otras disciplinas como en el metodológico y en el estrictamente filosófico. La racionalidad misma está en cuestión.

La estructura del trabajo es la siguiente: en la sección II. se abordan las teorías deontológicas; en III. las consecuencialistas; en IV. se examina detalladamente el consecuencialismo de la teoría de la decisión; en V. los aspectos no consecuencialistas en nuevas teorías de la decisión; finalmente en VI. las conclusiones.

Teorías deontológicas

En las teorías deontológicas se puede argumentar acerca de valores o fines, en particular, acerca de cuál es el mejor valor. Una acción no se selecciona por sus consecuencias, sino a pesar de sus consecuencias por el valor que le es inherente o al que está asociado.

La Escuela de Francfort es una conspicua representante de esta posición. Se conoce con esta denominación a un grupo de pensadores reunidos alrededor del *Institut für Socialforschung* en 1923 en Francfort que llegó a tener varias filiales en otros países. Un pequeño grupo, pero el más representativo, se nucleó alrededor de Max Horkheimer. Sus integrantes eran Pollock, Lowenthal, Adorno, Fromm y Marcuse. En 1934 el Instituto

se instaló en New York asociado a la Universidad de Columbia. Uno de los principales aportes de la escuela de Francfort es la Teoría Crítica, que tiene diferentes versiones.

Siempre rechazaron tajantemente lo absoluto bajo cualquier forma que se presentara, especialmente como la posibilidad de la verdad absoluta. Reivindicaron la dialéctica de Hegel, su idea de una lógica sustantiva, y su concepción de la razón (como *Vernunft* y no solo como *Verstand*). La Escuela llevó un combate en dos frentes, contra el irracionalismo y contra un excesivo empirismo que reducía la razón a *Verstand*, abdicando de una razón más abarcativa.

El método de la Teoría Crítica era muy diferente del tradicional, se expresaba a través de la crítica de otros pensadores y tradiciones filosóficas y no es un sistema filosófico cerrado, la Escuela de Francfort tenía especial aversión por ellos. Fue influenciada principalmente por Hegel, Marx y la Filosofía de la Vida. Sus características más notorias para nuestro fines son las siguientes:

La lógica formal es insuficiente, es necesaria una lógica sustantiva, que dependa del contenido; el método de las ciencias sociales y naturales son diferentes; no hay dualismo de hechos y valores; la razón descubre los hechos por detrás de las apariencias, y determina los valores; la investigación social contiene un componente histórico que debe considerar lo presente a la luz de las posibilidades históricas de realización.

La razón instrumental

La crítica de la razón instrumental constituyó uno de los aportes fundamentales de la Escuela de Francfort, y uno de los más influyentes. Seguiremos la exposición de Horkheimer, *Crítica a la razón instrumental* [1973]. Por razón instrumental entiende la capacidad de clasificar, deducir, calcular, sin necesidad de atender a la clase de contenido al que se aplica. En lógica es característico afirmar que la validez de un argumento depende de su forma y no de su contenido. Cuando tiene que habérselas con medios y fines, calcula los modos más adecuados de proceder para alcanzar el fin con los medios disponibles. Sin embargo, no determina si los fines son razonables o no, a lo sumo se sobreentiende que sirven a los intereses del sujeto. Esta clase de razón puede denominarse *razón subjetiva*.

La idea de un objetivo capaz de ser racional por sí mismo (...) le resulta a la razón subjetiva profundamente ajena, aún allí donde se eleva por encima de la consideración de valores inmediatamente útiles, para dedicarse a reflexiones sobre el orden social contemplado como un todo. [Horkheimer, p.16]

Durante mucho tiempo predominó otra concepción de la razón, estaba contenida no solo en la conciencia individual, sino también en el mundo objetivo. Se la puede encontrar entre otros en Platón, Aristóteles, la escolástica, el idealismo alemán,

Esta aspiraba a desarrollar un sistema vasto o una jerarquía de todo lo que es, incluido el hombre y sus fines. El grado de racionalidad de la vida de un hombre podía determinarse conforme a su armonía con esa totalidad. La estructura objetiva de esta –y no solo el hombre y sus fines- debía servir de pauta para los pensamientos y las acciones individuales. Tal concepto de la razón no excluía jamás a la razón subjetiva sino que la consideraba una expresión limitada y parcial de una racionalidad abarcadora, vasta, de la cual se deducían criterios aplicables a todas las cosas y a todos los seres vivos. El énfasis recaía más en los fines que en los medios. [Horkheimer, p.16]

La razón subjetiva es una capacidad del sujeto, y en esta concepción *racional* solo se predica con propiedad del individuo. La razón objetiva es un principio inherente a la realidad, incluido el hombre, y este mediante su intelecto, su razón subjetiva, es capaz de descubrir esa racionalidad incluida los fines. Mientras que en la razón subjetiva la razón es un acto, en la objetiva es sobre todo una cosa, un pensamiento y la relación que esa cosa o pensamiento guarda con un fin. Para la razón subjetiva,

No existe ninguna meta racional en sí, y no tiene sentido entonces discutir la superioridad de una meta frente a otras con referencia a la razón. Desde el punto de vista subjetivo, semejante discusión es posible cuando ambas metas se ven puestas al servicio de otra tercera y superior, vale decir, cuando son medios y no fines.[Horkheimer, p.18]

Las críticas contemporáneas a la razón, a sus límites, pueden entenderse en parte como la imposibilidad de concebir una razón objetiva como la descrita o de presentarla como una ilusión. Ese proceso comienza negando contenido objetivo a los fines, y se extiende finalmente a toda

realidad particular, “vaciadas de su contenido todas las nociones fundamentales se han convertido en meros envoltorios formales. Al subjetivarse, la razón también se formaliza.” [Horkheimer, p.19]. Una consecuencia de esta formalización es que la razón no puede determinar si un fin es de por sí deseable o mejor que otro. Tampoco puede determinar la validez de un ideal, o los principios éticos o políticos. Todas estas decisiones dependen de factores que rebasan a la razón. Los sistemas filosóficos de la razón objetiva asignaban a la ciencia el fin de descubrir la estructura objetiva. Por las razones dadas anteriormente, “Se oponían a toda teoría epistemológica que redujera la base objetiva de nuestra comprensión a un caos de datos descoordinados y que convirtieran el trabajo científico en mera organización, clasificación o cálculo de tales datos.” [Horkheimer, p.23]. También rechazaban una ciencia que sólo podía discutir acerca de valores ‘teóricos’ como verdad, grado de predictibilidad o de precisión.

Al perder fuerza el concepto de razón, al formalizarse, queda debilitada frente a los manejos ideológicos y al relativismo. La Escuela de Francfort reivindicó parcialmente la razón objetiva, especialmente argumentó en favor de una idea de razón que pudiera discutir acerca de fines y pudiera determinarlos, tanto como de los principios de la ética, de la economía y de la política. No aceptaba un dualismo de la razón como la kantiana (razón teórica y razón práctica).

El deontologismo es filosóficamente muy atractivo y pretende poner a nuestro alcance ámbitos típicamente humanos pero muy problemáticos. Sus principales debilidades han consistido en permanecer más como un programa que como una realización y en la ausencia de un conjunto de criterios para *fundamentar* las normas de manera análoga a la *fundamentación* por el método científico de una hipótesis empírica.

Las teorías consecuencialistas

El consecuencialismo, generalmente, niega que se pueda argumentar acerca de fines o valores. Las premisas de un argumento estándar están constituidas por enunciados verdaderos o falsos, lo mismo que sus conclusiones. En cambio una oración valorativa, no es ni verdadera ni falsa, es simplemente una norma. Estas inhiben o incitan a una acción. Por lo tanto, las normas no pueden formar parte de los argumentos.

Existen lógicas no estándar que implican argumentos que aceptan premisas no asertivas (ni verdaderas ni falsas), pero se puede optar por

privilegiar la lógica estándar para la metodología, como lo hace Popper (1972, pp.276-287) y rechazar todo otro tipo de inferencia. O también, como lo sugerimos al final del apartado anterior, aunque aceptemos otras lógicas alternativas, los criterios para “validar” una oración normativa no están claros, excepto en algunos contextos excepcionales como los jurídicos, y así se llega a la misma conclusión, solo se aceptan oraciones que puedan calificarse de verdaderas o falsas.

Una definición positiva del consecuencialismo implica la selección de una acción debido a las consecuencias a las que conduce. Una acción no se acepta o se realiza por su naturaleza intrínseca o por las circunstancias o condiciones que la rodean, por ejemplo, las motivaciones que la causan, sino solo por sus consecuencias.

El utilitarismo

El consecuencialismo se asocia usualmente al utilitarismo. Este surgió, primero, como una teoría ética y uno de sus principales expositores fue Jeremy Bentham. Formuló la ley básica del obrar: *actuar de acuerdo al propio interés, que se manifiesta en la búsqueda del placer y en evitar el dolor*. O en sus propias palabras con las que comienza *An Introduction to the Principles of Morals and Legislation*,

Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne. They govern us in all we do, in all we say, in all we think: every effort we can make to throw off our subjection, will serve but to demonstrate and confirm it. In words a man may pretend to abjure their empire: but in reality he will remain subject to it all the while. The *principle of utility* recognizes this subjection, and assumes it for the foundation of that system, the object of which is to rear the fabric of felicity by the hands of reason and of law. [cap.1, punto I]

Para los utilitaristas lo que mueve al hombre a actuar es la búsqueda de placer. Mediante el principio de utilidad una acción es correcta o incorrecta

según aumente o disminuya el placer. Este principio supone qué significa *utilidad*: la propiedad que produce placer,

By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing) or (what comes again to the same thing) to prevent the happening of mischief, pain, evil, or unhappiness [Bentham, cap. 1, punto III]

Estas afirmaciones son muy claras respecto del estatus ontológico de la utilidad: es una propiedad objetiva de las cosas. No es una reacción o estado de un sujeto provocado por un objeto; no es relativa al sujeto como lo será para el marginalismo económico. La objetividad de la utilidad constituye la base para que se pueda medir de la misma manera que la longitud, el tiempo o el peso.

El individualismo radical implícito en el utilitarismo es matizado porque reconocen que los individuos poseen sentimientos sociales, especialmente la simpatía que tendrá en A. Smith un representante destacado. Mediante los sentimientos sociales se reconoce que los demás también buscan el placer. De esta manera, el hombre debe obrar para alcanzar el placer para el mayor número de personas, es decir, *la felicidad máxima, obrar de manera de alcanzar la mayor felicidad para el mayor número*, es la medida de lo correcto e incorrecto.

Bentham expuso una aritmética de los placeres basada en dos principios: 1. Todos los placeres son cualitativamente iguales, por lo tanto pueden ser medidos; 2. Los placeres de diversas personas pueden ser comparados entre sí. El placer o el dolor, para un individuo o en sí mismos, pueden medirse teniendo en cuenta los siguientes elementos o dimensiones: la *intensidad, duración, certeza, proximidad y seguridad* [Bentham, cap. IV, punto II]. La medición del placer o dolor para un conjunto de personas debe agregar estas dos dimensiones suplementarias: *fecundidad y pureza* [Bentham, cap. IV, punto IV]. Indudablemente está pensando en valores aditivos, en términos contemporáneos, en una escala proporcional,

[Para un conjunto de individuos] *Sum up* the numbers expressive of the degrees of *good* tendency, which the act has, with respect to each individual, in regard to whom the tendency of it is *good* upon the whole: do this again with respect to each individual, in regard to whom the tendency of it is *good* upon the whole: do this again with respect to each individual, in regard to

whom the tendency of it is *bad* upon the whole. Take the *balance* which if on the side of *pleasure*, will give the general *good tendency* of the act, with respect to the total number or community of individuals concerned; if on the side of pain, the general *evil tendency*, with respect to the same community. [Bentham, cap.4, punto 6]

Tipos de consecuencialismo

El consecuencialismo aparenta una atrayente simplicidad, sin embargo, presenta muchos problemas de los cuales solo formulamos dos, ya convertidos en clásicos, para ilustrarlos.

Perjudicar a un inocente. Supongamos que tres personas necesitan un trasplante cada una y que se elige a una sana como donante el cual implicará su muerte. ¿Es lícito realizarlo al aumentar las consecuencias positivas por salvar tres vidas al precio de una?

No cumplir las promesas. Un individuo promete realizar una acción y en el momento que debe realizarla estima que las consecuencias serán más beneficiosas si no cumple su promesa. ¿El consecuencialismo es incompatible con el cumplimiento de las promesas?

Otra perspectiva para evaluar las complejidades del consecuencialismo reside en ver las diferentes variantes que presenta. Se puede formular de alguna de las siguientes maneras, entre otras muchas. El consecuencialismo es

Actual: cuando se toman en cuenta solo las consecuencias actuales contra las previsibles o las que se tenía intención de obtener;

Directo: si las consecuencias que cuentan son las derivadas de la acción y no se consideran los motivos, las reglas o las costumbres que impulsan otros actos semejantes;

Maximizador: se selecciona la acción maximizadora no meramente la que es satisfactoria o meramente mejora un estado de cosas;

Utilitarista: se toman en cuenta las acciones en el que el balance entre las consecuencias de placer (o útiles) predominan;

Universal: se consideran las consecuencias para todos los seres no solo para un individuo.

Para obtener consecuencias se debe partir de premisas. Según cuáles sean estas, las consecuencias serán diferentes. Esta es la clave para entender los tipos de utilitarismo recién enumerados y para resolver los problemas descriptos arriba. El otro aspecto problemático consiste en la evaluación de las consecuencias, cuáles son placenteras o dolorosas – típico del utilitarismo-, cuáles son mejores que otras, aceptables o no; criterios evaluativos que definen un tipo diferente de consecuencialismo no utilitarista.

El consecuencialismo de la teoría de la decisión

En la sección anterior hemos señalado el papel crucial que desempeñan las premisas específicas de las que se parte para obtener cierta clase de conclusiones. Es decir, no basta afirmar en un debate que se es consecuencialista sino también especificar de cuáles premisas y suposiciones adicionales se parte para obtener las consecuencias. Además, cuál es el o los criterios evaluativos para separar en dos clases por lo menos, a las conclusiones, las deseables (o las que producen placer, las mejores, las útiles), de las restantes. Premisas y criterios evaluativos forman el núcleo del consecuencialismo.

Si examinamos la teoría de la elección estándar bajo riesgo o incertidumbre está suficientemente claro desde von Neuman y Morgenstern [1967] cuáles son esas premisas y criterios evaluativos –lo cual no es tan claro en las discusiones éticas. Aunque la teoría está formalizada y axiomatizada explícitamente, realizamos una exposición más intuitiva.

Se suele representar los elementos involucrados en una decisión bajo riesgo o incertidumbre mediante una matriz en que las filas representan los cursos de acción o acciones, las columnas los estados del mundo y en cada celda aparecen resultados. Los resultados comúnmente se expresan en valores monetarios correspondientes a cada acción según el estado del mundo considerado y en la probabilidad asociada a ese estado -será la probabilidad frecuencial bajo riesgo y subjetiva bajo incertidumbre. El criterio de evaluación de los resultados es la suma de cada uno de los resultados ponderados por su probabilidad en cada fila que produce el *valor monetario esperado* de esa acción. O una evaluación alternativa puede consistir ahora en la utilidad que produce obtener cierto resultado monetario y encontrarse en ese estado del mundo, y la suma de cada uno

de los resultados ponderados por su probabilidad en una fila se denomina *utilidad esperada* –recordemos que en la elección bajo certeza la evaluación se reduce simplemente a la maximización de la utilidad. Se ejemplificará primero con los valores monetarios esperados por ser más intuitivos.

Representemos mediante una matriz la siguiente situación: en su estadía de diez meses en la ciudad de Posadas E.S. duda acerca de comprarse un auto nuevo o bien comprar un usado (y en cualquier caso venderlo al terminar su estadía). El usado puede descomponerse y obligarlo a alquilar otro, riesgo que es mínimo con el auto nuevo. Piensa recobrar dinero al vender el auto nuevo al final de su estadía y reducir la pérdida a 30.000 \$; y 12.000 con el auto usado. El alquiler del auto insume 2.400 \$ mensuales. La probabilidad de no alquilar comprando un auto nuevo es de .9 y de .5 con ese auto usado. La representación de estos hechos quedará de esta manera,

	No es necesario alquilar	Se debe alquilar
Comprar el auto nuevo	-30.000 \$.90 27.000 \$]	-54.000 \$ [-.10 [-5.400 \$]
Comprar el auto viejo	-12.000 \$.50 [-6.000 \$]	-36.000 \$ [-.50 [-18.000 \$]

FUENTE: Adaptado de Resnik, p.86

Mediante el método de los valores monetarios esperados se puede calcular las pérdidas de una y otra opción, con la estrategia de la primera fila E.S. incurre en un gasto esperado de -32.400 \$ y comprando un auto usado de -24.000 \$. Si se maximiza el valor esperado se decidirá por la segunda fila, comprar el auto usado, pues significa una pérdida menor (el valor esperado es mayor).

Ahora bien, en esta matriz se considera el costo del auto pero no están representados el placer de manejar un auto nuevo, el temor a que el auto viejo falle, etc. Esto indica que se puede generalizar el método y en lugar de considerar solamente los valores monetarios de las acciones también considerar esos aspectos y reemplazar los valores monetarios por la utilidad de cada uno de los resultados,

	No es necesario alquilar	Se debe alquilar
Comprar el auto nuevo	.90 $u = 500$.10 $u = 50$
Comprar el auto viejo	.50 $u = 100$.50 $u = 70$

FUENTE: Elaboración propia

Es decir, pasamos de (1) a (2) para la primera acción, del valor monetario esperado a la utilidad esperada,

$$(1) \quad .90 (30.000 \$) + .10 (-5.400 \$) = -32.400 \$$$

$$(2) \quad .90 (500) + .10 (50) = 455$$

Calculando la utilidad esperada para cada acción en la segunda tabla se obtiene 455 y 85 respectivamente y si maximiza la utilidad esperada elegirá ahora la primera acción, la de mayor utilidad esperada.

¿Cuáles son las consecuencias? Se advierten más fácilmente en la primera tabla, el gasto incurrido en la compra el automóvil nuevo o en la del usado, el posible gasto en el alquiler de un auto, las probabilidades objetivas (probabilidad frecuencial) de que ocurra uno u otro evento. En la segunda los valores monetarios son reemplazados por las utilidades y la probabilidad (subjetiva) de obtenerla. La evaluación en el primer caso se realiza mediante la estimación de los valores monetarios –como lo muestra la expresión (1). Para el caso de la segunda tabla se evalúa mediante la utilidad esperada. Si para cada acción los estados son

exclusivos y exhaustivos, entonces, la suma de la probabilidad de cada uno da 1, y puede definirse de la siguiente manera,

$$(3) \quad E = \sum_{i=1}^n p(A_i) u(c_i)$$

El uso de la utilidad esperada para decidir puede considerarse una regla descriptiva o prescriptiva que se remonta a Daniel Bernoulli [1783]. Si consideran dos acciones x e y la regla de decisión consiste en maximizar la utilidad esperada que puede formalizarse de la siguiente manera,

$$(4) \quad x \text{ es preferida a } y \quad \text{sii} \quad \sum_{i=1}^n p(A_i) u(c_i) > \sum_{i=1}^n p(A'_i) u(c'_i)$$

En las discusiones actuales acerca de la teoría de la decisión bajo riesgo o incertidumbre de hecho se han propuesto en reiteradas ocasiones cambios en las premisas o premisas suplementarias con las cuales se obtienen otras consecuencias. Generalmente estos cambios van asociados a los ajustes de esta teoría para que dar cuenta de los datos experimentales que se obtienen cuando se prueba la teoría e intentar mantenerla dentro del consecuencialismo y el cognitivismo.

Se demuestra con cierta facilidad que la teoría estándar merece el nombre de consecuencialista. Veremos, sin embargo, en la próxima sección que al explicar ciertos sesgos no previstos en la teoría, se introducen factores que la volverían no consecuencialista y no puramente cognitivista.

Nuevos aspectos no consecuencialistas de la TD

Kahneman y Tversky [1979] han estudiado extensamente y de manera experimental diversas anomalías en la teoría de la utilidad esperada. La primera que se expone (1979, p.266) consiste en elegir entre A (4000\$, .80) o B (3000\$). La mayoría de los sujetos eligen B (80%) respecto de A (20%). Sin embargo, cuando se les da a elegir entre C (4000\$, .20) o D (3000, .25) los sujetos eligen C (65%) y D (35%). Aunque ambas elecciones son inconsistentes.

Ahora bien, cuando se les da a elegir entre perder 4000\$ pesos con alta probabilidad o perder con seguridad 3000\$, invierten la elección anterior.

Es decir, A' (-4000, .80) o B' (-3000); 92 sujetos eligieron A' y solo 8 eligieron B' (1979, p.268).

¿Cómo interpretar estos resultados? La violación de la utilidad esperada se explica en el primer caso por el “efecto certeza”, es decir, por la aversión al riesgo; se elige la opción segura antes que una ganancia esperada superior. Cuando los signos se invierten representando pérdidas, hay un efecto “reflejo”, la aversión al riesgo en el dominio de las ganancias, se reemplaza por la búsqueda del riesgo en el dominio de las pérdidas. Esto muestra que las preferencias parecen estar determinadas por las actitudes o las emociones frente a ganancias y las pérdidas y el cambio de la aversión a la propensión al riesgo implican un punto de referencia, es decir, para los agentes es más significativo las variaciones de riqueza que los niveles de la misma –muy común en fenómenos perceptivos [cfr. Kahneman, pp.191-195].

Tversky y Kahneman presentan un interesante efecto denominado *marco*. Consiste en dos descripciones de opciones con pequeñas diferencias pero que no varían extensionalmente y que sin embargo los agentes deciden de manera diferente. El siguiente ejemplo ilustra el fenómeno. La gripe A que afectará a nuestro país este invierno causará la muerte de 600 personas. Se proponen dos métodos alternativos para hacerle frente con las siguientes consecuencias: Si se elige el tratamiento A se salvarán 200 personas; si se elige el B hay una probabilidad de un tercio que se salven 600 personas y dos tercios de que no se salve ninguna. La mayoría elige la alternativa A.

Se seleccionan al azar parte de los que respondieron y ahora se les pide decidir en base a esta presentación: Si se elige el tratamiento A', morirán 400 personas. Si se elige el tratamiento B' hay una probabilidad de un tercio que no muera ninguno y una probabilidad de dos tercios que mueran 600. Con esta formulación una gran mayoría ahora se inclina por B'. Es evidente la reacción afectiva involucrada en la decisión, la certeza de salvar vidas es altamente atractiva y una desproporcionada aversión a aceptar la muerte segura de personas [ver Kahneman, 2003, p.197-199]. Ante la carencia de formas canónicas de formulación de las opciones, es inevitable la intervención de factores contextuales en la toma de decisiones.

Interpretaciones de algunos sesgos de la teoría de la utilidad esperada

Seguiremos a Damasio (2010) en la explicación de estas anomalías o sesgos en la teoría de la decisión, aunque es una teoría más general que da cuenta de las interacciones inescindibles entre mente y cuerpo, en especial entre cognición y emoción. Este autor elaboró una hipótesis para explicar ciertos hechos observados en individuos que tenían lesiones cerebrales relacionadas con la incapacidad para tomar decisiones. Formula una hipótesis audaz sobre las interrelaciones inextricables entre emociones y racionalidad.

Veamos brevemente los hechos. Observó en su gabinete que ciertos sujetos con lesiones cerebrales estaban impedidos de evaluar distintas opciones aunque las generaban (Damasio, p.73). De manera más específica encuentra que las lesiones en las cortezas prefrontales, en el sector ventromediano, se asocia con deterioros en el razonamiento/toma de decisión y de las emociones/sentimientos, es decir, en los sistemas dedicados al razonamiento y la toma de decisiones está implicada la emoción y el sentimiento y se dedican en parte al procesamiento de las señales procedentes del cuerpo.

¿Cómo comprender esta alianza que no solo cuestiona explicaciones neurobiológicas y científicas sino también epistemológicas y filosóficas – recordar la posición extrema de Descartes respecto a la separación de estas dos *substancias*?

Un organismo es incapaz de sobrevivir si no regula biológicamente su interacción con el medio y con su propio cuerpo. Los impulsos e instintos generan comportamientos que contribuyen a la supervivencia, desde un extremo que implica salvar la vida al otro extremo que implica suministrar condiciones más ventajosas para el organismo o reducir condiciones dañinas.

Las estructuras cerebrales que hacen funcionar la regulación biológica básica regulan también el comportamiento y los procesos cognitivos.

Damasio no pretende enunciar un reduccionismo biológico, del comportamiento social al comportamiento biológico, sino por el contrario, lo sociocultural tiene propiedades irreductibles pero surgieron por interacción con el nivel biológico y no funcionan sin él (Damasio, p.151). Y en particular, la regulación no está subsumida en la región subcortical

(evolutivamente más vieja) y el razonamiento en la región neocortical. Ambas se interrelacionan y se implican mutuamente en el comportamiento humano; en particular, las emociones y sentimientos proporcionan *el puente* entre ambas regiones, entre los procesos racionales y no racionales (Damasio, p.155).

Emociones y sentimientos son procesos regulatorios importantes. Una emoción primaria es una pauta innata de reacción corporal específica (rubor, temblor, respiración agitada) desencadenada por un estímulo. Las emociones secundarias se desencadenan luego de un proceso evaluativo, voluntario y no automático entre el estímulo y el estado emocional del cuerpo lo cual mejora el sistema de protección de la emoción. En los sentimientos los procesos corporales de la emoción se superponen con los pensamientos de los contenidos específicos que le dieron lugar, como un paisaje o una canción. Los sentimientos suponen tanta cognición como cualquier otra imagen perceptual, lo que los diferencia e individualiza es que se refieren al cuerpo y nos ofrecen la cognición de él (Damasio, p.189). Por sus vínculos inescindibles con el cuerpo y por haberse originado filogenéticamente antes, tienen primacía y penetran nuestra vida mental.

Hemos señalado que cuando se razona y se decide hay que tener conocimiento sobre la situación que se decide, sobre las diferentes acciones, sobre los estados del mundo en cada acción planteada, sobre las consecuencias o resultados de cada acción en un estado del mundo y la valoración de los resultados. Siempre se ha destacado la intervención de diferentes procesos funcionales que posibilitan lo anterior como la atención, la memoria y el pensamiento, pero nunca se señalan el papel de las emociones y sentimientos “y no se oye casi nada sobre el mecanismo que genera un repertorio de opciones diferentes para su selección” (Damasio, p.197). En la selección de la respuesta desempeñan un papel fundamental las emociones y sentimientos.

Hay dos hipótesis básicas para explicar el proceso de selección de los diferentes elementos descriptos en el párrafo precedente: el ojo de la mente -la razón- selecciona o, de acuerdo a Damasio, el marcador somático desempeña también ese mismo papel.

Cuando un resultado malo conectado a una respuesta u opción aparece en la mente por fugaz que fuere se experimenta un sentimiento desagradable en el cuerpo. Porque tiene que ver con el cuerpo Damasio lo denomina somático y porque marca una imagen lo denomina marcador.

El marcador somático hace prestar atención a los resultados negativos o positivos. Si es negativo se dispara una alarma automática y puede conducirnos a rechazar inmediatamente una opción y elegir otras. Nos protege de pérdidas futuras y reduce el número de alternativas a examinar. Ahora actúa el proceso racional evaluativo pero luego de la drástica reducción de las opciones. En palabras de Damasio (p.205) “los marcadores somáticos son un caso especial de sentimientos generados a partir de emociones secundarias. Estas emociones y sentimientos han sido conectados, mediante aprendizaje, a resultados futuros predecibles de determinados supuestos.” Por el contrario un marcador positivo es una guía o incentivo.

Los marcadores no sustituyen a la razón, sino colaboran con ella desechando ciertas alternativas o resaltando otras. Es un conjunto cambiante que se va formando y evolucionando con nuestro desarrollo personal y social.

Las emociones según este enfoque nunca dejan de estar presentes tanto cuando se toman decisiones como en cualquier comportamiento en que se halle involucrada la racionalidad. En la teoría de las decisiones el factor emotivo puede ser muy importante, como en situaciones de aversión al riesgo, o menos importante y pasar inadvertidas por completo en situaciones que nos provocan indiferencia. Pero siempre están presentes, y lo hacen en la selección de las acciones a considerar, el despliegue de los estados del mundo a considerar, el conjunto de consecuencias que nos parece satisfactorio y en la evaluación, es decir, en todas las instancias de la decisión. La emoción no deteriora, paraliza o hace funcionar distorsionadamente la razón, por el contrario, es un elemento junto a la razón que no la vuelve estéril. Nos informa qué premisas seleccionar o continuar buscando, qué consecuencias son satisfactorias de las infinitas que se pueden extraer de un conjunto de premisas, y colabora en persuasión de ciertas conclusiones en el caso de contextos argumentativos no deductivos como los retóricos.

Las teorías que introducen las emociones como un aspecto explicativo de las decisiones sin reducirlas a un aspecto meramente cognitivo son muy variadas. Señalaremos solamente algunas mencionadas en Loewenstein et al. (2001, pp.268-9) alternativas a la expuesta. Estos autores proponen la hipótesis de los sentimientos bajo riesgo, que da cuenta de dos fenómenos muy abarcativos. La primera es cómo las evaluaciones cognitivas dan lugar a respuestas afectivas; la dirección inversa explica

las importantes influencias de las emociones en la cognición. Otra perspectiva la encontramos en Zajonc (1980) y LeDoux (1996) que comparan las evaluaciones cognitivas con las reacciones afectivas a estímulos y concluyen que las últimas son más rápidas y más básicas, proveen al organismo de una casi inmediata aunque tosca evaluación de las opciones a que se enfrenta. Estos resultados son compatibles con afirmaciones de Simon (1967, por ejemplo) en que los procesos psicológicos, en particular, las reacciones emocionales, sirven para detener o redirigir los procesos de decisivos en diferentes etapas. O los trabajos de Slovic (1991) y su equipo que propusieron una heurística afectiva que ilumina la importancia del afecto para las percepciones del riesgo y el comportamiento basado en decisiones de riesgo. Y por último mencionaremos a Elster quien también ha sido pionero en proponer una teoría no consecuencialista de la decisión mediante la introducción de normas o emociones (1991, 1998).

Conclusiones

Las perspectivas que implican la introducción de las emociones como factor explicativo de las decisiones tienen repercusiones en distintos campos. Comencemos por la economía. Si resulta inevitable considerar el papel explicativo de las emociones, podemos entenderlo desde al menos dos posiciones antagónicas. La primera consiste en aceptar el factor emoción en la toma de decisiones y recurrir a la psicología junto con la neurobiología como las disciplinas que se encargan de caracterizarlas. Esta ampliación explicativa podría conducir a una redefinición de los campos disciplinares formulando un nuevo recorte a través de un funcionamiento intradisciplinal, una especie de nueva ciencia biosocial. La mayor parte de la literatura mencionada en la última sección y la mayoría de las investigaciones de economía comportamental conduce a esta perspectiva.

La segunda, rechaza la disolución o absorción de la economía en el primer enfoque y propone volver a los orígenes de la definición del campo económico de la economía política clásica. La economía es el estudio de los mercados (enfaticando implícitamente los comportamientos agregados, holísticos para escapar a la explicación biopsicológica de las conductas individuales). Así, por ejemplo, L. Blau (2010, pp.122-131) señala el peligro de la disolución de la economía y reivindica la definición clásica.

Más allá de la economía, estas discusiones tienen impacto en el campo filosófico y especialmente en el campo ético. Sin abordarlo detalladamente resulta claro que el campo de las teorías éticas (pero también científicas) no se reduce a la dicotomía consecuencialismo/deontologismo. La negación del consecuencialismo no nos conduce inevitablemente a seleccionar acciones por los valores, sino a criterios distintos en que la racionalidad y las emociones tienen su papel –eventualmente si se aclarase cómo discutir acerca de valores estos también podrían introducirse. Se amplían la tipología de teorías éticas y los criterios evaluativos que suponen.

Por último, afecta a nuestra concepción de la racionalidad. Si bien las emociones impulsan y modifican los comportamientos, se habían considerado un factor distorsivo que el individuo perfectamente racional debía eliminar. A la luz de los avances científicos actuales el panorama es radicalmente diferente. No es consistente con los hechos una racionalidad puramente consecuencialista y *separada*, en la que no desempeñe un papel factores tradicionalmente dejados fuera de su caracterización como la emoción, que examinamos en este artículo en el contexto de la teoría de la decisión.

Bibliografía

- Bentham, J. (1823), *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press, 1907. Reproducida on-line en: <http://oll.libertyfund.org>
- Bernoulli, D. (1954) [1738], “Exposition of a New Theory on the Measurement of Risk”. *Econometrica*, v.22, n.1, January, pp.22-36. Translated from Latin into English by L. Sommer.
- Blaum, L. (2010), *El objeto de la economía política, ética y hermenéutica social*. Tesis. Universidad Tres de Febrero.
- Elster, J. (1991), *Juicios Salomónicos*. Gedisa.
- _____ (1998), “Emotions and Economic Theory”. *Journal of Economic Literature*, v.36, n°1, pp.47-74.
- Horkheimer, *Crítica a la razón instrumental* [1973]

- Kahneman, D., "Mapas de racionalidad limitada: psicología para una economía conductual", Revista Asturiana de Economía, nº28, 2003, pp.181-223.
- Kahneman, D., A. Tversky (1979), "Prospect Theory: A Analysis of Decision under Risk", Econometrica, v.47, nº2, pp.263-92.
- LeDoux, J. (1996), *The Emotional Brain*. Simon & Schuster.
- Loewenstein, G.F., E.U. Weber, Ch.K. Hsee, N. Welch (2001), "Risk as Feelings", Psychological Bulletin, v. 127, nº2, pp.267-286.
- Popper, K. (1972), *El conocimiento Objetivo*. Tecnos.
- Simon, H. (1967), Motivational and emotional controls of cognition. Psychological Review, 57, pp.386-420).
- Slovic, P., M. Lyman, N. Krauss, J. Flynn, J. Chalmers, G. Gesell (1991), Perceived Risk, stigma and potential economic impacts of a high-level nuclear waste repository in Nevada". Risk Analysis, 11, pp-683-696.
- Resnik. M. D. (1998), *Elecciones –Una introducción a la teoría de la decisión*. Gedisa.
- Von Neumann, J., O. Morgenstern (1967), *The Theory of Games and Economic Behavior*. Wiley.
- Zajonc, R.B. (1980) "Feeling and Thinking: Preference need no inference. American Psychologist, 35, pp.151-175.

PROSPECT THEORY: ¿UNA TEORÍA DE LA DECISIÓN NEOCLÁSICA?

Diego Weisman (CIECE-FCE-Universidad de Buenos Aires / CONICET)

Resumen

Expected Utility Theory (EUT), surge en forma de teoría axiomática en 1944, y rápidamente es reconocida por la profesión como la teoría de la decisión en condiciones de riesgo e incertidumbre capaz de brindar un fundamento adecuado para multitud de modelos de la economía neoclásica. En 1979 Daniel Kahneman y Amos Tversky, dos psicólogos interesados en el trabajo interdisciplinario, hacen pública su *Prospect Theory*, una teoría rival que incorpora al análisis de las decisiones humanas factores psicológicos, haciendo justicia a un viejo reclamo por el realismo que Herbert Simon había proferido en 1955. Así como EUT fundamenta a la economía neoclásica, *Prospect Theory* hace lo propio con *Behavioral Economics*: sus modelos incorporan agentes con características psicológicas que surgen del trabajo de laboratorio. ¿En qué se apartan los modelos de *Behavioral Economics* de los modelos neoclásicos?

La respuesta estándar a la pregunta precedente es que *Behavioral Economics* está basada en *Prospect Theory*, y ambos incorporan agentes no racionales. La tesis de este trabajo es que esa respuesta es errónea, si por racionalidad se entiende maximización. *Behavioral Economics* no se diferencia de la Economía Neoclásica por contar con agentes que no maximizan. La relevancia práctica del tema está dada porque muchos economistas toman a la maximización como sinónimo de racionalidad en el contexto de elección: los agentes racionales son agentes maximizadores. Como veremos este es el óbolo que los neoclásicos pagan a la tradición clásica, y en general es una convención inofensiva, salvo en el presente contexto, donde puede conducir a confusiones y enredos de diversa magnitud.

Introducción

E. Roy Weintraub define la economía neoclásica¹ por tres características:

- 1-Los agentes poseen preferencias racionales sobre resultados
- 2-Los individuos maximizan su utilidad, y las firmas sus beneficios
- 3-Las personas actúan independientemente sobre la base de información completa y relevante.

Expected Utility Theory (EUT), surge en forma de teoría axiomática en 1944, y rápidamente es reconocida por la profesión como la teoría de la decisión en condiciones de riesgo e incertidumbre capaz de brindar un fundamento adecuado para muchos modelos de la economía neoclásica. En 1979 Daniel Kahneman y Amos Tversky, dos psicólogos interesados en el trabajo interdisciplinario, hacen pública su *Prospect Theory*, una teoría rival que incorpora al análisis de las decisiones humanas factores psicológicos, haciendo justicia a un viejo reclamo por el realismo que Herbert Simon había proferido en 1955. Así como EUT fundamenta a la economía neoclásica, *Prospect Theory* hace lo propio con *Behavioral Economics*: sus modelos incorporan agentes con características psicológicas que surgen del trabajo de laboratorio. ¿En qué se apartan los modelos de *Behavioral Economics* de los modelos neoclásicos? La respuesta habitual suele hacer pasar la diferencia específica por la

¹ “Neoclassical economics is what is called a metatheory. That is, it is a set of implicit rules or understandings for constructing satisfactory economic theories. It is a scientific research program that generates economic theories. Its fundamental assumptions are not open to discussion in that they define the shared understandings of those who call themselves neoclassical economists, or economists without any adjective. Those fundamental assumptions include the following:

1. People have rational preferences among outcomes.
2. Individuals maximize utility and firms maximize profits.
3. People act independently on the basis of full and relevant information.

Theories based on, or guided by, these assumptions are neoclassical theories.” E.Roy Weintraub, The Concise Encyclopedia of Economics, entrada “Neoclassical Economics”, on line.

naturaleza de los agentes: en *Behavioral Economics*, a diferencia de los modelos neoclásicos, los agentes son irracionales, se afirma.

Esta respuesta no es indefendible, pero en este trabajo quisiera mostrar algunas objeciones posibles y señalar caminos a evitar. El término racionalidad es sin duda complicado, tanto por la cantidad de asociaciones normativas, descriptivas y prescriptivas que lleva consigo, como por las múltiples tradiciones que convergen sobre él, borrando sus contornos. Además, como posee el dudoso privilegio de ser un *common sensical term*, muchos sienten que es un viejo conocido que se acomoda amablemente a casi cualquier condición que uno le imponga, y del que nada puede sorprendernos. Pero es justamente el hecho de que sea un viejo conocido el que hace imprescindible analizarlo con doble esfuerzo.

La respuesta obvia a la pregunta precedente es que *Behavioral Economics* está basada en *Prospect Theory*, y ambos incorporan agentes no racionales. La tesis de este trabajo es que esa respuesta es errónea, si por racionalidad se entiende maximización. *Behavioral Economics* no se diferencia de la Economía Neoclásica por contar con agentes que no maximizan. La relevancia práctica del tema está dada porque muchos economistas toman a la maximización como sinónimo de racionalidad en el contexto de elección: los agentes racionales son agentes maximizadores. Como veremos este es el óbolo que los neoclásicos pagan a la tradición clásica, y en general es una convención inofensiva, salvo en el presente contexto, donde puede conducir a confusiones y enredos.

En la vida diaria usamos la racionalidad con dos propósitos muy diferentes, aunque interrelacionados. El primero es el de dar sentido a la conducta de otras personas. Lo llamaremos “uso hermenéutico” de la racionalidad. Para interpretar o entender las acciones de la gente que nos rodea *suponemos que son racionales*; ante un acto que nos sorprende, ensayamos y vamos descartando diversos intentos sucesivos de comprensión. Si al final del día ningún intento queda en pie, sea porque el acto es realmente desusado, sea por nuestra incapacidad de generar interpretaciones consistentes, simplemente renunciamos a comprender. En tales casos suele afirmarse “X está loco”, lo cual es una trabajosa manera de decir que su conducta es incomprensible para nosotros.

El segundo propósito es menos evidente, pero no menos habitual. Se utiliza generalmente cuando se pretende *anticipar* la conducta de alguien. Podemos denominarlo “uso predictivo” de la racionalidad. Supongamos

que seguimos por televisión las andanzas de un dirigente político, que en este momento formula declaraciones inusualmente amistosas en relación a un candidato del arco opositor. Una vez que le damos *sentido* a esta conducta, inventando fines potenciales que hacen compatible la conducta observada con la racionalidad de esa persona, podemos *anticipar* conductas: si los fines imputados al político son, por ejemplo, tejer una alianza con la oposición, podemos prever ciertas acciones. En caso de que la conducta observada sea completamente diferente a la anticipada, lo más probable es que se reconozcan que los fines del sujeto eran otros. Un filósofo de la ciencia expresaría lo anterior diciendo que el uso predictivo de la racionalidad da lugar a reconstrucciones empíricas corregibles, pero por el momento podemos prescindir de esta aparatosa forma de expresión y refugiarnos en el sentido común.

El ejemplo anterior muestra que hay dos usos diferentes para la racionalidad: el de construir interpretaciones y el de anticipar conductas. El uso predictivo supone el uso hermenéutico. Para anticipar conductas se necesita no solamente atribuir racionalidad sino además algún fin específico. Una vez establecido este fin, y dando por supuesto que esa persona es racional, se analiza el contexto para identificar de antemano la próxima jugada.

Muchos economistas comparten de manera implícita este doble uso de la racionalidad. En primer lugar, toman maximización y racionalidad como sinónimos laxos. En segundo lugar, utilizan la maximización como un principio metodológico que les permite comprender conductas, y *también* construyen modelos adjudicándoles fines concretos a los agentes para anticipar su conducta. En ambos casos se afirma que hay agentes maximizadores o racionales *en dos sentidos sumamente diferentes*. Como hemos dicho, confundirlos conduce a enredos de diversa magnitud. En lo que sigue se intentará echar luz sobre algunos de ellos, relativos al status de *Prospect Theory* como teoría de la decisión.

***Prospect Theory* como teoría de la decisión no racional**

Prospect Theory no es una teoría de la decisión racional; más aun, no *puede* serlo. Un requisito mínimo de racionalidad, algo que debe estar en el núcleo de cualquier teoría posible de la racionalidad es la consistencia, un requisito sin duda mínimo pero lamentablemente ausente del campo de las decisiones humanas. Como *Prospect Theory* aspira a ser

descriptivamente correcta en el campo de las decisiones humanas, es mejor que se desprenda de la racionalidad si no quiere fallar en sus pronósticos.

Diversos experimentos muestran que los seres humanos violan el principio de extensionalidad². Diseñado para responder a la pregunta por la igualdad de conjuntos³, el principio de extensionalidad reclama que dos descripciones alternativas del mismo referente sean reemplazables *salva veritate*. En el campo de las acciones humanas, los individuos respetan este principio si y solo si las decisiones son *independientes* de las descripciones, o dicho de otra manera si y solo si el orden de preferencias permanece idéntico ante cambios en la descripción. Esto implica que si alguien prefiere las obras de Goethe a las obras de Cervantes⁴, debería preferir las obras de Goethe a las obras del Manco de Lepanto. Pero supongamos que nuestro hombre revierte ahora sus preferencias por alguna razón (quizás siente aprensión por las personas tullidas, o algo por el estilo). En tal caso, difícilmente afirmaríamos que estamos ante un decisor *racional*. Si este es el caso, ninguna teoría de la decisión que quiera describir la conducta de decisores como el representado podrá ser *racional*. Como vimos, hay un amplio consenso en relación a que *Prospect Theory* no es una teoría de la decisión racional y que además no *puede* serlo:

“Because invariance [is] normatively essential and descriptively invalid, a theory of rational decision cannot provide an adequate description of choice behavior.(..) Consequently, the dream of constructing a theory that is acceptable both descriptively and normatively appears unrealizable” (Kahneman, 1986. p 255)

Pero ¿qué significa que algo sea una teoría de la decisión *racional*? ¿Qué significa el término “racional” en este contexto?

Racionalidad en Economía

² El más conocido sin dudas es el de *Asian Disease*.

³ En teoría de conjuntos, se dice que dos conjuntos son extensionalmente equivalentes si y solo si tienen los mismos miembros.

⁴ Este ejemplo fue sugerido por el prof. Gustavo Marqués.

Los economistas suelen emplear el término *maximización* como sinónimo laxo de racionalidad⁵. Los agentes racionales son agentes maximizadores, que actúan siguiendo su propio interés y lo hacen de la mejor manera posible: es decir, maximizando. Que los agentes sean racionales (maximizadores) puede entenderse como una hipótesis vacía de contenido o como una hipótesis sustantiva, que describe al ser humano o al menos a una subclase de ellos.

Boland⁶ ha dicho que es fútil criticar el principio neoclásico de maximización, y sugiere que de hecho es irrefutable. Por su parte, algunos críticos (v.gr., Caldwell) sostienen que de hecho algunos economistas la toman como una hipótesis tautológica y otros como una hipótesis empírica, y hay quienes, como Simon (1955) o Kahneman (2001), argumentan que no solo es refutable, sino que de hecho está refutado. Obviamente no pueden tener todos razón, a menos que se refieran a cosas distintas. Y de hecho, se refieren a cosas distintas.

La confusión surge, creo, porque hay dos principios de maximización. El primero es un principio puramente vacío y formal, un principio que hunde sus raíces en los clásicos, y más específicamente en su imperativo de explicar las conductas de mercado no por simpatía o buenos sentimientos, sino como el resultado de acciones auto interesadas. En sus orígenes fue un imperativo metodológico, relativo a las *explicaciones* a construir, no un imperativo ontológico respecto a la *naturaleza* humana⁷. Retomado por las

⁵ Desde Adam Smith y Ricardo la racionalidad como maximización ha sido una constante: “ Among the classical economists, such as Smith and Ricardo, rationality had the limited meaning of prefer more to less (...) Their rationality hypothesis was the maximization of profit by the firm, although this formulation was not explicitly achieved in fully generality until the 1880s” Arrow (1986) p.S388.

⁶ De acuerdo con Boland, el Principio Conductual de Maximización es tiene un carácter metafísico y cumple una función metodológica: “In the case of neoclassical economics, the maximization hypothesis plays [a] methodological role. Maximization is considered fundamental to everything; even an assumed equilibrium need not actually be put beyond question, as disequilibrium in a market is merely a consequence of the failure of all decision makers to maximize. Thus, those economists who put maximization beyond question cannot ‘see’ any disequilibria”.

⁷ Este punto puede ser controvertido; un par de líneas adicionales. Sostengo que la hipótesis del auto interés no afirmaba originalmente que los seres humanos sean egoístas *per se*, sino que las *interpretaciones* de su conducta deben suponer

marginalistas, el auto interés devino en maximización, jugada estratégica que abrió la puerta convenientemente al uso de herramientas matemáticas. La justificación conceptual detrás de la maximización como sucedáneo del auto interés ha sido posiblemente la siguiente: el célebre carnicero o panadero que menciona Adam Smith buscan su beneficio guiados por su propio interés, lo que los lleva a tratar de obtener la máxima ganancia -es decir a maximizar. La maximización es la forma *racional* de seguir el propio interés o alternativamente, la búsqueda del propio interés está detrás de la maximización.

El principio de maximización así entendido no postula ningún “fin” particular: establece más bien una instrucción al teórico que puede parafrasearse de la siguiente manera

“Explique toda conducta humana como si fuera el resultado de una maximización de algo”,

o bien

“Construya modelos que supongan agentes maximizadores de alguna cosa”

En términos de Boland la instrucción deriva del carácter metafísico de la hipótesis, que está en el núcleo del programa de investigación neoclásico y por lo tanto es irrefutable por decisión metodológica. En su reconstrucción el principio afirma:

Principio Conductual de Maximización (PCM): *Todos los agentes maximizan algo*

El carácter metafísico de PCM explica que se pueda aplicar a *cualquier* conducta humana. Para hacerlo sólo basta ingenio suficiente como para postular fines adecuados. Como el PCM no postula ninguno, no sorprende que se acomode a cualquiera de ellos. Es una suerte de casillero vacío a

el mínimo grado de simpatía posible. Visto de esta manera, refleja un principio de parsimonia bastante razonable, en el sentido que constriñe al teórico en la construcción de modelos: si quiere modelar conductas de mercado, hágalo con el individuo más despojado en términos éticos posibles. De otra manera, siempre se podría atribuir los fenómenos de crecimiento a la simpatía y los de escasez a la antipatía humanas, lo cual puede ser útil para condenar moralmente a la sociedad, pero no para comprenderla (incidentalmente creo que esta ingenuidad epistemológica es la que Marx le critica a la afirmación de Proudhon de que “la propiedad es un robo”)

completar de manera libre. La Madre Teresa maximiza su utilidad “altruista” de la misma manera que Bill Gates maximiza su utilidad terrenal y el individuo que se arroja por la ventana porque cree que el fin del mundo se acerca maximiza su utilidad percibida. Cualquier conducta es económicamente racional desde el punto de vista de PCM, lo que equivale a decir que es irrefutable. En palabras de Gary Becker:

“[My] analysis assumes that individuals maximize welfare as they conceive it, whether they be selfish, altruistic, loyal, or masochistic”

“Unlike Marxian analysis, the economic approach I refer to does not assume that individuals are motivated solely by selfishness or gain. It is a method of analysis” (Becker, G. 1992)

Resulta claro que tanto Boland como Becker hacen referencia al principio de maximización *metodológico*, PCM, y en tal sentido son herederos directos tanto de la tradición de los clásicos como del uso hermenéutico de la racionalidad que solemos hacer habitualmente para inteligir la conducta de la gente que nos rodea. Este uso es diferente del segundo, que entra en escena una vez que se fijan los fines.

En efecto, el segundo principio de maximización surge cuando se especifica algún fin: por ejemplo, cuando se afirma que los agentes maximizan su utilidad esperada, como en EUT. Una vez que se especifican fines, *el principio adquiere contenido sustantivo*. En el contexto de la Teoría de la Decisión Estándar (*Ordinal Utility Theory* para certeza y *Expected Utility Theory* para condiciones de incertidumbre) los agentes son racionales si:

- 1-Poseen preferencias racionales
- 2-Maximizan su utilidad esperada

Ambas notas pueden reducirse a la segunda, porque maximizar la utilidad esperada supone la existencia de una función de utilidad esperada, la cual surge a partir de preferencias completas, continuas, transitivas, e independientes. Los requisitos puestos a las preferencias están hechos para garantizar un orden bien formado, algo así como una lista de objetos (loterías en el caso de EUT) que mostraría las preferencias del individuo. Sobre ese listado, el principio de maximización establece que se elije el

más alto disponible. En términos explícitos, el Principio Sustantivo de Maximización dice que en circunstancias de riesgo o incertidumbre:

Principio Sustantivo de Maximización (PSM): *Todos los agentes maximizan su utilidad esperada*

A diferencia del PCM, este principio es sustantivo, tiene contenido empírico, se utiliza para anticipar la conducta de los individuos –y no solamente para darle sentido- y *por lo tanto es refutable*. Su función es muy diferente al primer principio, porque se la usa para predecir o anticipar conductas. Si se me permite el énfasis:

PCM ≠ PSM

Así como la racionalidad tiene dos usos –el uso hermenéutico y el uso predictivo-, la maximización se utiliza en dos sentidos. Recordemos que muchos economistas utilizan el término “maximización” como sinónimo de racionalidad, una costumbre venerable por su origen, pero que engendra confusiones. Ahora puede verse que tanto Boland como sus críticos tienen razón: el principio de maximización (PCM) es metafísico e irrefutable, y el principio de maximización (PSM) es refutable y de hecho está refutado para los partidarios de Kahneman y Tversky. No sorprende entonces la heterogeneidad de respuestas que se ha generado en torno a esta hipótesis:

“(...)Tanto defensores como opositores de la hipótesis han discutido extensamente acerca del status del postulado de maximización. A veces la hipótesis es concebida como poseyendo contenido empírico y a veces es formulada como una tautología; a veces constituye la pieza central de un sistema teórico y a veces un artificio apenas mencionable” (Caldwell, 1981)

La confusión entre ambos usos también es patente en los siguientes pasajes de Kahneman (1986):

*“The assumption of rationality has a favored position in economics. It is accorded all the methodological privileges of **a self-evident truth**, (...), a **tautology** (...). Furthermore, the assumption of rationality is protected by a formidable set of defenses (...) **that restrict the significance of any observed violation of the model.**”* (subrayado propio)

Es fácil entrever varios errores metodológicos aquí. El pasaje sugiere que el supuesto de racionalidad posee privilegios metodológicos de ser una verdad tautológica, y además los contraejemplos suelen chocar contra las defensas erigidas por los epígonos de la economía neoclásica. Ahora bien, cualquier estudiante de lógica sabe que una tautología no puede tener contraejemplos empíricos, porque su verdad es *a priori*, por lo que no es posible que ambas afirmaciones se refieran a la misma hipótesis. La dificultad desaparece cuando se reconocen dos hipótesis de maximización/ racionalidad distintas, una con privilegios metodológicos (tautológica *avant la lettre*) y otra con contenido empírico y por ende capaz de sustentar contraejemplos.

Otro de los múltiples ejemplos está en Kahneman, Wakker y Sarin (1997) En el contexto de la diferencia entre la utilidad revelada en las decisiones –*decision utility*- y la utilidad experimentada como producto de esas decisiones –*experienced utility*- se afirma:

*“The elimination of experienced utility from economic thought was justified by important considerations, but it was not costless. Perhaps the heaviest cost was that the exclusive concern with decision utility removed some important problems from the reach of empirical research. In particular, **the proposition that people maximize utility was granted the status of a maintained hypothesis, which is used to constrain the interpretation of other facts but is not itself subject to test.** Admitting experienced utility as a measure of outcomes turns utility maximization into an empirical proposition, which will probably be found to provide a good approximation to truth in many situations and to fail severely in others”* (pág. 397, subr. propio)

Es cierto que *experienced utility* convierte a la maximización de la utilidad (en sentido metodológico, es decir PCM) en una hipótesis **empírica**, porque instancia el “algo” que se maximiza –la utilidad experimentada- y de esa manera permite anticipar conductas. Pero *exactamente lo mismo sucede con EUT*, que instancia el “algo” que se maximiza como utilidad esperada. Ambas están al mismo nivel, como instanciaciones del PCM que permiten arribar a hipótesis con contenido empírico, aplicables en algunos contextos y capaces de “fallar severamente en otros”. Como en el uso predictivo de la racionalidad, una vez que se fijan los fines el modelo se torna empíricamente corregible –es decir, puede fallar.

Hasta ahora hemos argumentado que los dos principios que nos permitían darle sentido y anticipar la conducta de los políticos de la televisión se utilizan también en contextos más académicos, aunque lamentablemente se los suele confundir. EUT es una teoría de la decisión racional que interpretada descriptivamente intenta anticipar conductas, lo cual sugiere que están allí operando tanto PCM como PSM. Avancemos un paso más. ¿Qué hace que EUT sea una teoría de la decisión *racional*? Básicamente la presencia de tres factores interrelacionados:

- 1-Presencia de PCM: explica la conducta de los agentes como maximizando *algo*
- 2-Presencia de PSM: es capaz de anticipar la conducta de los agentes suponiendo que maximizan algo *específico*: la utilidad esperada
- 3-Presencia de preferencias racionales: postula un orden transitivo, completo, continuo e independiente que hace posible la maximización en términos matemáticos.

Así, cuando se afirma que EUT es una teoría de la decisión racional se dice que modela individuos maximizadores (en los sentidos 1 y 2) y que poseen preferencias “ordenadas”.

Racionalidad en *Prospect Theory*

Lo llamativo es que las tres notas que suelen vincularse con racionalidad están presentes en *Prospect Theory*!

- a) Maximización: *Behavioral Economics* suele utilizar individuos maximizadores. Por ejemplo, en *Behavioral Life Cycle Hypothesis*, el “planner” es un sujeto maximizador:

“The doer is assumed to be pathologically myopic, concerned only with current period consumption. At date t the doer is assumed to possess a subutility function $u_t(c_t)$. We assume diminishing marginal utility ($u_t(-)$ is concave in c_t), and also nonsatiation (u_t is strictly increasing in c_t). In contrast, the planner is concerned with maximizing a function of lifetime doer utilities” (Thaler and Sheffrin, pág. 612)

En otros modelos la conducta maximizadora es aún más explícita. Por ejemplo, en *Equity Premium Puzzle*, la maximización es clara:

“The “prospective utility” of a gamble, G, which pays off x_i with probability π_i is given by

$$V(G) = \sum \pi_i v(x_i),$$

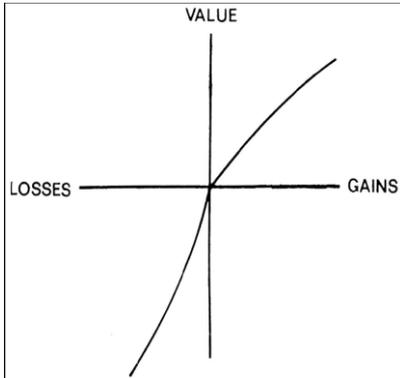
where π_i is the decision weight associated assigned to outcome i .”

*“Therefore, every year Y will solve her asset allocation problem by choosing the portfolio that **maximizes her prospective utility** one year away, just as X does”* (Benartzi and Thaler, 1995, pág. 79 y ss. Subrayado propio)

- b) Preferencias bien formadas: ¿Son las preferencias sobre descripciones de objetos bien formadas? En la medida en que la función valor de *Prospect Theory*, graficada más abajo, sea una función continua y derivable, creo que respuesta es positiva.

Si lo anterior se acepta (aunque admito que es un punto débil de la argumentación, y confieso que no tengo idea cómo resolverlo), entonces podemos pasar revista a las características de *Prospect Theory*:

- 1-Tiene la maximización como principio conductual (PCM)? Si
- 2-Tiene la maximización como hipótesis sustantiva (PSM) ? Si
- 3-Tiene preferencias transitivas, completas, y continuas (Posiblemente si)



Función valor (Kahneman y Tversky, 1979, p. 279)

Entonces hay racionalidad en *Prospect Theory*. Es una instanciación diferente a la de EUT, y ahora podemos ver claramente que la diferencia reside en el *objeto* de la maximización.

Aunque no hay ninguna razón para quedarse con el uso estrecho del término, creo que en la medida en que los economistas continúen tomando a la maximización como un sinónimo laxo de racionalidad, la confusión es inevitable. Si los agentes de *Prospect Theory* realmente maximizan la Utilidad Prospectiva, entonces desde el punto de vista convencional puede considerarse una Teoría de la Decisión *Racional*, a pesar de que:

“It is noteworthy that the everyday usage of the term “rationality” does not correspond to the economists’ definition as transitivity and completeness, that is, maximization of something. The common understanding is instead the complete exploitation of information, sound reasoning, and so forth” (Arrow, 1986, p.S390-1)

Los agentes que violan extensionalidad no son racionales en términos de EUT, porque sus preferencias no guardan un orden bien formado, lo cual impide contar con una función de utilidad esperada, lo que a su vez impide que maximicen la utilidad esperada. Pero el PMC obliga a interpretar *cualquier* conducta observada como un caso de maximización, y además garantiza que puede hacérselo: obedeciendo este imperativo, Kahneman y Tversky demuestran que la conducta anómala puede modelarse como

un caso de maximización: la maximización de la utilidad experimentada, o de la utilidad prospectiva.

Los individuos modelados de esta manera buscan maximizar la utilidad experimentada, lo que parece además una conducta defendible desde el punto de vista normativo. Después de todo, la idea de que las acciones futuras se llevan a cabo para estar mejor en relación al estado presente es una de las tradiciones asociadas con el término racionalidad.

Notablemente, Kahneman ya entrevía en 1981 la posibilidad de que la psicología de la decisión tuviera una carga *normativa*. Para dar el salto, es necesario cambiar el significado de la racionalidad, dándole notas sustantivas:

“The psychology of choice is also relevant to the normative question of how decisions ought to be made. [...] the modern theory of rational choice has adopted the coherence of specific preferences as the sole criterion of rationality. [...] Consistency is only one aspect of the lay notion of rational behavior. As noted by March (26), the common conception of rationality also requires that preferences or utilities for particular outcomes should be predictive of the experiences of satisfaction or displeasure associated with their occurrence. [...] In particular, predictive considerations may be applied to select the decision frame that best represents the hedonic experience of outcomes” (Kahneman y Tversky, 1981)

La estrategia de Kahneman es postula que existe otro concepto de racionalidad que no tiene que ver con la consistencia, sino con la tradición de Hume: son racionales quienes maximizan utilidad experimentada, o felicidad.

Comentarios finales

De las tres notas que menciona E. Roy Weintraub, ¿cuáles diferencian a *Behavioral Economics* de la Economía Neoclásica? No la maximización. Más aun, la permanencia de ese rasgo en común ha posibilitado el diálogo, mediado además por el uso de herramientas matemáticas convencionales que pululan en varios modelos *behavioral*. Con la maximización ha entrado también la racionalidad, al menos un el sentido en el que habitualmente la entienden los economistas *mainstream*.

Quizás la diferencia más profunda deba buscarse en la tercera característica definitoria: en *Behavioral Economics* las personas *no actúan*

de manera independiente sobre la base de información completa y relevante. El sujeto ha cambiado, y es ahora un individuo que deja de ser libre y autónomo para convertirse en objeto del paternalismo. Lo que ha muerto no es la maximización como herramienta teórica, o el uso de la matemática, sino el sujeto iluminista que estaba en la base de la economía liberal. Incidentalmente esto explicaría por qué razón hay tantos malos entendidos con el término racionalidad: en la tradición psicológica, la racionalidad está vinculada más a la capacidad de procesar información que a la maximización⁸. En sus propios términos –en la tradición psicológica-, claramente tanto *Prospect Theory* como *Behavioral Economics* son ambas teorías no racionales.

Bibliografía

- Arrow, K. (1986), Rationality of Self and Others in an Economic System, The Journal of Business, Vol. 59, No. 4, Part 2: The Behavioral Foundations of Economic Theory. (October), pp. S385-S399.
- Bernoulli, D. (1954), “Exposition of A New Theory on the Measurement of Risk”, Econometrica, 22(1), pp. 23 – 36 (El trabajo original fue publicado en 1738).
- Druckman, James N. (2001), “The Implications of Framing Effects for Citizen Competence”. Political Behavior, Vol. 23, No. 3, Special Issue: Citizen Competence Revisited (Sep., 2001), pp. 225-256
- Entman, Robert M. (1991), “Framing U.S. coverage of international news: contrasts in narratives of the KAL and Iran air incidents”. Journal of Communication, 41: 6-27.
- Entman, Robert M. (1993), “Framing: toward clarification of a fractured paradigm”. Journal of Communication 43: 51-58.
- Fagley, N. S., and Miller, Paul M. (1987), “The effects of decision framing on choice of risky vs. certain options”. Organizational Behavior and Human Decision Processes 39: 264-277.
- Friedman, M. y Savage, L. J., (1948), “The Utility Analysis of Choices Involving Risk”. Journal of Political Economy, LVI, pp. 279 – 304.

⁸ Ver Arrow, *op.cit.*

- _____ (1952), "The Expected utility Hypothesis and the Measurability of utility" . The Journal of Political Economy, Vol. LX, N° 6).
- Hausman, D., (1992), *The Inexact and Separate Science of Economics*. Cambridge University Press, Cambridge.
- Kahneman, D., (1999), "New Challenges to the Rationality Assumption", en Arrow, Colombatto, Perlman and Schmidt, (eds), óp. cit, pp.203-219.
- _____ (2003a), "Maps of Bounded Rationality". The American economic Review, Vol. 93, N° 5, pp.1449-1475.
- _____ (2003b), "A psychological Perspective on Economics", AEA, p. 162.
- Loomes, G. and Sugden, R., (1982), "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty". Economic Journal, 92, pp.805–824.
- Machina, M., (1982), "'Expected Utility' Analysis Without the Independence Axiom". Econometrica, Volume 50, N° 2.
- Marschack, J., (1950), "Rational Behavior, Uncertain Prospects and Measurable Utility". Econometrica, Vol. 18, N° 2, pp.111–141.
- Schoemaker, P.J., (1982), "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations", JEL.
- Starmer, Chris, (2000), "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk". Journal of Economic Literature, Vol., XXXVIII, pp.332-382.
- Stigler, G.J., (1950), "The Development of Utility Ytheory. I". The Journal of Political Economy, Vol. 58, N° 4, pp.307–327.
- Tversky, A., (1999), "Rational Theory and Constructive Choice", en Arrow, Colombatto, Perlman and Schmidt, (eds), óp. cit, pp.185-197.
- Tversky, A. and Kahneman, D., (1981), "The Framing of Decisions and the Psychology of Choice". Science, Vol. 211, N° 4481, pp.453-458.
- _____ (1986), "Rational Choice and the Framing of Decisions". The Journal of Business, Vo, 59, N° 4, pt. 2.

EL PRINCIPIO EPISTÉMICO DE CONFIANZA EN TANTO CONDICIÓN DE POSIBILIDAD DE LA DECISIÓN RACIONAL.

Pablo García y Sandra Maceri (FCE – Universidad de Buenos Aires /
CONICET)

Resumen

Este trabajo versa sobre las concepciones de racionalidad que subyacen a las diferentes maneras de analizar los actos decisionales, a saber, la racionalidad bajo certeza y la racionalidad limitada. La tesis fundamental es que ambas concepciones suponen el principio epistémico de confianza. Argumentamos a favor de nuestra formulación del principio epistémico de confianza que ampliamos respecto de su formulación original como supuesto de la teoría de la racionalidad en sus dos vertientes. Nuestra reformulación abarca no sólo al hecho de que confío en mis facultades y en la de los demás sino también el hecho de que tengo confianza gradual en las cosas en las que creo. Sostenemos que esta enunciación del principio epistémico de confianza se constituye como la condición de posibilidad tanto para la racionalidad bajo certeza como bajo incertidumbre. En efecto, tanto la racionalidad bajo certeza como bajo incertidumbre suponen el principio ampliado de confianza en el sentido en que tanto el “decisor neoclásico” como el “decisor limitado” toman decisiones porque sus capacidades epistémicas lo incluyen. De este modo, el principio epistémico de confianza se constituye como la condición de posibilidad en la toma de decisiones porque sin la presunción de este principio epistémico el acto de decisión no sería posible.

Introducción

Como es sabido, la teoría de la decisión es central en economía, y suele también denominarse teoría de la racionalidad.¹ De este modo las nociones de decisión, elección y racionalidad han llegado a constituirse en conceptos inseparables para el pensamiento económico. Con palabras de

¹ Scarano, E. (2009), p. 11.

Hausman: “La Microeconomía retrata a los agentes individuales como agentes que eligen racionalmente. Muchas de sus generalizaciones que ofrece concernientes a como la gente elige son también indicaciones sobre como los agentes deben elegir racionalmente”.²

En este trabajo nos ocuparemos de las concepciones de racionalidad que subyacen a las diferentes maneras de analizar los actos decisionales, para adentrarnos en la noción de “confianza epistémica”, que parece sostener las distintas concepciones de racionalidad. La teoría estándar de la decisión presenta dos vertientes: i) la decisión en condiciones de certeza y ii) la decisión en condiciones de riesgo, es decir, en situaciones de incertidumbre.³ Ambas vertientes presuponen el principio epistémico de confianza en tanto su condición de posibilidad y, en este sentido, a priori,⁴ aunque no por ello vacío. Sostendremos que la condición de posibilidad tanto para la racionalidad bajo certeza como para la imperfecta o racionalidad limitada es el principio epistémico de confianza, que ampliaremos respecto de la formulación original de Zagzebski⁵, el cual suele quedar en segundo plano a la hora de explicar el problema de la racionalidad en nuestro modo de decidir en tanto agentes racionales.

A continuación, argumentaremos a favor de nuestra propuesta sobre el principio epistémico de confianza como supuesto en ambas concepciones de racionalidad.

Desarrollo

Todos los días tomamos decisiones. Por lo general, esas decisiones consisten en elecciones: elegimos el medio de transporte para ir a nuestro trabajo, el libro que vamos a seguir para dar nuestras clases, el restaurante en el que vamos a almorzar, etc.⁶ Los economistas consideran

² Hausman, D. (2002), cap.1.

³ Marqués, G. (2009), pp. 173-174

⁴ Sobre teorías a priori en economía, cf. Scarano (2006).

⁵ Zagzebski, L. (2009).

⁶ Como bien sostienen Hausman, D. (2002), cap.1. y Resnik 1998 “qué es elegir racionalmente” no es una cuestión fácil de resolver.

En este trabajo nos circunscribiremos al problema de la toma de decisiones individuales sin tener en cuenta el problema de la toma de decisiones en el orden social excelentemente tratado por Resnik a través de su principio de razón

que la elección surge a partir de una conjunción de preferencias (deseos), expectativas (creencias) y restricciones

Hossein Arsham ofrece el siguiente esquema del proceso, que consta de una serie de pasos: (i) identificar la situación decisoria y comprender los objetivos que guían la decisión, (ii) identificar las alternativas posibles, (iii) descomponer y modelizar el problema, (iv) elegir la mejor alternativa, (v) llevar a cabo un análisis de sensibilidad, (vi) si se requiere un mayor análisis repetir los pasos anteriores, y si el análisis se considera suficiente implementar la alternativa elegida.

Esta secuencia es aplicable, en principio, tanto al *homo economicus* en tanto modelo teórico que pretende explicar cómo actuaría en condiciones ideales un sujeto perfectamente racional como a un agente racionalmente limitado. La toma de decisiones se relaciona con la combinación de la información sobre probabilidades y la información sobre deseos e intereses bajo incertidumbre.

Comenzaremos presentando la racionalidad neoclásica, continuaremos con la racionalidad limitada para luego centrarnos en el principio epistémico de confianza.

La racionalidad bajo certeza

La teoría de la racionalidad en condiciones de certeza⁷ asume que, en una situación de decisión, el medio, la información disponible, las creencias y

insuficiente. Según Resnik, ante una toma de decisión, sin siquiera saber qué lugar ocuparemos en el conjunto de la sociedad, cada agente puede conjeturar "la cantidad de utilidad que le correspondería si fuera una persona determinada en una sociedad concreta, puesto que se trataría simplemente de la cantidad de utilidad que estuviera al alcance de esa persona [...] Esta utilidad será sencillamente la media de las utilidades disponibles en esa sociedad. De ahí que el principio de razón insuficiente sea el que lleve al agente a seleccionar una sociedad para la cual la utilidad media sea máxima" (Resnik, 1998, p. 81). En el mismo sentido, no nos adentraremos en la teoría de la elección pública, la cual analiza "la toma de decisiones por parte del Estado suponiendo que los gobiernos y los políticos tienen una función objetivo que tratan de optimizar. Los políticos tratan de maximizar las posibilidades de ser elegidos, son maximizadores de voto". Mochón, F. y Beker, V. (1993), p. 101.

⁷ Hausman, D. (2002), cap.1.

los análisis personales son óptimos, las estimaciones de probabilidades son fácilmente realizables, el individuo tiene a su alcance información sobre todas las alternativas posibles y sin costo, y que dispone de un sistema completo y consistente de preferencias que le permite hacer un perfecto análisis de todas ellas. No se presentan, pues, dificultades ni límites en los cálculos matemáticos que debe realizar para determinar cuál es la mejor, y por lo tanto, se garantiza que la alternativa elegida es óptima.⁸ Con palabras de Hausman “El modelo básico de elección implícito en la teoría microeconómica estándar y en muchas alternativas a ésta, toma la elección o acción de un agente *A* como el resultado de sus creencias, deseos o preferencias. La elección es racional cuando está determinada por deseos y preferencias racionales”⁹. Esta asunción de la teoría de la racionalidad perfecta supone, a su vez, una racionalidad estándar.

De acuerdo con la racionalidad estándar, la microeconomía postula que los individuos son agentes que eligen de manera racional. Sin embargo, no se restringe el comportamiento racional a los agentes económicos exclusivamente, pues sería completamente extraño que los agentes se comportaran de ese modo únicamente en la esfera económica y no en las restantes. De esta manera, se puede abstraer primero una teoría de la

<http://home.ubalt.edu/ntsbarsh/Business-stat/matrix/matrix.htm>

⁸ Cf. Pascale, R., a quien tomamos especialmente en cuenta para este punto.

Pascale menciona aquí al “óptimo global” en el sentido de “la mejor solución posible a un problema. Es un ideal que en algunos problemas puede ser realmente inalcanzable en el corto plazo metódicamente sin las herramientas apropiadas.”

⁹ Hausman, D. (2002), cap.1. Además, “Se considera que los agentes tienen preferencias racionales si sus preferencias son completas y transitivas, y que eligen racionalmente si sus preferencias son racionales y no existe opción factible que los agentes prefieran más que la elegida. Las preferencias del agente *A* son completas si para toda opción *x* e *y*, *A* prefiere *x* antes que *y*, y antes que *x* o es indiferente entre *x* e *y*. Las preferencias de *A* son transitivas si para toda opción *x*, *y*, *z*, si *A* prefiere *x* a *y* e *y* a *z* entonces *A* prefiere *x* a *z*, y si *A* es indiferente entre *x* e *y*, e *y* y *z*, entonces *A* es indiferente entre *x* y *z*.” Hausman agrega el requerimiento de la independencia del contexto en la racionalidad estándar. En relación con estos requerimientos, el autor se extiende sobre las funciones de utilidad ordinal y la preferencia revelada, temas que no desarrollaremos en el marco de este trabajo.

decisión racional y luego, mediante restricciones adicionales, explicar las decisiones económicas.

Como ya dijimos, los economistas consideran que la elección surge a partir de una conjunción de preferencias (deseos), expectativas (creencias) y restricciones. Si bien las preferencias se consideran dadas y no sujetas a explicación, las elecciones realizadas a partir de un conjunto de creencias y preferencias serán, en cambio, racionales sólo si cumplen con determinadas condiciones¹⁰.

En adelante nos proponemos enfocar el problema de la elección racional y, en general, de lo que se entiende por “racionalidad”, siguiendo en gran medida el punto de vista propuesto por Popper. En efecto, de acuerdo con el principio de racionalidad de Popper, aceptaremos que, describiendo la situación que un actor enfrenta (considerando objetivos y restricciones), se asume que éste elige racionalmente.¹¹ Ante determinadas situaciones de conflicto siempre decidimos racionalmente.

En su formulación más general, el principio de racionalidad puede enunciarse como sigue: “los individuos obran siempre de un modo racional adaptado a la situación en la que se encuentran”. Según Popper, “no estamos por tanto en presencia más que de una ley de animación –el principio de una acción apropiada a la situación; por supuesto es un principio *casí vacío*. Este principio es conocido bajo el nombre de ‘principio de racionalidad’, término que ha conducido a innumerables malentendidos”.¹²

Es difícil dirimir si se trata de un principio de orden empírico-psicológico o de una consecuencia del postulado metodológico según el cual “debemos envolver o comprimir todo nuestro esfuerzo teórico, toda nuestra teoría explicativa, en el análisis de la *situación*”.¹³ En virtud de la lógica situacional,¹⁴ el principio en cuestión puede reformularse en los siguientes términos: “cuando hemos analizado nuestra situación (una situación en la que nos encontramos), sólo suponemos una cosa y ni una más para

¹⁰ Scarano, E. (2008, julio-diciembre).

¹¹ Borella, A. (2006).

¹² Popper, K. R. (1968), p.136.

¹³ *Ibid.*, p. 38.

¹⁴ Para una análisis de la lógica situacional cf. Scarano, E. (2007).

actuar acorde con ella.”¹⁵ Entonces los actores obran según este modelo de conducta (racional). La adopción del principio de racionalidad puede ser considerada, por tanto, como el subproducto de un postulado metodológico.¹⁶ Este principio no desempeña el papel de una teoría empírica explicativa ni de una hipótesis contrastable. Ciertamente el principio de racionalidad no desempeña el papel de una proposición empírica o psicológica dado que, al menos para la mayoría de los autores,¹⁷ no es empíricamente refutable. Para otros autores,¹⁸ sin embargo, este principio no es universalmente verdadero, es decir, infieren, es falso.

Sea como fuere,¹⁹ en cualquier caso, todos los autores coinciden en que la teoría de la racionalidad perfecta supone, a su vez, una racionalidad perfecta a la hora de tomar decisiones en el sentido aquí explicado.

La racionalidad limitada

Herbert Simon²⁰ fue uno de los primeros críticos de la modelización de los agentes económicos que tienen ilimitada capacidad de procesamiento de la información. En este contexto, presentó su propuesta de racionalidad acotada o limitada para describir una concepción más realista que la neoclásica o la racionalidad bajo certeza de las capacidades de resolución humana de problemas. La propuesta de Simon sugiere “emplear el concepto de procedimiento de racionalidad para distinguir entre la racionalidad de una decisión (resultados de una elección) considerada independientemente de la forma en que se la haga y la racionalidad de una decisión en cuanto a la forma en que se haga”.²¹ Desde esta perspectiva, Simon argumentó por qué en economía se debería reemplazar el principio neoclásico de optimización o maximización por un principio de satisfacción.²²

¹⁵ Borella, A. (2006).

¹⁶ *Ibidem*.

¹⁷ Popper, por ejemplo, quien trata el principio de racionalidad “como si fuera a priori”. Popper, K. R. (1968), 140.

¹⁸ Ésta será nuestra interpretación de Boland, *infra*.

¹⁹ Este trabajo no trata puntualmente sobre el *status* del principio de racionalidad.

²⁰ Simon, H. (1955).

²¹ Simon, H. (1976).

²² Simon, H. (1957).

En términos generales, podemos decir que Simon²³ sustituyó la figura del agente que decide racionalmente por el que decide con racionalidad limitada como consecuencia de los tres hechos siguientes: (i) ante determinada situación conflictiva, los seres humanos deciden, eligen, de acuerdo con una información incompleta, (ii) dada su condición de seres humanos, sólo pueden explorar un número limitado de posibilidades de acción, y (iii) también por ser humanos, son incapaces de exactitud en los resultados.

Según Simon, “las personas no somos conscientes y deliberadamente irracionales, aunque algunas veces sí lo somos, pero no poseemos ni los conocimientos ni el poder de cálculo que permita alcanzar un nivel muy alto de adaptación óptima”, por lo tanto, no existe la “perfecta racionalidad” implicada en la teoría económica clásica.²⁴

En resumen, se trata de un proceso de decisión de un individuo considerando limitaciones tanto de conocimiento (información) como de capacidad computacional. La racionalidad es limitada desde dos direcciones: desde el punto de vista del entorno “objetivo” del decisor, ya que no tiene acceso a la información completa e influyen en él factores exógenos como la cultura, las organizaciones en las que está inmerso, etc., y desde el proceso mental del decisor, pues éste no siempre tiene ni la conciencia perfecta de su estructura de preferencias, ni la capacidad completa de cálculo, y lo afectan factores como la experiencia, la memoria, la percepción, las creencias y la sensibilidad personal.²⁵

Estos factores hacen, a nuestro criterio, que seamos “tan racionales como podemos ser”, en el sentido de que somos imperfectamente racionales en la toma de decisiones. Los economistas y los teóricos de la decisión hablan de “situaciones de incertidumbre cuando las probabilidades de los

El primer párrafo de este punto es una modificación de la tesina en curso de Karina Amarilla “Avances en la crítica de la Economía Racional luego de que D. Kahneman recibiera el premio novel de Economía en el año 2002”.

²³ Por ejemplo, Simon, H. (1957).

²⁴ *Idid.*

²⁵ http://pisis.unalmed.edu.co/vieja/cursos/analisis_decisiones/ComportamientodelUuario/racionalidadLimitada/CapituloRL.doc

resultados, o incluso el rango de resultados de la acción, son desconocidos”.²⁶

Ahora bien, decidamos bajo racionalidad en condiciones de certeza o bajo incertidumbre y riesgo, influidos o no por los factores mencionados, la toma de decisiones resulta posible sólo si suponemos el principio epistémico de confianza

El principio de confianza

Sea bajo certeza o sea limitada, la facultad humana de la razón para la toma de decisiones es (o debe presuponerse) confiable. La toma de decisiones supone la confianza tanto en los otros como en mí mismo.²⁷

Para algunos autores, por ejemplo Zagzebski, la confianza en uno mismo tiene dos partes: (i) el supuesto de la fiabilidad de mis facultades (Zagzebski parece referirse acá no sólo a la razón),²⁸ y (ii) el hecho de que confío efectivamente en mis facultades.²⁹

Estas partes se observan en la enunciación del principio: “cualquier razón que tengo para confiar epistémicamente en mí mismo es una razón que se aplica a muchas otras personas. No tengo ninguna razón para pensar que soy epistémicamente privilegiado. Debo, y de hecho, confío en los demás”.³⁰

A estos dos ítems de Zagzebski, agregaremos otro sobre el cual nos extenderemos: (iii) el hecho de que tengo confianza gradual en las cosas en las que creo.

Es cierto, sin embargo, que tenemos más confianza en algunas cosas que creemos que en otras (que también creemos). El grado de confianza decrece en cuanto se nos presenta evidencia en contra (de eso que creemos). En el curso de este proceso se está cada vez menos dispuesto a afirmar de un modo ‘unqualified’ la proposición que se cree o a decir de modo ‘unqualified’ que se la cree. No parece haber de modo obvio un punto preciso en el que desaparece la creencia ‘unqualified’. Esto sugiere

²⁶ Hausman, D. (2002).

²⁷ Cf. Luhmann, N. (2005).

²⁸ Zagzebski. L. (2009), p.43.

²⁹ *Ibid*, p.44.

³⁰ *Ibid*. p.43.

que la creencia es un fenómeno que admite grados³¹ y estos grados pueden calcularse sólo porque adherimos “previamente” al principio de confianza. Es decir: una cosa es la confianza que conforma el principio epistémico de confianza planteado en este trabajo y otra la confianza en tanto una especie de actitud psicológica que podemos calcular para la toma de una decisión.

Como ya señalamos, las explicaciones racionales del comportamiento de los individuos apelan a creencias y deseos. Estas explicaciones dependen no sólo de lo deseado sino de la intensidad con que se lo desea pero también con la probabilidad otorgada a los distintos cursos de acción a mano para que conduzcan a los resultados deseados. Llevar o no un paraguas dependerá de la probabilidad otorgada a que llueva, y del grado de deseo de no mojarse. Cuanto más el sujeto desee no mojarse, o cuanta más probabilidad otorgue a la posibilidad de que llueva, más racional será llevar el paraguas.

Pero la creencia binaria podría reducirse a la graduada, o viceversa, y en ese caso las restricciones para una serían también restricciones para la otra. Además, si fueran fenómenos independientes, cada uno sería pasible de restricciones racionales diferentes. En principio, estas concepciones de la creencia parecen admitir dos modelos formales diferentes. Según Christensen³², la concepción binaria parece ser fácilmente compatible con la aplicación más o menos directa de las restricciones que impone la lógica deductiva. La idea es que el conjunto de las relaciones entre las proposiciones creídas debe, en particular, ser (1) consistente y (2) respetar la clausura deductiva. La concepción graduada, por su parte, parece sujetar la creencia a otro tipo de restricciones. De acuerdo con ella, no hay ningún conjunto de proposiciones que el individuo crea sino actitudes gradualmente distintas que el individuo tiene con respecto a las proposiciones. En un extremo, están las proposiciones (‘claims’) de cuya verdad el individuo está “cierto”, y en el otro, aquellas de cuya falsedad el individuo está “cierto”. En el medio, hay un espectro mayoritario de proposiciones respecto de cuya verdad el individuo exhibe un grado intermedio de confianza. El modelo formal estándar de grados racionales ideales de creencia apela al cálculo

³¹ La explicación de los grados de creencia está tomada de Christensen, D. (2004).

³² Christensen, D. (2004).

de probabilidades. Los grados de creencia se representan en una escala de 1 (para las proposiciones en cuya verdad se tiene certeza) a 0 (para las proposiciones en cuya falsedad se tiene certeza). Los grados de creencia de un agente racional ideal deben obedecer las leyes del cálculo de probabilidades (deben ser 'probabilísticamente coherentes').

Según Thomas Bayes debemos preguntarnos cómo tomar una mejor decisión comparando información confiable.³³ En muchos casos, el decisor puede necesitar la opinión de un especialista para reducir sus incertidumbres con respecto a la probabilidad de cada uno de los estados de la naturaleza y, recién entonces, tomar la decisión en pro de obtener la máxima utilidad.

Es cierto que en la teoría de utilidad esperada ya está cuantificada la confianza pero en el sentido de creencia, de estados de creencias de los decidores, es decir que el grado de confianza en una creencia probabilística está en la misma probabilidad. Sin embargo, no se trata de la cuantificación del principio de confianza en tanto supuesto epistémico, como condición de posibilidad de toda decisión, tal como lo presentamos en este trabajo.³⁴ Una cosa es medir la utilidad esperada, y hasta los grados de creencia en la utilidad esperada, y otra es que esta misma medición suponga un principio epistémico para llevarse a cabo. De todos modos, tanto los cálculos de la lógica de binaria de Christensen como los de las probabilidades subjetivas de Bayes no miden exactamente la utilidad esperada sino el grado de certidumbre en las cosas en las que creo para decidir cuál es la mejor decisión que puedo tomar.

En términos generales³⁵, el teorema de Bayes trata del ajuste de las probabilidades a la luz de nueva evidencia empírica, incluyendo en sí mismo grados de confianza, es cierto, pero este ajuste sólo es posible porque confío en mis propias capacidades epistémicas (primera parte del principio de confianza) o porque confío en las capacidades epistémicas de otro (segunda parte del principio de confianza), tal el ejemplo de la recurrencia a un agente externo especialista.

³³ <http://www.bayesian.org/>

³⁴ Agradecemos al Lic. Diego Weisman su comentario respecto de este punto.

³⁵ No incluimos ejemplos puntuales de los cálculos de Bayes por razones de espacio.

Esto último también es aplicable si la decisión fuera individual, sin apelación a un agente externo. Por ejemplo,³⁶ supongamos que creo que mañana va a llover con una probabilidad (subjetiva) del 70%. Tengo entonces el siguiente enunciado E1:

E1 “La probabilidad de que mañana llueva es 0.7”

Si alguien me dijera ahora cuál es el grado de confianza que tengo en mi creencia, hay dos caminos:

A- Estoy seguro de mi creencia, es decir:

E2 “La probabilidad de E1 es de 1”

B- Estoy incierto de mi creencia, en cuyo caso podría asignarle a mi incertidumbre un número probabilístico, digamos del 50%, con lo cual:

E3 “la probabilidad de E1 es de 0.5”

Suponiendo que tenga sentido E3, por propiedad de las probabilidades, se podría reducir a

E4: “La probabilidad de que mañana llueva es del 0.35”

De otra manera, habría un regreso al infinito, y nunca se podría tomar una decisión. Las probabilidades asignadas a los eventos ya tienen contemplada toda la información, es decir mis grados de confianza de segundo orden, tercer orden, y así hasta el infinito, con lo cual parece que lo único que hace falta para tomar decisiones es la probabilidad asignada al evento.

Ahora bien, el grado de confianza en una creencia es absoluto, o sea, de probabilidad 1. Varía el grado de confianza en la efectiva realización del contenido de la creencia. Por ejemplo, creo que mañana hay un 70 % de probabilidades de lluvia. En efecto, no puedo decir que mi creencia de que lloverá es de un 70% y de que no lloverá es de un 30%, mi creencia de que para mañana hay un 70% de probabilidades de lluvia es segura en un 100%. Si no fuera así, de todos modos tendría una creencia absoluta (segura) de que creo con un 70% de probabilidad que mañana lloverá. De esta manera, se evitaría la estrategia de las probabilidades combinadas,

³⁶ Ejemplo del Lic. Diego Weisman

porque se eliminaría la creencia probabilista en un nivel cero, que es el de la creencia última, evitando así el regreso infinito.

Sea como fuere, aunque el grado de confianza en una creencia probabilística estuviera en la misma probabilidad es fundamental establecer la siguiente diferencia: una cosa es el grado de confianza “medible” o “calculable” en sí misma y otra que para que ello sea posible debo adherir al supuesto de que, en el caso del ejemplo, mis facultades (y la de los demás) son fiables y también lo son las cosas que me resultan confiables. Mis facultades epistémicas, la de los otros y mi confianza en la información que manejo es, en conjunto, previa (a priori) a cualquier cálculo de probabilidad.

En síntesis, una cosa es el grado de confianza en una creencia probabilística (aún admitiendo que estuviera en la misma probabilidad) y otra la confianza como principio, (como una proposición).

Recurramos o no a la ayuda de un agente externo para la eliminación de errores en la toma de decisiones, coincidimos con Zagzebski en que no somos epistémicamente privilegiados, decidamos bajo certeza o bajo incertidumbre. Ésta es la razón por la cual parece cumplirse el principio de confianza enunciado por Zagzebski: “cualquier razón que tengo para confiar epistémicamente en mí mismo es una razón que se aplica a muchas otras personas”. Llamamos a este principio “principio restringido de confianza”, al cual completamos, en nuestra versión ampliada del principio de confianza: “cualquier razón que tengo para confiar epistémicamente en mí mismo es una razón que se aplica a muchas otras personas y a las cosas en las que creo”. Si las elecciones suponen el principio de confianza, de ser afectada esta confianza se generarán consecuencias negativas respecto de nuestra posibilidad de elegir racionalmente, de modo que sería necesario recurrir a la adivinación o la corazonada. Por esta razón, la elección debe estar respaldada por el ejercicio de capacidades epistémicas que efectivamente lleven a éxitos en la conducta elegida, es decir, una predominancia de elecciones exitosas llevará a incrementar el principio de confianza y sustentar la racionalidad en las elecciones.

El principio de racionalidad, tanto en la versión que expone la teoría económica tradicional bajo la forma del agente maximizador como en su versión atenuada propuesta por Simon, está “inseparablemente unido” al

principio de confianza en su versión ampliada en el sentido en que ambas versiones presuponen el principio epistémico de confianza en tanto su condición de posibilidad.

El principio ampliado de confianza podría entenderse entonces como un principio a priori, aunque no por ello vacío. Es decir que la condición de posibilidad tanto para la racionalidad bajo certeza como para la imperfecta o racionalidad limitada es el principio epistémico de confianza en su versión ampliada y, por ser condición de posibilidad, no es un principio empírico. Podría argumentarse que, si no es empírico, es a priori, y bien podría serlo porque, por un lado, su verdad no depende de la experiencia y, por otro, es, a la vez, necesariamente verdadero. Tampoco se trata de una tautología (es decir que tampoco es vacío en este sentido) dado que nos brinda cierta información sobre nuestras capacidades epistémicas, es decir que tiene cierto contenido empírico.

Lo importante es, a nuestro criterio, que el principio en cuestión no es un principio vacío por su posibilidad de ser a priori.³⁷ Entonces: no es un principio empírico, la evidencia empírica no puede refutarlo, y es, en algún sentido, un principio a priori. Al respecto, Barry Smith³⁸ señala que existe en los economistas austríacos un "apriorismo falibilista", una doctrina que parece explicar de manera satisfactoria la aparición de enunciados "a priori" en la teoría sin que necesariamente sean vacuos, es decir, que poseen contenido empírico. Smith sostiene que existe en la economía una "proto-disciplina", una estructura pre-teórica que presuponemos como válida y que nos permite estructurar la experiencia, pero a su vez esta estructura pre-teórica está sujeta a control empírico: la experiencia puede corregirla o modificarla. Algo similar, entendemos, ocurriría con el principio ampliado de confianza, ya que este principio podría ser puesto en cuestión y eventualmente modificado ante el caso, por ejemplo, de anomalías "recalcitrantes" que se sustraen a nuestros intentos de explicación, poniendo en crisis, en última instancia, nuestras capacidades epistémicas mismas.

En todo caso, es cierto que no se trata de un principio de orden psicológico sino de carácter epistémico, aunque tenga implicancias en la psiquis de los decisores. Tanto la racionalidad bajo certeza como la

³⁷ Sobre el apriorismo y la refutabilidad en el conocimiento económico, cf. García, P. (2003).

³⁸ Smith, B. (1999), chap. 10.

limitada son teorías de corte subjetivista, tanto el “decisor neoclásico” como el “decisor limitado” toman decisiones porque sus capacidades epistémicas incluyen el principio de confianza. Sin la presunción de este principio epistémico el acto de decisión no sería posible. El principio epistémico de confianza se constituye de este modo como la condición de posibilidad en la toma de decisiones.

Además, como tuvimos oportunidad de notar, la tesis popperiana acerca de la lógica situacional presupone que nuestras capacidades epistémicas son tales que pueden encontrar criterios adecuados para la toma de decisiones en condiciones donde la información completa resulta imposible. Pero la fundamentación de los criterios, así como la evaluación de la información obtenida, son falibles y revisables intersubjetivamente, lo cual implica que confiamos en las capacidades epistémicas de los demás, bajo ciertos criterios. Esos criterios, a su vez, son falibles y revisables y en este sentido se revisan bajo el supuesto de una racionalidad limitada. De manera que el principio de confianza no es compatible con una fundamentación última ni de los criterios para tomar decisiones ni del conocimiento sobre el cual se fundan esos criterios. Las condiciones de posibilidad del conocimiento resultan, de esta manera, tan falibles y revisables como el conocimiento mismo. Esta bien podría ser una razón de peso a favor de la conocida posición de Boland³⁹ en el sentido de que hay conocimiento científico (en el sentido de conocimiento fundado en la experiencia) y, en general, pensamiento científico (o sea, una manera de conducir nuestras capacidades epistémicas de modo de obtener conocimiento fundado aunque falible) sin método científico, aunque éste será tema de un próximo trabajo.

³⁹ Cf. Boland, L. (1994).

Bibliografía

Smith, B. (1999), *Austrian Philosophy*. Chicago: Open Court.

Boland, L. (2008), "Kuhn vs. Popper by way of Lakatos and the Cold War". [on line]Journal of Economic Methodology [on line]. Disponible en: <http://www.sfu.ca/~boland/fullerJEMessay.pdf>

_____(2003), "Recognizing knowledge: Prescription vs. Explanation", Energeia, 2, 221-28.

_____(1994), "Scientific thinking without scientific method: two views of Popper". En Roger Backhouse (ed.) *New Directions in Economic Methodology*. London: Routledge.

_____(1981), On the futility of criticizing the neoclassical maximization hypothesis, *American Economic Review*, 71, pp.1031-1036.

Borella, A. (2006), "Notas sobre el principio de racionalidad". Revista Libertas, XIII, 45.

Christensen, D. (2004), *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford University Press. (Federico Matías Pailos, Resumen del capítulo 2).

García, P. (2003), "Apriorismo y refutabilidad en el conocimiento económico: sobre una tesis de Lawrence Boland". Epistemología e Historia de la Ciencia, Córdoba (Argentina), volumen 9, número 9, pp.174-177.

Hausman, D. (2002), Racionalidad y teoría de la utilidad (cap. 1). Demanda y elección del consumidor (cap. 2). En *La Ciencia Inexacta y Separada de la Economía*. Alejandro Francetich (trad). Disponible en:

[http://www.economicasunp.edu.ar/episteme21/textos_y_revistas/Hausman Daniel-](http://www.economicasunp.edu.ar/episteme21/textos_y_revistas/HausmanDaniel-)

Luhmann, N. (2005), *Confianza*. Barcelona: Anthropos Editorial.

Marqués, G. (2007), "The Expected Utility Theory". Working Paper. FCE, UBA.

_____(2009), "¿De qué sirve maximizar la utilidad esperada?". En G. Marqués (comp.), *Racionalidad, economía e interdisciplinariedad*,

CIECE, Facultad de Ciencias Económicas, UBA, ISBN: 978-950-29-1182-3, Noviembre de 2009, pp.298.

- Pascale, R. "Racionalidad perfecta, racionalidad acotada y economía cognitiva". Disponible en:
<http://www.ccee.edu.uy/ensenian/catfinemp/material/racionalidad%20acotada.pdf>
- Popper, K. R. (1968), "La racionalidad y el status del principio de racionalidad", *Revista de Occidente*, VI, 2da. Época, 65, pp.133-146.
- _____(1975), "La ciencia normal y sus peligros". En: Lakatos y Musgrave (eds.), *Crítica y el crecimiento del conocimiento*. Barcelona: Grijalbo.
- _____(1991), Conferencia con motivo del otorgamiento del doctorado *Honoris causa*, Universidad Complutense. Madrid.
- Scarano, E.R. (1999). "Individualismo subjetivo". En: E. Scarano, G. Marques (comps.), *Epistemología de la economía*, A-Z Editora, pp.43-58.
- _____(2007), "Una perspectiva metodológica de la escuela austríaca: Las propuestas de Ludwig von Mises". En: Miroli, A. y H. Ruggeri, *Ensayos de filosofía y metodología de las ciencias económicas*. Buenos Aires: CIECE - UBA, pp.175-203.
- _____(2007), "Los vínculos entre la economía y la metodología de las ciencias sociales popperiana". En *XLII Reunión Anual, Asociación Argentina de Economía Política*, Universidad Nacional del Sur.
- _____(2008), "La teoría estándar de la racionalidad: Perspectivas metodológicas". *Economía*, XXXIII, 26, pp.63-92.
- _____(2006), "Teorías a priori en economía: criterios para su evaluación". *Endoxa*, Madrid: nº21, pp.375-394. Disponible en:
http://www.econ.uba.ar/www/institutos/epistemologia/marco_archivos/Trabajos%20Completos%20IX%20Jornadas%20en%20PDF/Scarano%20Eduardo.pdf
- _____(2009), "La evolución de la noción de valor económico y la teoría de la decisión". En G. Marqués (comp.), *Racionalidad, economía e interdisciplinariedad*, CIECE, Facultad de Ciencias Económicas, UBA, ISBN: 978-950-29-1182-3, Noviembre de 2009, pp.11-58.

- Simon, H. (1957), "A Behavioral Model of Rational Choice". En: *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: Wiley.
- _____(1976) "From Substantive to procedural rationality". En: Latsis, S. E. (ed.), *Methods and Appraisal in Economics*. UK: Cambridge University Press, pp.130.
- _____(1996), *Models of my life*. Cambridge, Massachusetts: The Mit Press, Massachusetts Institute of Technology.
- _____(1998), *The Sciences of the artificial*. Cambridge, Massachusetts. The Mit Press, Massachusetts Institute of Technology.
- Resnik, M. (1998), *Elecciones. Una introducción a la teoría de la decisión*. Barcelona: Gedisa.
- Zagzebski. L. (2009), "Confianza epistémica y conflicto epistémico". Diánoia. LIV, 62, pp.27-45.

EL DESARROLLO DE BEHAVIORAL ECONOMICS Y SU RELACIÓN CON LA TEORÍA DE LA ELECCIÓN RACIONAL

Leonardo Ivarola (CIECE – FCE – Universidad de Buenos Aires)

Bajo la concepción de la teoría estándar existe el supuesto de que las personas toman decisiones racionalmente, dejando de lado factores emocionales, pensamientos intuitivos, etc. Tampoco sus elecciones se ven alteradas por la forma en que un resultado es descrito. No obstante, investigaciones llevadas a cabo tanto por psicólogos de la conducta como por neurocientíficos sostienen que las personas no se comportan racionalmente, esto es, conforme a los axiomas de la teoría estándar, por lo cual esta última resulta ser ineficiente a la hora de funcionar como modelo descriptivo de decisión. El presente artículo analiza algunas de las tesis alternativas a la teoría de la elección racional, examinando tanto sus aspectos críticos como los cambios propuestos a la misma (en aquellos casos que sea posible). Una fundamentación neurológica y psicológica es ofrecida, a los efectos de dar cuenta del funcionamiento de los mecanismos en la toma de decisiones.

El artículo se dividirá de la siguiente manera: en la Sección I se expondrá brevemente la teoría de la elección racional, poniendo especial énfasis en sus axiomas y demás supuestos. En la Sección II se analizarán las anomalías de esta teoría a la luz de tres programas de investigación conductista: *Teoría Prospectiva*, *Efectos Marco*, y *Heurísticas y Sesgos*, examinando sus tesis principales. En la Sección III se brindarán, sobre la base de argumentos neurológicos y psicológicos, explicaciones acerca de la naturaleza del proceso decisorio de las personas, los cuales destacan la presencia casi determinante del pensamiento intuitivo. Finalmente, en la sección IV se propondrá una visión ecléctica de las posturas racional y conductista, examinando los contextos de aplicación de cada uno de los modelos ofrecidos.

La teoría de la elección racional

La *Teoría de la Elección Racional* en condiciones de certeza o *Teoría de la Utilidad* es un modelo normativo de decisión por el cual los

individuos eligen entre alternativas de decisión o cursos de acción. Se asume que las variables intervinientes en el proceso de elección son las creencias, las preferencias, y las acciones de las personas, y una elección es considerada racional cuando las preferencias son consistentes con las respectivas acciones. En particular, esta racionalidad es definida a partir de los axiomas de completitud, transitividad y continuidad. Siguiendo a Hausman (1992), las preferencias del agente A son

- Completas, si para toda opción x e y , A prefiere x antes que y , y antes que x o es indiferente entre x e y ,
- Transitivas, si para toda opción x , y , z , si A prefiere x a y e y a z , entonces A prefiere x a z ; y si A es indiferente entre x e y , e y y z , entonces A es indiferente entre x y z
- Continuas, si para cada opción x los conjuntos superior e inferior son ambos cerrados.

El postulado de racionalidad, definido a partir de los axiomas mencionados, solamente requiere que el consumidor sea capaz de clasificar los bienes en orden de preferencia. Este ordenamiento se expresa a partir de una función de utilidad que asocia números con alternativas, pero estos números solo suministran una clasificación ordinal de dichas preferencias. Por otro lado, ciertos estados mentales como impulsos, motivos o mecanismos psíquicos, son considerados irrelevantes (Friedman, 1953; Gul y Pesendorfer, 2008).

No obstante, muchas de las veces las elecciones de los agentes se dan en condiciones de incertidumbre, esto es, se toman decisiones sin tener completa información sobre las opciones relevantes (Frank, 2005). Así, se desarrolla la Teoría de la Utilidad Esperada (*Expected Utility Theory* o EUT), con el propósito de explicar esta clase de procesos decisorios. En la teoría de la utilidad esperada, las elecciones son entre alternativas con riesgo (o *prospectos*), es decir, resultados que incluyen cierta probabilidad de obtenerlos. Para que una decisión bajo riesgo sea racional, se introducen nuevos supuestos que acompañan a la transitividad, completitud y continuidad, a saber:

- Postulado de Reducción: una lotería compleja es equivalente a una lotería simple.

- Principio de independencia: si dos loterías difieren en un premio, entonces las preferencias entre las dos loterías deben ser idénticas a las preferencias entre los premios de ambas (Hausman, 1992).

Al igual que en la teoría de la utilidad, en EUT se les otorgan números a los prospectos, determinando un ranking de preferencias. Estos números son denominados valores de “utilidad esperada”, que se obtienen de la ponderación entre el grado de utilidad de un resultado y su respectiva probabilidad.

Anomalías en EUT

A diferencia de la visión tradicional en economía, el común de las decisiones de las personas no se comporta siguiendo los axiomas de la teoría de la elección racional. Específicamente hablando, numerosas investigaciones – principalmente del lado de la psicología experimental y de la neurociencia - destacan la presencia de desviaciones sistemáticas respecto de EUT. La presente sección muestra tres programas de investigación desarrollados por Tversky y Kahneman que cuestionan a la teoría de la elección racional como modelo descriptivo de decisión: *Teoría Prospectiva*, *Efectos Marco*, y *Heurísticas y Sesgos*.

TEORÍA PROSPECTIVA

Las elecciones entre prospectos con riesgo exhiben varios efectos que son inconsistentes con los principios básicos de la teoría de la utilidad. Estos efectos se ilustran sobre la base de problemas de elección tomados de Kahneman y Tversky (1979). En lo que sigue se desarrolla una crítica de la teoría de la elección racional bajo riesgo a partir de los efectos “certeza”, “reflejo” y “aislamiento”. Luego se presenta el modelo descriptivo de decisión denominado “Teoría Prospectiva”, propuesto como alternativa teórica a EUT.

Crítica a EUT

Efecto Certeza

El *efecto certeza* refiere a que las personas por lo general desprecian o sub-ponderan los resultados meramente probables, en comparación con resultados que son obtenidos con certeza. Para ilustrar esto, supongamos que una persona debe elegir entre dos prospectos A y B, tales que:

A: 80% de probabilidad de ganar \$4000

B: \$3000 con certeza.

Kahneman y Tversky (1979) notan que la mayoría de las personas se inclinan hacia la alternativa B. Pero supongamos ahora que esta misma persona debe elegir entre dos prospectos C y D, tales que

C: 20% de chance de ganar 4000

D: 25% de chance de ganar 3000.

En este último caso, Kahneman y Tversky observan que las personas se inclinan en su mayoría hacia C. Los ejemplos presentados muestran como los individuos violan la teoría de la utilidad esperada, la cual afirma que si se prefiere B a A, entonces cualquier probabilidad p añadida a B y a A no debería cambiar el orden de preferencias. No obstante, se puede observar que si se reduce la probabilidad de ganar de 1 a 0.25, el impacto será superior que si se reduce de 0.8 a 0.2. Este patrón de conducta ha sido denominado “efecto certeza”, y explica la razón por la cual las personas son aversas al riesgo en elecciones que involucran ganancias seguras y amantes al riesgo en elecciones que involucran pérdidas seguras (Kahneman y Tversky, 1979).

Efecto Reflejo

Según este efecto, de acuerdo al signo de los resultados, las personas suelen cambiar sus actitudes, siendo aversas al riesgo cuando se tratan de ganancias, y propensas al riesgo cuando se tratan de pérdidas. Por ejemplo, supongamos que en comparación con el primer

problema presentado (elegir entre el prospecto A o B), ahora se debe elegir entre dos prospectos A' y B' que involucran resultados negativos:

A': -4000 con el 80%

B': -3000 con certeza.

Kahneman y Tversky notan que la mayoría de las personas tienen una marcada preferencia por la opción A', lo que refleja como las personas pasan de ser aversas al riesgo en las ganancias (se prefiere B a A) a propensas al riesgo en las pérdidas (se prefiere A' a B'). Esta idea va en contra de los preceptos de la teoría de la utilidad esperada, que considera que la utilidad total de una lotería es la utilidad esperada de sus resultados.

Efecto Aislamiento

La teoría tradicional sostiene que los bienes y servicios son portadores de utilidad, o en otros términos, que los individuos valoran los estados finales de una dotación. En contraste, el efecto aislamiento muestra que en realidad las personas dejan de lado los componentes semejantes que comparten las alternativas, y focalizan en los elementos que los distinguen (Tversky, 1972; Kahneman y Tversky, 1979; Kahneman, 2003). Por ejemplo, supongamos que se presentan dos alternativas de decisión A y B, tales que

A: 1000 con probabilidad del 50%

B: 500 con certeza.

Supongamos que tanto para A como para B se otorgan previamente 1000, como el equivalente de una ganancia segura. A causa del ya mencionado efecto certeza, las personas elegirán en su mayoría la opción B. Comparativamente, supongamos dos nuevas alternativas C y D, tales que

C: -1000 con probabilidad del 50%

D: -500 con certeza.

Supongamos que tanto para C como para D se otorgan previamente 2000. A causa de la reversión de preferencias (o el *efecto reflejo*), las personas optarán por C. Kahneman y Tversky comprueban

satisfactoriamente tanto esta predicción como la anterior. No obstante, en términos de estados finales el prospecto A es idéntico al C, y el B al D, esto es, tanto en A como en C las loterías se componen de una probabilidad de 50% de ganar 1000 y 50% de ganar 2000, y tanto en B como en D existe una certeza de ganar 1500. Esto muestra que los individuos no valorarían los estados finales de una dotación, sino que focalizarían en las ganancias y pérdidas en relación con un punto de referencia dado.

Teoría

Sobre la base de las anomalías presentadas, un modelo alternativo de decisión bajo riesgo es propuesto por Kahneman y Tversky denominado “Prospect Theory”, el cual plantea tres grandes cambios teóricos respecto de la concepción tradicional: **a)** el concepto de utilidad, **b)** la aversión a las pérdidas, **c)** la ponderación de las probabilidades.

- a) En vez de la conocida *función de utilidad*, se incorpora una *función de valor* en la cual las personas valoran las ganancias o pérdidas monetarias en relación con un punto de referencia. Esta idea tiene su génesis en un principio psicológico de percepción y juicio, por el cual dicha percepción *depende de la referencia*: Los atributos percibidos de un estímulo focal reflejan el contraste entre dicho estímulo y un contexto de estímulos previos y concurrentes (Kahneman, 2003). Adicionalmente, la intensidad de un estímulo percibido disminuye en la medida que los valores de la variable independiente aumentan. Expresado esto en términos de ganancias–pérdidas, Kahneman y Tversky postulan que es más fácil distinguir (o valorar) una diferencia que va de una ganancia de \$100 a \$150, que una diferencia que va de \$1100 a \$1150.
- b) El concepto de *aversión a las pérdidas* implica que el *malestar* por perder un bien es superior que el bienestar por ganarlo. Tradicionalmente, se suponía que ambas utilidades debían de ser iguales en valor absoluto. No obstante, estudios empíricos mostraron que esta relación de aversión es acerca de 2-1, haciendo que la función de valor sea más empinada en el dominio de las pérdidas que en el de las ganancias (Tversky y Kahneman, 1991). Esto tiene fuertes implicaciones sobre el concepto de intercambio: en la

microeconomía *mainstream* existe un precepto básico por el cual, dados dos bienes A y B de igual utilidad, entonces cualquier intercambio entre los mismos no modificaría la situación final de los agentes. En contraste, bajo la idea de aversión a las pérdidas cualquier intercambio entre los mismos dejaría a las personas en una situación peor de la inicial. Por ejemplo, si se posee A, y posteriormente se lo intercambia por B, la des-utilidad por “perder” A siempre será superior a la de “ganar” B. En consecuencia, los individuos experimentarían cierta renuencia a ese tipo de intercambios. Esta anomalía ha sido denominada *efecto dotación*, y ha sido comprobada en numerosos experimentos (véase Kahneman, Knetsch y Thaler, 1991; Tversky y Kahneman, 1991; Kahneman, 1994). Sin embargo, un estudio más reciente llevado a cabo por Novemsky y Kahneman (2005) revela ciertas limitaciones de la aversión a las pérdidas. Particularmente, aquellos bienes que se prevén ser dados a cambio de otros (como por ejemplo pagar en efectivo) no revelan aversión a las pérdidas. Esto se debe a una cualidad de las personas por categorizar a los mismos como bienes de cambio. En contraste, los bienes de consumo sí padecen de aversión a las pérdidas.

- c) En EUT un individuo elige un prospecto racionalmente en tanto pondere la probabilidad de un resultado por la utilidad del mismo. Sin embargo, las personas no evalúan a los resultados de esa manera, sino que por el contrario transforman las probabilidades en “ponderadores de decisión”. Estos ponderadores de decisión miden el impacto de los eventos sobre la deseabilidad de los prospectos, y no meramente la probabilidad de esos eventos (Kahneman y Tversky, 1979). En particular, las probabilidades muy bajas suelen ser *sobreponderadas*, mientras que las probabilidades más altas suelen ser *subponderadas*: sea p la probabilidad de ocurrencia de un evento x , y sea $\pi(p)$ el ponderador de decisión, cuando p es relativamente baja, resulta que $\pi(p) > p$, y cuando p es relativamente alta, $\pi(p) < p$. Esta hipótesis permite no solo predecir decisiones inciertas de las personas (Kahneman y Tversky, 1992), sino también explicar los efectos certeza y reflejo mencionados arriba.

EFFECTOS MARCO

Dentro de la teoría de toma de decisiones racionales en situaciones de riesgo se utiliza un supuesto que emerge de un análisis lógico de juegos de azar: el de *extensionalidad* (Arrow, 1982) o *invarianza* (Tversky y Kahneman, 1986). Según el principio de extensionalidad, las elecciones deberían ser independientes de sus descripciones. Así, se asume que los individuos poseen la facultad de realizar una representación *canónica* de los resultados, es decir, “la invarianza captura la intuición normativa de que los cambios de formas que no afectan a los resultados actuales no deberían afectar a las elecciones” (Tversky y Kahneman, 1986). Sin embargo, experimentos muestran que las personas por lo general aceptan de manera pasiva la formulación dada. Por ejemplo, en Tversky y Kahneman (1981) se sometió a una muestra de 152 personas a la elección entre dos programas A y B para combatir una supuesta peste asiática en los estados Unidos que mataría a 600 personas. Los programas son detallados a continuación:

Programa A: se salvan 200.

Programa B: 1/3 de probabilidad de que esas 600 personas se salven, y 2/3 de que ninguna se salve.

El resultado fue un 72% a favor del programa A. Por otro lado, también se sometió a otra muestra de 155 personas a la elección entre dos programas C y D:

Programa C: mueren 400 personas

Programa D: 1/3 de probabilidad de que no muera nadie, y 2/3 de probabilidad de que todos mueran.

En esta última muestra, el 78% se inclinó por el programa D. No obstante, se puede apreciar que *extensionalmente* el programa A es idéntico al C, y el B al D. La única diferencia estriba en que en el primer caso los programas son planteados en términos de vidas salvadas, mientras que en el último caso los programas son planteados en términos de vidas perdidas. Así, ante un cambio en la descripción de resultados, las personas cambian sus preferencias, violando de esta manera el principio de extensionalidad. La invarianza no se cumple en los efectos marco (o *framing effects*), en los que descripciones equivalentes extensionalmente hablando llevan a elecciones diferentes, al modificar la notabilidad relativa de los diferentes aspectos del problema (Kahneman, 2003). Por último,

nótese que las hipótesis formuladas en prospect theory se satisfacen aquí: cuando se tratan de vidas “ganadas” (o salvadas), las personas suelen ser aversas al riesgo, pero cuando se tratan de vidas perdidas, las personas se vuelven propensas a contraer riesgos.

HEURÍSTICAS Y SESGOS

Tversky y Kahneman (1974) desarrollaron un programa de investigación dirigido a estudiar diversos juicios sobre acontecimientos inciertos. En dicho trabajo se llegó a la conclusión de que “la gente se basa en un número limitado de principios heurísticos que reducen la compleja tarea de valorar probabilidades y predecir valores a operaciones de juicio mucho más sencillas. En general, estas heurísticas son muy útiles, pero a veces llevan a errores severos y sistemáticos” (Tversky y Kahneman, 1974, p. 1124). Kahneman y Frederick (2002, 2005) han revisado estos estudios sobre la heurística del juicio y han propuesto una formulación en la que la reducción de las tareas complejas a operaciones más sencillas se logra mediante una operación de sustitución de atributos. “Los juicios se supone que están mediados por heurísticas cuando el individuo valora un atributo objetivo concreto de un objeto de juicio sustituyéndolo por otra propiedad de ese objeto –el atributo heurístico– el cual nos viene más rápidamente a la cabeza” (Kahneman y Frederick, 2002, p. 53). Un caso conocido de estas heurísticas fue desarrollado por Tversky y Kahneman (1983). En dicho estudio se sometió a una muestra de personas a un problema de elección en donde se les facilitó la descripción de una mujer y de ocho posibles resultados correspondientes a sus actividades cotidianas y a su empleo. La descripción es detallada a continuación:

Linda tiene 31 años, está soltera, es muy franca y muy brillante. Se especializó en filosofía. En su época de estudiante se preocupó muchísimo de las cuestiones relacionadas con la discriminación y la justicia social y participó también en manifestaciones antinucleares.

Las opciones relevantes fueron:

Linda es una cajera de banco (opción 6).

Linda es una cajera de banco y un miembro activo del movimiento feminista (opción 8).

La opción 8 es de tipo conjuntiva, y naturalmente la probabilidad de que ocurran dos eventos conjuntamente es inferior a que cada uno ocurra por separado. Así, la probabilidad de una conjunción $P(A \text{ y } B)$ no puede exceder la probabilidad de sus constituyentes $P(A)$ y $P(B)$, debido a que la extensión (o el conjunto de posibilidad) de la conjunción es incluida en la extensión de sus constituyentes (Tversky y Kahneman, 1983). No obstante, la mayoría de las personas suelen responder de modo inverso: en dicho estudio, el 89 por ciento de los encuestados ubicó a la opción 8 en una posición más elevada probabilísticamente que sus constituyentes. Tversky y Kahneman denominaron a dicho efecto como la “falacia de la conjunción”, y concluyeron que esta clase de juicios está mediada por la regla heurística de representatividad: se sustituye el atributo objetivo de probabilidad por el más accesible atributo de similitud o representatividad.

Sin embargo, las heurísticas del juicio pueden ser incorporadas dentro de una clase más amplia de *heurísticas del prototipo* que comparten mecanismos psicológicos modernos (Kahneman, 2003). Particularmente hablando, siempre que se piensa o se observa un conjunto (grupo, categoría) que es lo suficientemente homogéneo como para tener un prototipo, la información relacionada con este último es accesible automáticamente (Posner y Keele, 1968; Rosch y Mervis, 1975). Esta accesibilidad en los prototipos tiene consecuencias importantes en los juicios sobre conjuntos (Kahneman, 2003). Por ejemplo, se ha observado la existencia de sesgos a la hora de hacer valoraciones de bienes públicos. Desvousges *et. al.* (1993) sometieron a una muestra de personas a que expresen o indiquen su disposición a contribuir con dinero a los propósitos de prevenir el ahogamiento de aves migratorias. El número de pájaros que se salvarían variaba en las diferentes submuestras. A partir de imágenes de pájaros que se ahogaban, así como otras en las cuales estos se salvaban de ahogarse, se generó una alteración en el aparato emocional de los individuos. Esto provocó que las disposiciones a pagar no guarden correlación alguna con la cantidad de pájaros salvados. Para este caso, las disposiciones a pagar fueron \$80, \$78 y \$88, para salvar a 2.000, 20.000 y 200.000 pájaros, respectivamente. También, en Kahneman, Knetsch y Thaler (1986) se observó que los encuestados en Toronto estaban dispuestos a pagar lo mismo para limpiar los lagos de una pequeña zona de Ontario, que para limpiar todos los lagos de dicha provincia. Este sesgo ha sido denominado “olvido de la extensión” (*extension neglect*), y responde a la idea de intransitividad de las elecciones, puesto que el incremento en la extensión

de una categoría aumenta el valor de los atributos extensionales, pero no modifica los valores de sus atributos prototípicos (Kahneman, 2003). Otro sesgo relevante ha sido el de “violación de la monotonicidad”: List (2002) observó la violación de dominancia en relación con una hipotética fijación de precios. Específicamente, propuso un experimento en donde los que comerciaban con cromos de deportes le daban un valor más elevado a un conjunto de 10 cromos “sin usar – casi en perfecto estado” que a un conjunto en el que se incluían los mismos 10 cromos y tres más “en mal estado”. Esta violación de la monotonicidad también se ha observado en la evaluación de episodios prolongados: en Kahneman *et al* (1997) se hizo un experimento clínico en donde se sometía a un conjunto de pacientes a una colonoscopia, y se observó que el desagrado experimentado no guardaba relación alguna con el tiempo, sino con la intensidad del estímulo. Más aun, en algunos casos el procedimiento se extendió temporalmente, reduciendo momento a momento la intensidad del dolor. En estos últimos casos, la evaluación global fue considerablemente más agradable.

Fundamentos científicos

Investigaciones tanto de la psicología conductual como de la neurociencia afirman que la conducta de las personas es preponderantemente intuitiva en vez de racional. El presente apartado pretende mostrar los fundamentos científicos de esta tesis. Dos versiones son presentadas en el siguiente análisis: una relativa a la Psicología, y otra relativa a la Neuro-Economía.

PSICOLOGÍA

Desde la perspectiva psicológica existen dos procesos cognitivos, denominados Sistema 1 y Sistema 2 (Stanovich y West, 2000). Las operaciones del sistema 1 son **intuitivas**: se realizan de manera espontánea, sin esfuerzo, asociativas, y a menudo están cargadas emocionalmente. Las operaciones del sistema 2, en cambio, son **racionales**: se realizan más lentamente, puesto que requieren de un considerable esfuerzo mental previo. Kahneman (2003) enuncia ciertas características particulares de cada uno de los sistemas:

Sistema 1: Es rápido, paralelo, automático, sin esfuerzo, asociativo, emocional, y de aprendizaje lento.

Sistema 2: Es lento, consecutivo, controlado, con esfuerzo, regido por reglas, flexible y neutral.

Investigaciones indican que la mayor parte de las conductas suelen ser intuitivas, o provenientes del sistema 1 (Gilbert, 2002; Wilson, 2002; Epstein, 2003). Por ejemplo, un conocido experimento llevado a cabo por el profesor Shane Frederick del MIT Sloan School of Management a estudiantes universitarios consiste en resolver el siguiente problema: “Un bate y una bola cuestan en total 1,10. El bate cuesta \$1 más que la bola. ¿Cuánto cuesta la bola? Intuitivamente, el común de las personas suele responder “\$0,10”. Esta respuesta es claramente incorrecta, si se la analiza racionalmente. No obstante, más del cincuenta por ciento de los estudiantes suele con la respuesta incorrecta, lo que pone de manifiesto lo poco que se controla el producto de un pensamiento asociativo realizado sin esfuerzo (Kahneman, 2003). Sin embargo, no debe pensarse que existe una correlación directa entre intuición y respuesta incorrecta, puesto que numerosas veces, las personas adquieren ciertas habilidades que las hacen responder de manera intuitiva, esto es, sin pensar demasiado, como puede ser el manejar o jugar al ajedrez.

Una propiedad que caracteriza al pensamiento intuitivo es que vienen a la mente de manera espontánea. Higgins (1996) ha denominado “accesibilidad” a esta propiedad por la cual los procesos mentales vienen a la cabeza fácilmente. Ahora bien, algunos atributos, sean ya relacionados con la percepción o con el juicio, son más o menos accesibles que otros. Por ejemplo, es más accesible la identificación natural de un estímulo percibido, que si sobre esa percepción debe realizarse una determinada transformación. Dicha accesibilidad engloba un número suficientemente amplio de variables, *v. gr.*, pueden tener connotaciones emocionales. Particularmente, la evaluación de estímulos como buenos o malos suele hacerse de manera rápida y sin esfuerzo. Bargh (1997) realizó un experimento en donde se les presentaban a los participantes determinados estímulos, y se les dotó de una palanca para que accionen hacia adentro o hacia afuera, en caso de agrado o desagrado. Las respuestas fueron sumamente rápidas, mucho antes de que se registrara conscientemente el significado del estímulo. La accesibilidad puede darse también en destellos físicos: si se muestran

simultáneamente una carta verde grande y una azul pequeña, la verde nos vendrá a la mente primero (Kahneman, 2003).

NEUROECONOMÍA

El sistema cerebral de los seres humanos se compone de tres niveles distintos: a) el cortex o neocortex, b) el sistema límbico, y c) el sistema reptiliano, de los cuales solo los primeros dos serán analizados. En a), se dan los procesos pensantes o **racionales** de los individuos: la información proveniente de las sensaciones es elaborada en el neocortex mediante distintos circuitos cerebrales antes de iniciar una respuesta. En contraste, en b) se dan las reacciones emotivas. Particularmente, el sistema límbico hace referencia a la parte no consciente, la cual se conecta con la consciente (o cortex), transfiriéndole información a este. Todas las emociones relacionadas con la supervivencia – *v. gr.*, el hambre – se generan en este sistema, así como también todas aquellas acciones impulsivas e instintivas. Está demostrado que las acciones ordenadas desde el sistema límbico son anteriores a las del cortex. Por ejemplo, en la percepción sensorial la información proveniente de los sentidos es llevada al tálamo, y a partir de este se genera un envío de información hacia la amígdala (perteneciente al sistema límbico) y hacia la neocorteza. Según LeDoux (1999), entre el tálamo y la amígdala existe un conjunto pequeño de neuronas que hacen una conexión directa, lo cual permite una entrada de información directa desde los estímulos sensoriales percibidos. Esto hace que la información llegue antes a la amígdala que a la neocorteza. De ello, inferimos que las reacciones que se den en el sistema límbico serán anteriores a las respuestas del neocortex.

Por otro lado, el cerebro se compone de dos hemisferios: el derecho y el izquierdo, y se encuentran unidos por el cuerpo calloso. El hemisferio derecho se relaciona con el pensamiento intuitivo, y da una visión más global o abarcativa de las percepciones. Dentro de este hemisferio se encuentran neuronas con axones más largos que los del izquierdo, por lo que se conectan con neuronas que están más distantes. Esto hace que dicho hemisferio esté mejor “equipado” que el izquierdo para extraer conclusiones asociativas, debido a que cuenta con varios módulos de actividad simultánea para ello (Braidot, 2005). El hemisferio izquierdo, en cambio, se relaciona con el pensamiento lógico – racional, y da una visión más detallada del entorno. Las neuronas se encuentran densamente amontonadas, y con conexiones cortas y apretadas. Esto

hace que exista una mayor predisposición para el desarrollo de actividades que requieren de grandes esfuerzos de concentración.

Se ha notado que gran parte del comportamiento humano está derivado del funcionamiento del hemisferio derecho (Braidot, 2005). Específicamente hablando, la información proveniente de las sensaciones son infinitas, pero lo que se procesa de ella es una cantidad significativamente menor. Es decir, si bien podemos obtener gran parte de información, solo podemos procesar una pequeña parte de ella; es lo que se da en la parte consciente. Neurobiológicamente, no existe un análisis consciente de cada uno de los atributos percibidos para luego realizar un procesamiento lógico del cual se puedan extraer conductas típicamente racionales. Esto no significa que nunca se tomen decisiones basadas en el proceso de información, y que por tanto siempre existan respuestas intuitivas de las personas. Por el contrario, algunas veces – principalmente con decisiones trascendentales, tales como la compra de una casa – las conductas están basadas en un procesamiento mayor de la información (Braidot, 2005).

Hacia un eclecticismo racional – conductista

Las investigaciones analizadas arriba reflejan la necesidad de incorporar elementos de la psicología y de la neurociencia al estudio de la economía, lo cual contradice la idea de Gul y Pesendorfer (2008) de que al ser la economía una ciencia que estudia los procesos de optimización bajo restricciones, no importa el *cómo* en la toma de decisiones. No obstante, si bien las anomalías mencionadas sugieren cambios importantes en la teoría de la elección racional (tales como reemplazar el término de “cestas” por el de “cambios patrimoniales”, o el de incorporar aversión a las pérdidas), la misma puede ser “salvada” por dos argumentos: (I) relativo a la naturaleza normativa de EUT, y (II) relativo al contexto de aplicación del modelo.

En (I), la teoría de la elección racional puede verse exenta de las críticas anteriores si se la piensa como teoría *normativa*, esto es, como modelo que proponga reglas a seguir. Un ejemplo claro es el de los *costos irre recuperables*: en la teoría tradicional existe la concepción de que los costos irre recuperables no deben tenerse en cuenta a la hora de hacer un análisis económico. Un costo irre recuperable es aquel en el cual se incurre de todos modos, se tome una decisión o no. Un caso análogo (aunque

inverso) sucede con los costos de oportunidad: por lo general, las personas no lo tienen en cuenta, aunque si fueran racionales, deberían considerarlos. Claramente, un análisis normativo se encuentra exento de toda crítica que apunte hacia lo descriptivo, y la teoría estándar no sería la excepción. Analizar las condiciones que posibilitan una elección racional no implica con necesidad que las personas realmente actúen de esa manera, sino sólo que así (esto es, siguiendo las “reglas” que propone la visión tradicional) es como deberían actuar. Estos enunciados, por ser reglas, no son susceptibles de ser verdaderos o falsos. No obstante, no debe confundirse esta noción con la postura instrumentalista, debido a que no se tratan de herramientas *útiles* para predecir. De ser así, la teoría de la elección racional debería ser evaluada por su adecuación empírica, tal como lo formularía una usual corriente antirrealista. Por el contrario, sería más coherente analizar a dicha teoría como inherentes de contrastación, pero dependientes de una normatividad más exhaustiva. Así, la única posibilidad de reemplazar la teoría tradicional por otra es que la nueva proponga nuevas reglas que establezcan las condiciones de posibilidad de una elección más eficiente.

En cambio, el argumento (II) se relaciona con el hecho de que los modelos que hacen uso de la hipótesis de racionalidad pueden ser aplicados en contextos específicos. En particular, decisiones tomadas de manera rápida podrían ser predichas con alto nivel de precisión por modelos que incorporasen, por ejemplo, reglas heurísticas. Por el contrario, decisiones más importantes – como comprar un inmueble, o realizar una fuerte inversión – podrían ser predichas por el modelo de elección racional, debido a que al requerir de una elaboración mayor, estas decisiones implican procesos mentales en el área de neocortex (o sistema 2). Al respecto, Roth (1996) intenta salvar la hipótesis de racionalidad argumentando que la aplicabilidad de los modelos racionales depende del nivel de análisis requerido para el problema. Así, Roth divide los niveles de análisis en modelos económicos estándares (subdivididos en modelos de riesgo neutral y modelos de utilidad esperada) modelos psicológicos, y modelos neurobiológicos (subdivididos en modelos de interacción entre mecanismos neurales largos y modelos bioquímicos) (véase LeDoux, 2002; Chorvat y McCabe, 2005; Sanfey *et al.*, 2003).

Bibliografía

- Arrow, K. (1982), "Risk Perception in Psychology and Economics". Economic Inquiry, vol. 20, nº 1.
- Bargh, J. A. (1997): "The Automaticity of Everyday Life", en Wyer, R. S., Jr. (ed.), *The automaticity of everyday life: Advances in social cognition*, vol. 10, Erlbaum, Mahwah, pp.1-61.
- Braidot, N. (2005), *Neuromarketing*, Puerto NORTE-SUR.
- Chorvat, T. y McCabe, K. (2005), "Neuroeconomics and Rationality". Chicago-Kent Law Review, Vol. 80, p.101. Available at SSRN: <http://ssrn.com/abstract=748264>.
- Desvousges, W., Johnson, F., Dunford, R., Hudson, S, Wilson, K. y Boyle, K. (1993), "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability". en Hausman J. A. (ed.), *Contingent valuation: A critical assessment*, North Holland, Amsterdam, pp.91-164.
- Epstein, S. (2003): "Cognitive-Experiential Self Theory of Personality". En Millon, T. y Lerner, M. J. (eds.), *Comprehensive handbook of psychology: Personality and social psychology*, vol. 5, Wiley & Sons, Hoboken, pp.159-184.
- Fishburn, P. (1968), "Utility Theory". Management Science, Vol. 14, No. 5, January, pp.335- 378.
- Frank, R. (2005), *Microeconomía y Conducta*, Mc Graw Hill, 5º edición.
- Friedman, M. (1953), "The Methodology of Positive Economics". *Essays in Positive Economics*, Chicago: University of Chicago Press.
- Gilbert, D. (2002), "Inferential Correction". En: Gilovich, T.; Griffin, D. y Kahneman, D. (eds.), *Heuristics and biases: The psychology of intuitive thought*, Cambridge University Press, Nueva York, pp.167-184.
- Gul, F. y Pesendorfer, W. (2008), "The case for mindless economics". En: *The foundations of positive and normative economics: a handbook*, eds. Andrew Caplin, and Andrew Schotter, Oxford: Oxford University Press, pp.3-40.
- Hausman, D. (1992), *The Inexact and Separate Science of Economics*. Cambridge, University Press, Cambridge.

- Henderson, J. y Quandt, R. (1982), *Teoría Macroeconómica, una aproximación matemática*. Editorial Ariel, Barcelona. [1a. Ed: 1962].
- Higgins, E. (1996), "Knowledge Activation: Accessibility, Applicability, and Saliency.", En: Higgins, E. T. y Kruglanski, A. W. (eds.), *Social psychology: Handbook of basic principles*, Guilford Press, Nueva York, pp.133-168.
- Kahneman, D. (1994), "New Challenges to the Rationality Assumption". Journal of Institutional and Theoretical Economics, CL, pp.18-36.
- _____(2003), "Maps of Bounded Rationality". The American Economic Review, Vol. 93, Nº 5, Dec., pp.1449–1475.
- Kahneman, D. y Frederick, S. (2002), "Representativeness Revisited: Attribute Substitution in Intuitive Judgment", en Gilovich, T.; Griffin, D. y Kahneman, D. (eds.), *Heuristics and biases: The psychology of intuitive thought*, Cambridge University Press, Nueva York, pp. 49-81.
- _____(2005), "A Model of Heuristic Judgment". En: K. J. Holyoak & R. G. Morrison [eds.], *The Cambridge Handbook of Thinking and Reasoning*, Cambridge University Press, pp.267-293.
- Kahneman, D. and Riepe, M. (1998), "Aspects of Investor Psychology – Beliefs, preferences and biases investment advisors should know about". Journal of Portfolio Management, v. 24, nº4, Summer.
- Kahneman D. and Thaler, R. (1991), "Economic Analysis and the Psychology of Utility: Applications to Compensation Policy". AER, v.81, nº 2, May, pp.341-346.
- Kahneman, D. and Tversky, A. (1979) "Prospect Theory: An Analysis of Decision Under Risk". Econometrica, v.47, nº 2 , pp.263-292.
- _____(1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty". *Journal of Risk and Uncertainty*, 5, October, pp.297-324.

- Kahneman, D., Knetsch, J. and Thaler, R. (1986), "Fairness as a Constraint on Profitseeking: Entitlements in the Market". American Economic Review, vol. 76, nº 4, september, pp.728-741.
- _____(1991), "The Endowment Effect, Loss Aversion, and Status Quo Bias". Journal of Economic Perspectives, Vol. 5, N°1, pp.193-206.
- Kahneman, D., Wakker, P. and Sarin, R. (1997), "Back to Bentham? Explorations of Experienced Utility". Quarterly Journal of Economics, vol. 112, nº 2, may, pp.375-405.
- Knetsch, J. (1989), "The Endowment Effect and Evidence of Nonreversible Indifference Curves". American Economic Review, American Economic Association, vol. 79(5), pp.1277-84.
- LeDoux J. (1999), *El cerebro emocional*. Barcelona: Ariel-Planeta.
- _____(2002) *Synaptic Self: How Our Brains Become Who We Are*. New York, Viking.
- List, J. (2002): "Preference Reversals of a Different Kind: The 'More Is Less' Phenomenon". American Economic Review, vol. 92, nº 5, pp.1636-1643.
- Novemsky, N. y Kahneman, D. (2005), "The Boundaries of Loss Aversion". Journal of Marketing Research, Vol. XLII, 119-128.
- Posner, M. y Keele, S. (1968): "On the Genesis of Abstract Ideas". Journal of Experimental Psychology, Pt. 1, vol. 77, nº 3, pp.353-363.
- Rosch, E. y Mervis, C. (1975): "Family Resemblances: Studies in the Internal Structure of Categories". Cognitive Psychology, vol. 7, nº 4, pp.573-605.
- Roth. A. (1996), "Adaptive Behaviour and Strategic Rationality: Evidence from the Laboratory and the Field", in Kenneth J. Arrow et al., *The Rational Foundations Of Economic Behaviour*.
- Sanfey, A., Rilling, J., Aronson, J., Nystrom, L., and Jonathan D. Cohen, J. (2003), "The Neural Basis of Economic Decision-Making in the Ultimatum Game", Science, Vol. 300, no. 5626, pp.1755-1758.

- Stanovich, K. y West, R. (2000): "Individual Differences in Reasoning: Implications for the Rationality Debate?". Behavioral and Brain Sciences, vol. 23, n° 5, octubre, pp.645-665.
- _____ (2002): "Individual Differences in Reasoning: Implications for the Rationality Debate?". En: Gilovich, T.; Griffin, D. y Kahneman, D. (eds.), *Heuristics and biases: The psychology of intuitive thought*, Cambridge University Press, Nueva York, pp.421-440.
- Tversky, A. (1972), "Elimination by Aspects: A Theory of Choice". Psychological Review, 79, pp.281-299.
- _____ (1999), "Rational Theory and Constructive Choice". K. Arrow, et. all., eds, op. cit.
- Tversky, A. and Kahneman, D. (1974), "Judgment under Uncertainty: Heuristics and biases". *Science*, 185, 1124-1 131.
- _____ (1981): "The Framing of Decisions and the Psychology of Choice". Science, vol. 211, n° 4481, pp.453-458.
- _____ (1983), "Extensional versus intuitive reasoning: The conjunction fallacy in probability Judgment". Psychological Review, Vol. 90(4), 983, pp.293-315.
- _____ (1986), "Rational Choice and the Framing of Decisions". Journal of Business, vol. 59, n° 4, pp.251-278.
- _____ (1991), "Loss Aversion in Riskless Choice: A Reference- Dependent Model," The Quarterly Journal of Economics, MIT Press, vol. 106(4), pp.1039-1061.
- Wilson, T. (2002): *Strangers to ourselves: Discovering the adaptive unconscious*, Harvard University Press, Cambridge.

UN REALISMO *EMPÍRICO* EPISTEMOLÓGICAMENTE ADECUADO PARA *BEHAVIORAL ECONOMICS*

Gustavo Marqués (CIECE – FCE – Universidad de Buenos Aires)

Introducción

Muchos -quizás la mayoría- de los autores encolumnados en la corriente de *Behavioral Economics* (BE), califican su manera de modelar como *realista* y la contraponen a la manera convencional de modelar. He aquí algunas de las principales tesis que aquellos reivindican y que justificarían su pretensión de contar con modelos más adecuados que los convencionales (todas las citas provienen de Camerer y Loewenstein, 2002, draft).

1) “We share the positivist view that the ultimate test of a theory is the accuracy of its predictions” (p. 2).

2) “Loss aversion is more *realistic* than the standard continuous, concave, utility function over wealth, as demonstrated by hundreds of experiments” (p.3)

3) “Behavioral economics increases the explanatory power of economics by providing it with more realistic psychological foundations” (p.1)

4) “But we also believe that, *ceteris paribus*, better predictions are likely to result from theories with more *realistic* assumptions” (p. 2)

(1) afirma que las predicciones cuentan en la evaluación empírica de una teoría. Esto es un punto no controversial, al que la mayoría de los economistas adhiere sin reservas. Es importante mencionarlo porque aunque la mayoría de los economistas que practican BE (los llamaremos economistas BE, para abreviar) típicamente contraponen su visión metodológica y su práctica de modelar a las tesis expresadas por Friedman (1953), no difieren con él en este punto en particular.

La tesis (2) clarifica el sentido en que los economistas BE reclaman mayor realismo para sus modelos: se distinguen de los modelos convencionales

en que incorporan factores explicativos que han sido identificados y corroborados experimentalmente. Las tesis (3) y (4) se fundan en la tesis (2). La primera sostiene que sustituyendo en los modelos convencionales preferencias y expectativas que los experimentos (y el conocimiento básico disponible a la fecha) han mostrado que los individuos *no* tienen, por características más realistas (en el sentido expresado en (2)), su capacidad *explicativa* debe aumentar. La segunda sostiene que *ceteris paribus*, la incorporación de características conductuales bien corroboradas, como es el caso de *loss aversion*, al modelar el comportamiento de elección de los individuos, aumenta la probabilidad de obtener *mejores predicciones*.

Este trabajo combina aspectos descriptivos y normativos. De una parte, se ofrece una reconstrucción de la epistemología implícita en la práctica de modelar de *Behavioral Economics*. De otra parte, se intenta brindar un fundamento epistemológico a la reivindicación de muchos economistas conductistas, quienes señalan que los modelos de *Behavioral Economics* son mejores que los convencionales, en sentido de que son realistas o más realistas que éstos. Se clarificará en qué consiste el tipo de realismo reclamado por BE y se mostrará que es un realismo de carácter empírico, fundado en evidencia experimental. Esto permitirá entender mejor el nexo que los economistas BE establecen entre realismo y explicación, y se mostrará que los modelos BE son una clase particular de modelos de posibilidad (“possibility models”), y que es debido a esta particularidad que resultan genuinamente explicativos. Asimismo, se analizará la adecuación de las tesis (3) y (4), que afirman la existencia de un vínculo entre mayor realismo, por un lado, y mayor capacidad explicativa y predictiva, por el otro. Sobre la base del papel que asignamos a los modelos, sostendremos que de ambas sólo la tesis (3) puede ser defendida.

La caja de herramientas de *Behavioral Economics*. Factores causales

Resultados alcanzados en las últimas décadas en el área de la psicología cognitiva y la psicología experimental muestran que los individuos poseen un repertorio de *propensiones* $P = (p_1, p_2, \dots, p_n)$ y de *heurísticas* $H = (h_1, h_2, \dots, h_k)$. Estas son utilizadas en circunstancias particulares bien establecidas, y han sido descubiertas mediante una extensa gama de estudios y experimentos. Entre los elementos de P se encuentran “loss

aversion”, “miopía”, “adaptación”, “saliencia”, “focus illusion” y “mental accounts”. Y entre los elementos de H, la regla default, “fifty – fifty” y “regla 1/n”. Un resultado sobresaliente que arroja la psicología cognitiva es que cuando un sujeto se enfrenta a una situación en que su riqueza cambia respecto de un estado dado, la utilidad (en valores absolutos) que obtiene de este cambio es mayor cuando el cambio representa una pérdida que cuando representa una ganancia. Sea Δx el cambio referido. Lo que la psicología cognitiva ha mostrado es que $|U(-\Delta x)| > |U(\Delta x)|$. Este fenómeno, que refiere meramente al (diverso) impacto emocional que generan la ganancia o la pérdida de un bien (que puede ser dinero), dado un estado inicial de riqueza, es llamado *loss aversion*. En principio, nada hace pensar que esto esté relacionado con la teoría de la decisión (un sujeto que manifiesta *loss aversion* puede no estar adoptando decisiones). Sin embargo, se ha mostrado que *loss aversion* resulta influyente de maneras diferentes en diferentes tipos de decisiones, generando patrones de comportamiento característicos. Dos de ellos sobresalen por su importancia: el *endowment effect* y la diferente propensión al riesgo manifestada ante descripciones alternativas de una misma situación de elección.

El primer patrón surge en condiciones de intercambio. Sujetos que padecen *loss aversion* valorarán más un bien cuando lo poseen que cuando no lo poseen. Este fenómeno ha sido comprobado experimentalmente en numerosas ocasiones y en muy diversas circunstancias (Knetch, 1989; Kahneman et.al., 1991). La denominación del mismo como *endowment effect* se debe a Thaler, 1980. Por otra parte, Tversky and Kahneman (1981, 1986) han descubierto experimentalmente que *loss aversion* es una de las causas de la reversión de preferencias. Esto ocurre cuando una misma situación de elección es descrita en términos de ganancias o pérdidas. Este descubrimiento representa una extensión del dominio de aplicación de *loss aversion*: se observa que en ciertas condiciones influye sobre la actitud ante el riesgo.

Un elemento adicional importante de las investigaciones psicológicas que estamos refiriendo es que no se limitan a descubrir la presencia del fenómeno de *loss aversion*, sino que han proporcionado una *medida cuantitativa* del mismo (si bien abarca un cierto rango de valores y puede ser relativa a la comunidad humana en que se la mide). A esta se la conoce como *coeficiente de loss aversion*, que es definido de esta manera:

CLA: $|U(-\Delta x)| / |U(\Delta x)|$.

Se ha observado que CLA oscila entre 2,25 y 2,50, aproximadamente.

Papel de los Modelos en *Behavioral Economics*

Aunque la teoría psicológica proporciona un repertorio amplio de propensiones y heurísticas, su empleo con propósitos teóricos en economía no es obvio ni inmediato. El conocido experimento de la Gripe Asiática – el hecho de que los individuos manifiesten aversión al riesgo cuando una cierta situación de elección les es descrita en términos positivos (vidas salvadas) y, por el contrario, se muestren propensos a correr riesgos cuando la misma situación les es descrita en términos negativos (vidas perdidas) –, concierne inicialmente al dominio de la teoría pura de la decisión y por sí mismo no sugiere aplicaciones al campo de la teoría económica propiamente dicha. En principio, no puede decidirse de antemano si, por ejemplo, guarda alguna relación con reconocidos patrones de ahorro o inversión. En segundo lugar, las circunstancias en que las diferentes propensiones o heurísticas emergen o predominan pueden ser muy diferentes entre sí. Por eso no todas ellas están actuando en todo momento y en toda circunstancia. Además, una cierta propensión o heurística pueden dar lugar a un patrón T en circunstancias C y a un patrón T' diferente en circunstancias C'. ¿Cuándo cabe esperar uno u otro resultado? Para dilucidar estas cuestiones se necesita un modelo.

La tarea del modelo es explorar qué propensión o heurística específica es causalmente relevante, dadas determinadas circunstancias, en relación a un patrón de comportamiento económico específico p . La naturaleza y función de los modelos BE se clarifica cuando su producción es descrita como una secuencia en dos etapas. En la primera, los modeladores cuentan con un arsenal de hipótesis bien testadas acerca de propensiones y heurísticas, así como con el conocimiento de los efectos que éstas producen en circunstancias experimentales bien definidas. Por ejemplo, como ya señalamos, se sabe que sujetos que poseen *loss aversion*, cuando enfrentan una situación de intercambio, se comportan de una manera regular que es descrita como *endowment effect*: valúan más un bien si lo poseen que si no lo poseen. En un segundo momento, los modeladores se proponen resolver con estos recursos un problema particular, que trasciende los resultados ya alcanzados: explicar la

formación de un cierto patrón observado p de conducta económica, el cual es conocido independientemente y nada hace presumir que posee vinculación con alguna de las propensiones y heurísticas conocidas. Para ello construyen un modelo, que consiste en dar forma a la conjetura de que una propensión o heurística particular (o un conjunto de ellas), digamos xi , explica el patrón p . Por ejemplo, Benartzi y Thaler (1995) muestran que el *Equity Premium Puzzle* (EPP) -un cierto patrón de distribución del ingreso entre bonos y acciones, que resulta anómalo desde la perspectiva convencional- se explica asumiendo que los individuos poseen *loss aversion* y padecen de *miopía* (es decir, reciben información frecuente acerca de los rendimientos futuros de cada tipo de inversión).

EPP configura un patrón económico de naturaleza completamente diferente al *endowment effect*. Aquél no se produce en el marco de una situación de intercambio: no resulta de que los propietarios de acciones (bonos) les ponen a éstos un precio mayor cuando los poseen que cuando desean comprarlos. EPP refleja el hecho de que los inversores (a largo plazo) no prefieren invertir en acciones más que en bonos, pese a que el rendimiento de los primeros es claramente superior. En este caso, no existe *endowment effect* por dos razones. En primer lugar porque se supone que tal cosa no se produce respecto de la posesión de dinero. Y, además, porque, aunque ello ocurriera, sería insustancial en este caso, porque lo que se describe es una asignación anómala de recursos entre dos tipos de bienes alternativos, igualmente *no* poseídos por los individuos. EPP surge en un escenario en que los sujetos no oscilan entre los roles de comprador y vendedor, sino que son siempre compradores.

Aprovechando los resultados alcanzados, podemos precisar mejor el papel de los modelos BE. Digamos que xi es una propensión o heurística, acreditada empíricamente, y p un patrón de conducta económica conocido pero que carece todavía de vinculación probada con xi . Lo que se propone un modelo BE es confirmar la hipótesis de que xi puede ser relevante para explicar p (o que xi puede ser parte de la causa de p). Un modelo teóricamente exitoso logra mostrar que, bajo circunstancias C, existe una *conexión* entre xi y p . Análogamente, un modelo fracasa cuando no logra establecer este vínculo. Dicho modelo no es un test de la existencia de xi , porque ésta ya ha sido testada y confirmada con anterioridad. En consecuencia, un modelo exitoso (no exitoso), que procura explicar p a partir de xi , *no* es una confirmación (refutación) de xi .

Volviendo al modelo propuesto por Benartzi y Thaler (1995), si éste fracasara en su propósito, ello no afectaría en nada el hecho comprobado independientemente de que los sujetos padecen *loss aversion*. El fracaso mostraría únicamente que, aún cuando hubiera una conexión entre esta propensión y EPP, el modelo no ha tenido éxito en su pretensión de mostrarla. Es importante aclarar que por “fracaso” del modelo no debe entenderse aquí que sus predicciones (o consecuencias empíricas) resultan contradichas por los hechos. Esto puede no ocurrir porque, por construcción, muchos modelos BE intentan dar cuenta de hechos (patrones) ya conocidos. Lo que el modelo aporta, en sentido estricto, es una *conexión* entre propensiones y heurísticas, por un lado, y patrones observados, por el otro. El modelo fracasa cuando fracasa en establecer esta conexión.

Despistados, quizás, por una concepción excesivamente logicista, que ignora o subestima la importancia de los aspectos pragmáticos presentes en la construcción de los modelos, muchos autores sostienen una interpretación de su papel muy diferente al que acabamos de presentar. Ello ha oscurecido en buena medida la clásica discusión en torno al realismo de los supuestos. En su debate con los realistas de su época, Friedman concluye que la única manera de determinar si un enunciado componente del modelo (se trate de una hipótesis o un “supuesto”) es o no “realista” (es decir, empíricamente adecuado) es examinar las predicciones del modelo en que participa.

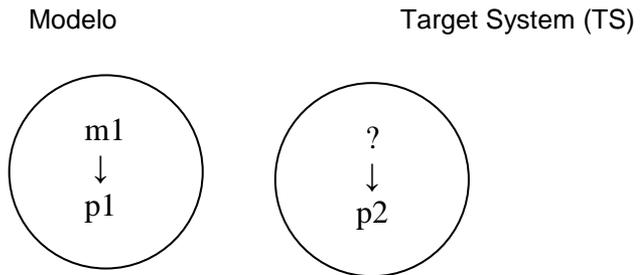
En mi opinión, su argumentación pasa por alto la importante distinción entre dos usos bien diferentes de una hipótesis o teoría. Aún admitiendo a rajatabla su tesis de que “...the only relevant test of the *validity* of a hypothesis is comparison of its predictions with experience” (Friedman, 1953, pp. 8-9), uno puede rechazar la idea de que todos los portadores de verdad de un modelo se someten a test en el único y mismo acto de obtener predicciones del modelo y compararlas con los patrones observables. Como hemos visto, mecanismos como *loss aversion* u otras propensiones y heurísticas son testados experimentalmente antes (e independientemente) de su inclusión en modelos de *Behavioral Economics*.

Es importante pues no confundir dos contextos argumentativos diferentes: aquel en que una hipótesis es usada con propósitos de *test* y aquel en que se la emplea a los efectos de brindar una *explicación*. Aunque sea cierto que todo test de un enunciado se realiza únicamente examinando sus consecuencias (predicciones), no es cierto que toda vez que se emplea un enunciado con propósitos argumentativos se esté intentando testarlo. Este aspecto pragmático pasó desapercibido en la argumentación de Friedman, quien, al parecer, creyó que *toda* premisa que ocurra en un modelo es sometida a test por las predicciones que se obtienen del mismo. Ello no es así. Del mismo modo en que no todo enunciado que forme parte de un modelo afirma lo que literalmente parece afirmar (Musgrave, 1981), no todo enunciado que efectivamente afirma lo que parece es incorporado a un modelo con el propósito de ser testado. Como hemos sugerido, puede incluírsele para averiguar si mantiene alguna vinculación con ciertos fenómenos que el modelo desea explicar. Este es el papel que desempeñan hipótesis como *loss aversion* o *mental account* en los modelos de *Behavioral Economics*.

Para ilustrar la diferencia entre testar una hipótesis y emplearla con propósitos explicativos, recordemos que diversos experimentos han permitido obtener una medida del coeficiente de *loss aversion* (CLA), que establecen entre 2,25 y 2,50. Esta medida es empleada luego en diferentes modelos para explicar y predecir fenómenos a primera vista no vinculados con tal magnitud. Dicho coeficiente no es arbitrario ni modificable a voluntad por quien lo incluye en un modelo. Para diferenciarse, como *Behavioral Economics* desea hacerlo, de la manera *mainstream* (irrealista) de modelar, no podría postular *loss aversion* al modelar las preferencias de los individuos y asignar a esta propensión un valor cualquiera, que le resulte conveniente (es decir, no sería “realista” proceder de esta manera). El valor de *loss aversion* es establecido previa e independiente a su inclusión en el modelo. Funciona como el análogo de una constante física, que impone una rigidez a las fórmulas en que participa. Por ello, dicho coeficiente no resulta afectado por el éxito o fracaso del modelo mismo.

Los modelos de *Behavioral Economics* como modelos de posibilidad realistas

Que los componentes de **P** y **H** incorporados a los modelos BE sean reacciones humanas cuya existencia ha sido determinada previamente en situación de control experimental es importante para distinguir el proyecto BE del de la economía convencional, en que no se presume contrapartida real a las variables explicativas. La conocida exposición de Friedman ilustra este punto. La teoría de la firma, argumenta, no sostiene que los empresarios realizan lo que sus axiomas describen literalmente o implican (por ejemplo, no afirma que aquellos comparan las curvas de ingreso y costo marginal para decidir su nivel de producción). Los empresarios sólo se comportan *como si* lo hicieran. Lo mismo puede decirse de *Expected Utility Theory* (EUT): esta teoría no se compromete con que los agentes, de hecho, poseen preferencias completas, transitivas, etc. Más bien afirma que los agentes exitosos al final del día de mercado se comportan *como si* tuvieran esas preferencias. Llamaremos “estrategia ‘como si’” a esta manera de interpretar el contenido de las teorías o modelos. Algunos economistas BE rechazan *todo* uso de la estrategia “como si” por considerarlo ilegítimo (“irrealista”). Sin embargo, esta crítica es incorrecta, ya que esta misma estrategia es empleada en los modelos BE. Para mostrarlo consideremos el siguiente gráfico:



Supongamos que observamos en TS el patrón p2, presumiblemente originado por algún mecanismo desconocido. Para explicarlo construimos un modelo, como el representado en la parte izquierda del gráfico. Allí, a partir de aislar un mecanismo m1 obtenemos un patrón p1 (igual a, o compatible con, p2). Un usuario del modelo puede entonces

legítimamente afirmar que los individuos “reales” de TS se comportan *como si* sus decisiones estuvieran originadas en el mecanismo m_1 (independientemente de cuán probable resulte que “realmente” lo estén). Puede que m_1 esté efectivamente actuando en TS, pero el usuario del modelo no necesita asumir este compromiso para que el empleo de la expresión “como si” esté legitimado, ya que en un contexto semejante la expresión sólo quiere decir que asumiendo m_1 es posible deducir en el modelo patrones análogos a los observados (lo cual es estrictamente cierto en este caso).

El rechazo a-crítico de la estrategia “como si” es erróneo por una segunda razón, aún más importante que la primera. Esta estrategia se adecua perfectamente a lo que se conoce como *modelos de posibilidad* (“possibility models”): modelos que muestran qué hace (o haría) *posible* ciertos resultados observados. Y enseguida se advierte que tanto los modelos convencionales como los modelos BE son modelos de posibilidad. Ambos muestran qué mecanismos harían posible (podrían generar) un patrón económico dado. En el caso ilustrado anteriormente, Benartzi and Thaler (1995) muestran que los inversores se comportan como si padecieran miopía y *loss aversion* respecto de los retornos que provienen de sus inversiones en bonos y acciones.

Sin embargo, los modelos convencionales y los modelos BE son dos tipos muy diferentes de modelo de posibilidad. Esto se torna evidente cuando se explicitan las restricciones que en cada caso se imponen a la solución del problema que abordan. Los modelos convencionales son diseñados para mostrar qué mecanismo *posible* haría posible un patrón dado p . Los modelos BE, en cambio, afrontan una restricción adicional: tratan de mostrar qué mecanismo, *compatible con el conocimiento científico disponible* (específicamente, con aquel conocimiento aportado por la *psicología experimental*), haría posible p . Esta interpretación convierte a los modelos BE en una subclase de la clase más amplia de modelos de posibilidad.

Para formular esta idea de una manera algo más formal, supongamos que Y es la clase de todos los mecanismos posibles y X la subclase de Y que contiene a todos aquellos mecanismos conductuales descubiertos en sede experimental. Dado un patrón económico p , que demanda explicación, el economista convencional puede modelarlo apelando a *cualquier* mecanismo (quizás sea más exacto decir “cualquier mecanismo

perteneciente al conjunto diferencia $Y - X$). Su única restricción es que p pueda ser deducido a partir de algún mecanismo *concebible*. El modelador BE, en cambio, se auto impone condiciones más exigentes para obtener p : debe hacerlo a partir de alguno de los elementos de X (y debe construir un modelo para probar esta conexión). Esto explica la naturaleza *realista* de los modelos BE. Que sean también modelos de *posibilidad* se ve en lo siguiente. Sea x el elemento seleccionado en X . El modelo que conecte exitosamente x con p no prueba que este elemento está efectivamente causando p en esta circunstancia particular. Lo único que muestra es que x *podría* ser la causa (o una de las causas) de p . Esto es inmediatamente traducible a los términos de Friedman: el modelo muestra que los agentes involucrados se comportan *como si* sus decisiones fueran causadas por x . La diferencia entre este uso de la estrategia “como si” y el prevaleciente en los modelos convencionales es que dicha estrategia es empleada ahora invocando un dominio de mecanismos causales más restringido, fundado empíricamente y compatible con los resultados del conocimiento científico. Este es el uso que los economistas BE aprecian como *realista* y propiamente *explicativo*.

Conclusiones

Los argumentos que hemos expuesto anteriormente acerca del tipo de factores e hipótesis que son usados en los modelos BE, su potencialidad explicativa y el papel que estos modelos tienen en su estrategia argumental, nos permitieron ofrecer una reconstrucción de la concepción *realista* que se halla implícita en la práctica de modelar de BE, la cual toma en cuenta de manera esencial el papel de la evidencia empírica. En nuestra visión, su práctica modeladora aprovecha la disponibilidad de tipos de conducta (propensiones o heurísticas) descubiertos en sede experimental. Estos recursos son luego incorporados en modelos particulares, diseñados con el propósito de determinar si existe una conexión entre alguno de estos tipos de conducta y un conjunto de fenómenos económicos acerca de los cuales inicialmente no cabía presumir que tuvieran alguna relación con ellos.

Behavioral Economics asigna a dichos factores un genuino papel causal. En principio, *cualquier* propensión o heurística confirmada (es decir, que se ha mostrado experimentalmente que los individuos poseen o usan), podría estar involucrada en la generación de un cierto patrón de conducta

de interés teórico para la economía. Se necesita un modelo para probar que dicha conexión puede ser establecida. El modelador BE goza pues de una gran libertad para buscar una solución a su problema, pero ésta no es ilimitada: puede elegir *cualquier configuración de propensiones o heurísticas de entre las corroboradas*, pero *no* postular cualquier cosa. En particular, los economistas conductistas no intentan explicar un patrón de comportamiento económico **p** mediante propensiones o heurísticas *disconfirmadas o antojadizas*. Tampoco modelando las preferencias y expectativas de los agentes de cualquier manera que resulte matemáticamente conveniente. No confunden un ejercicio deductivo exitoso con una explicación. Por ello, si se sabe (o se cree) que hay ciertas propensiones o capacidades que los *homo sapiens no* tienen, no se apela a ellas como recurso explicativo.

Una característica importante de los modelos BE es que son *explicativos*, en un sentido no meramente deductivo y diferente a que posean capacidad predictiva. Como sugiere la tesis (3), *Behavioral Economics* asocia esta capacidad con su mayor realismo. Pero esta pretensión es exagerada. El *realismo* de BE está en la base de su aptitud explicativa, pero no la garantiza, porque los factores que incorporan los modelos BE *no son, por sí mismos, suficientes* para dotarlos de capacidad explicativa, ya que nada asegura que tal o cual propensión o heurística particular, por bien comprobada que esté experimentalmente, pueda ser relacionada con un cierto patrón económico que se desea explicar. Para que la fuerza explicativa del modelo se concrete se requiere que el modelo tenga éxito en conectar el mecanismo seleccionado con el patrón observado **p**. Aún así, este éxito sólo muestra *qué factor o factores harían posible* la ocurrencia de **p**. Para que esta *fuerza explicativa potencial* se concrete se necesita de información empírica adicional que muestre que ese mecanismo incorporado en el modelo el que efectivamente ha generado **p**.

La tesis (4), que considera al mayor realismo de las hipótesis de los modelos BE como condición *suficiente* de su mayor capacidad predictiva en relación a los modelos convencionales, tampoco parece fácilmente defendible. Incorporar a los modelos propensiones o heurísticas comprobadas, no garantiza, de por sí, mejores predicciones, porque puede que se seleccione un factor que no es relevante para la ocurrencia del tipo de conducta de la que se desea dar cuenta (o no lo es en las

condiciones estudiadas). También es controversial una interpretación más débil de (4), según la cual sólo se pretendería que la incorporación de dichas propensiones y heurísticas es *necesaria* para dotar a los modelos BE de mayor capacidad predictiva que la que poseen los modelos convencionales. En la medida en que los modelos BE resultan exitosos para dar cuenta de las anomalías que afectan a los modelos convencionales podría sostenerse que predicen mejor este tipo de fenómenos (o, más prudentemente, que la manera BE de modelar parece ser promisoria para el objetivo de predecir mejor este tipo de patrones). Sin embargo, si solo se trata de predecir (no de explicar) el ingenio matemático convencional puede arreglárselas para corregir las anomalías. Por último, la necesidad de incorporar a los modelos económicos los recursos aportados por la psicología cognitiva no puede ser defendida en general, ya que hay otras hipótesis conductuales, igualmente realistas, que son distintas a las específicamente empleadas por BE (por ejemplo, de tipo ético, como “reciprocidad”) que podrían tener efectos causales sobre algunos fenómenos económicos y han sido incorporadas con éxito a otros tipos de modelos no convencionales.

En la interpretación de los modelos BE que hemos ofrecido en este trabajo desempeña un papel importante la distinción entre experimentos y modelos. Se trata de dos contextos argumentativos diferentes. Tal tipo de distinciones resultan frecuentes en metodología. Por ejemplo, la que se realiza entre premisas y conclusiones o hipótesis y datos. Un mismo enunciado puede ser tratado como una hipótesis (premisa) o un dato (conclusión), según el contexto en que se lo use. Debido probablemente a que intenta identificar modelos y experimentos, Mäki (2005) pasa por alto el papel de estos últimos en la identificación de mecanismos cognitivos y psicológicos operantes en las decisiones humanas. Sin embargo, esta distinción es importante para defender la concepción realista de *Behavioral Economics*. En ella los experimentos anteceden a los modelos, permitiendo el reconocimiento de tipos de conducta o reacciones psicológicas que pueden ser luego intencionalmente “isolados” en el marco de un modelo con el propósito de explorar sus implicancias económicas.

Bibliografía

- Barberis, N. and Thaler, R. (2003), "A Survey of Behavioral Finance", in G.M Constantinides, M. Harris and R. Stuls, Ed., *Handbook of the Economics of Finance*, Elsevier Science B.V.
- Benartzy, Sh., and Thaler, R. (1995), "Myopic Loss Aversion and The Equity Premium Puzzle". The Quarterly Journal of Economics.
- Camerer, C., and Loewenstein, G. (2002), "Behavioral Economics: Past, Present and Future". Disponible en:
(<http://www.hss.caltech.edu/~camerer/ribe239.pdf>)
- Friedman, M., (1953), "The Methodology of Positive Economics". En: *Essays in Positive Economics*, Chicago: University of Chicago Press.
- Ho, T.H., Lim, N. and Camerer, C. (2005), "Modeling the Psychology of Consumer and Firm Behavior with Behavioral Economics"
<http://www.hss.caltech.edu/~camerer/JMRFinal.pdf>
- Kahneman, D. (2003a), "Maps of Bounded Rationality: Psychology for Behavioral Economics", AER, Vol. 93, N° 5, pp.1449–1475.
- _____(2003b), "A psychological Perspective on Economics", AER, Vol. 93, N° 2, pp.162-168.
- Kahneman, D. and Tversky, A. (1979), "Prospect Theory: An Analysis of Decision Under Risk". Econometrica, 47, pp.263-291.
- Kahneman, D., Knetsch, J.L. and Thaler, R.H. (1991), "The Endowment Effect, Loss Aversion, and Status Quo Bias". Journal of Economic Literature, Volume 5, N° 1, pp.193–206.
- Knetsch, J.L., (1989), "The Endowment Effect and Evidence of Nonreversible Indifference Curves". The American Economic Review, vol. 79, n° 5, pp.1277–1284.
- Maki, U. (2005), "Models are experiments, experiments are models". Journal of Economic Methodology, 12:2, pp.303-315.
- _____(2009), "Models and truth. The functional decomposition approach". European Philosophy of Science, 2007.
- Musgrave, A. (1981), "Unrealistic Assumptions' in Economic Theory: The F-twist Untwisted". Kyklos, 34, pp.377-87.

- Shefrin, H., and Thaler, R. (1988), "The Behavioral Life-Cycle Hypothesis". Economic Inquiry, 26 (4).
- Tversky, A. and Kahneman, D. (1981), "The Framing of Decisions and the Psychology of Choice". Science, New Series, Vol. 211, N° 4481, pp.453–458.
- _____(1986), "Rational Choice and the Framing of Decisions". The Journal of Business, Vol. 59, N° 4, Part 2, pp.S251–S278.

TIPOS NEUROECONÓMICOS COMO TIPOS INSTITUCIONALES. UN ESQUEMA DE INTERPRETACIÓN

Alejandro G. Miroli
(FCE-CIECE-Universidad de Buenos Aires / Instituto de Profesorado A-
1390 "Alfredo Palacios"/USAL)

Introducción

Este trabajo aborda el desarrollo de la noción de tipo institucional como noción central para la interpretación de las categorías neuroeconómicas. En §1 se exponen tres problemas de los tipos disciplinares presentes en los casos de disciplinas que integran niveles diferentes de la realidad v.g. la neuroeconomía; en §2 se expone el marco mínimo del debate sobre las clases naturales/sociales en las que se encuadra el análisis filosófico de los tipos disciplinares, en §3 se desarrolla la teoría de los racimos homeostáticos de propiedades propuesta por Richard Boyd (1988) y se propone que dicha teoría elucida en forma adecuada la noción de tipo disciplinar que involucra diferentes niveles de integración mereológica o tipos institucionales y se analiza un caso de aplicación.

§1.

En un trabajo anterior, (Miroli, 2008) se expuso un encuadre general para las relaciones entre neurociencias y teoría económica. En ese trabajo decíamos: "...la interpretación desde un léxico/ontología FH permite una integración no conflictiva de los resultados experimentales y los niveles jerárquicos con primacía del componente proléptico-intencional...opera un *methateoretial framework* no reductivo... la interpretación funcional-homuncular es perfectamente intuitiva... y la interacción entre proceso neural y efector volitivo se analiza *en términos de los tipos institucionales que postula FH*". Este párrafo introduce la noción de tipo institucional para la integración del vocabulario categorial neuroeconómico, pero deja sin examinar la pertinencia y el carácter de dicha noción.

La postulación de tipos institucionales como categorías¹ propias de una formación interdisciplinar es una respuesta al problema de la integración de niveles heterogéneos de realidad como herramienta que trata de evitar la tentación reduccionista, que muchas veces aparece.

En cada caso se trata de determinar condiciones de verdad de oraciones del tipo “X /Pred/ Y” donde X y Y son tipos que corresponden a niveles de integración y de clausura epistémica diferentes v.g. el neurológico y el económico. El diagnóstico tradicional era que dichas oraciones cometían errores categoriales y que dicha predicación era un sinsentido, tal como definía Ryle en su trabajo pionero sobre dicha cuestión

““Two proposition-factors are of different categories or types, if there are sentence-frames such that when the expressions for those factors are imported as alternative complements to the same gap-signs, the resultant sentences are significant in the one case and absurd in the other” (1938/1971, 181); dicho de otra manera el error categorial supone el fracaso de la sustitución de términos que preserve el sentido de la forma oracional considerada.

Sin embargo, dicha noción supone una multitud de problemas pues la noción de sentido, o la correspondiente noción de sinsentido o absurdo son nociones dependientes de contexto. En términos más austeros es posible reconocer tres problemas cuya resolución o no generaría o impediría los errores categoriales interdisciplinares.

Conmensurabilidad

¹ En el resto del trabajo asumimos que los términos del vocabulario científico tienen un significado reglamentado, que difiere del significado usual en los lenguajes naturales. Esa regimentación es resultado de la exigencia de determinar condiciones de verdad y de referencia libres de contexto –incluso libres del contexto de las relaciones semánticas que las palabras tienen en cada lengua natural- y condiciones para la coantificación de dichas términos científicos –en especial para la modelización numérica y la medición. En ese sentido las categorías tienen intensiones determinadas por condiciones necesarias y suficientes de identificación, y taxonomías explícitas, lo que no sucede en los lenguajes naturales. En el resto del trabajo usaremos los términos *categorías* y *tipos disciplinares* como sinónimos.

El establecimiento de condiciones de verdad categóricas para un esquema oracional “X /Pred/ Y” supone que las relaciones entre X y Y son relaciones medibles en algún sentido, y que dicha medición está protocolizada. O sea que incluso que X y Y sean tipos categorialmente equivalentes, o sea tipos disciplinares –y no meramente términos generales del lenguaje ordinario- hay una exigencia adicional a las exigencias de comprensión que una oración general tiene: el problema de su relación supone el problema de la conmensurabilidad estricta; o sea que los enunciados que empleen categorialmente tipos disciplinares deben diferir en un punto importante de la mera yuxtaposición de tipos disciplinares. Para apreciar el problema se puede dar un ejemplo en general²:

- (1) “El comportamiento criminal se incrementa al incrementarse el nivel de degradación urbana.”

Esta sentencia es un típico enunciado interdisciplinario que yuxtapone categorías del derecho penal y de la sociología urbana. Como tal el enunciado no tiene ningún problema para su comprensión y puede aparecer en el contexto de un debate político o de una columna periodística, pero si se lo quiere entender categorialmente sí ofrece problemas: en primer lugar “Comportamiento criminal” no es un término descriptivo sino normativo, como todo término normativo su significado es de textura abierta y la extensión del término depende de las valoraciones sociales v.g. la apostasía es un comportamiento criminal en Irán – castigado con la pena de muerte- pero es un comportamiento penalmente indiferente en la Argentina; en segundo lugar “Degradación urbana” es un término ambiguo ya que puede significar degradación en tanto un espacio urbano exhibe una desviación importante respecto de (i) un promedio de los espacios urbanos reales o de (ii) un estándar fijado. Lo que nunca ocurrirá es que “comportamiento criminal” sea un término del urbanismo. Por otro lado (1) es compatible con una variedad enorme de ratios de variación diferentes, dado que “incremento” también es un término ambiguo.

Pero en la medida que los enunciados disciplinares categóricos suponen enlaces reales, hay un problema porque dichos enlaces no se dan entre “comportamiento criminal” y “degradación urbana” tal como se los

² Para precisar el tema, hemos evitado dar un ejemplo que contenga predicados biológicos.

presentó, sino entre “comportamiento de cierto tipo” y “entorno urbano que exhibe tales características”. Para que (1) exhiba tales enlaces reales debe hacerse una paráfrasis que determine en forma fija la extensión de X y Y:

- (2) “El comportamiento tal y cual tiene la relación funcional F con las variaciones de tales y cuales rasgos del entorno urbano primario de los agentes examinados”

Un ejemplo de esto aparece en la discusión sobre el rol del miedo en las decisiones financieras i.e. “One complication is that unlike risk, which is an objective measure, it is difficult to measure fear, for the following reasons:

1. In many cases, fear is socially undesirable. Consequently, in surveys, people sometimes deny experiencing fear in particular situations.
2. Because of its emotional character, fear is often subconscious and difficult to articulate.
3. Fear can manifest itself suddenly and be present only fleetingly. Therefore, to be valid, the significance of fear for financial decision-making should be measured temporally close to its occurrence.

As a result, if one wanted to measure the influence of fear on financial decision-making, the first issue to deal with would be the appropriate method of measuring fear.” (Koning 2007, 7-8). Medir no es solamente brindar una escala de algún tipo, sino justificar que el contenido cognitivo y referencia de *miedo* como término general del lenguaje natural, queda capturado por *miedo** como tipo disciplinar³ y en general que la noción

³ El problema de la conmensurabilidad no se reduce a un problema de medición sino que –al ser un problema categorial involucra toda la estructura lógica del vocabulario teórico- puede verse en la inmensa discusión sobre la noción general de inteligencia del lenguaje natural y la categoría de inteligencia* que se desarrolló a partir de la aplicación de diversos test como el Test de Matrices Progresivas de Raven: en ese caso el acto –metodológico- de medir era al mismo tiempo el acto –epistemológico- de imponer una determinada categoría de inteligencia y el acto –ideológico- de naturalizar una propiedad psicológico-social; ello ha sido señalado tanto por algunos de los críticos de la identificación Inteligencia(noción general) = Ingeligencia* (Categoría psicológica); cf. Stephen Jay Gould *La falsa medida del*

cualitativa propia del vocabulario común queda capturada por la categoría teórica que se define explícitamente en el marco de una teoría.

Homología etiológica.

Las relaciones funcionales no son dadas sino que son relaciones empíricas que aparecen en la investigación experimental, cuasi experimental o modelística. Además de la dependencia funcional, la integración de categorías de disciplinas diferentes y epistémicamente clausuradas debe incluir una homología etiológica, ya que dicha integración tiene como objetivo -más allá de establecer correlaciones- el hallazgo de genuinos y novedosos explicadores para tipos de comportamiento, o sea determinar enlaces etiológicas. Precisamente el programa de una disciplina neuro-económica, como convergencia de las neurociencias y la economía se justificó de ese modo:

“Economics is typically defined as the science characterizing the optimal allocation of scarce resources...economics models individuals valuing rewards and choosing among alternatives... this definition of economics as it maps economic decisions straightforwardly into the neural substrates that produce these decisions <cada decisión puede desagregarse en tareas> ... (i) obtaining information from the environment regarding possible actions, (ii) valuing those actions, and (iii) choosing between them. Each of these three tasks is, in principle, measurable. Further, this hierarchy of how decisions are made can further be broken down into sub-tasks, including determining one's objective s, filtering incoming information, accessing memories of related events, using heuristics and identifying constraints on cognitive processing (e.g. energy or time constraints)... too, are measurable.” (Zack, 1737); en este punto la preservación del mapeo entre los niveles sub-agenciales y agenciales que

hombre, Leon Kamin *Ciencia y política del IQ*, Ashley Montagu, (ed). *Race and IQ*. Oxford University Press, 1999, ISBN [978-0195102208](#) y Peter H. Schönemann “Factorial definitions of intelligence: Dubious legacy of dogma in data analysis”, en I. Borg (Ed.), *Multidimensional data representations: When and why*. Ann Arbor: Mathesis Press, 325-374; precisamente el tema de la complejidad de una categoría disciplinar de inteligencia dio origen a teorías alternativas que se han propuesto para superar dicha identificación v.g. entre otros las teorías de Howard Gardner, Daniel Goleman y Robert Stenberg.

señala P. Zack es precisamente el resultado de la determinar la conmensurabilidad entre las categorías involucradas.

Pero el objetivo de las oraciones interdisciplinarias con tipos disciplinares diferentes no es meramente correlacional sino que “... the neuroeconomics research programme seeks to discover **proximate causes of choice behaviour.**” (Zack, 1737) de modo tal que dicho conocimiento permitiría ajustar y afinar la explicación e interpretación de tipos de comportamientos v.g. “...to answer fundamental questions... such as: why do two individuals faced with the same information and incentives make different choices? Why does the same individual sometimes make choices that are inconsistent? How much is choice behaviour affected by childhood development, if at all? Currently, most answers to economic questions focus on average choices, rather than individual or temporal variation in choices <de modo los modelos y teorías que integren datos neurológicos permitirían> ...an understanding of the interpersonal and intertemporal variation in choices is fundamental to effective public policy.” (Zack, 1738).

Ese programa de reconocimiento e investigación de las causas próximas de tipos de comportamientos se ha visto favorecida por el empleo creciente de técnicas no invasivas de Resonancia Magnética Nuclear y de sistemas de procesamiento electrónico de datos que integran la información neural, lo que ha permitido vincular tipos muy específicos de comportamiento con tipos muy específicos de fenómenos neurales-endócrinos.

Co-extensionalidad.

En el marco del problema mente y cerebro, un tema central fue el de la posible conexión o no entre categorías psicológicas y categorías neurológicas; en este debate fue central la tesis de Jerry Fodor sobre las ciencias especiales. Esta tesis fue presentada en un modo ya clásico: “...the reasons why economics is unlikely to reduce to physics are paralleled by those which suggest that psychology is unlikely to reduce to neurology...if psychology is reducible to neurology, then for every psychological natural kind predicate there is a co-extensive neurological natural kind predicate, and the generalization which states this co-extension is a law. Clearly, many psychologists believe something of the sort. There are departments of 'psycho-biology' or 'psychology and brain science...There are no firm data for any but the grossest correspondence between types of psychological states and types of neurological states

(Fodor, 104-59). Esta tesis introduce el tema de las clases naturales, en un sentido central la relación entre lo neurológico y lo comportamental involucra entidades y relaciones de cierto tipo. Efectivamente si las oraciones “X /Pred/ Y” introducen predicados etiológicos, en atención a determinar las causas próximas de comportamientos de un cierto tipo, tales predicados deben tomar caso del tipo Y y enlazarlo con cada caso del tipo Y en forma estable. Pero según el autor esto no ocurre pues “...it is entirely possible that the nervous system of higher organisms characteristically achieves a given psychological end by a wide variety of neurological means. If so, then the attempt to pair neurological structures with psychological functions is foredoomed . If so, then the attempt to pair neurological structures with psychological functions is foredoomed < dado que >...if (token) psychological events are (token) neurological events, it does not follow that the natural kind predicates of psychology are co-extensive with the natural kind predicates of any other discipline (including physics). That is, the assumption that every psychological event is a physical event does not guaranty that physics (or, *a fortiori*, any other discipline more general than psychology) can provide an appropriate vocabulary for psychological theories...” (Fodor, ídem).

En este sentido la exigencia –semántica- de coextensionalidad es una exigencia de la afirmación –ontológica- de relaciones causales entre las clases naturales que aparecen en los esquemas de oración que vinculan categorías de distintas disciplinas. Y dado que la coextensionalidad fracasa, no se pueden afirmar en las ciencias especiales las mismas relaciones etiológicas que se afirman en las ciencias básicas –que genéricamente se llaman ciencias físicas.

El texto de Fodor plantea el problema de forma absolutamente clara, con independencia de la propia solución que él ofrece. Es fácil entender que donde Fodor dice psicología se puede entender en general ciencias del comportamiento, incluso ciencias de un tipo especial de comportamiento: el que aplica recursos escasos en situaciones de elección entre fines disyuntos. La importancia de esta presentación es que pone el problema de la comunicación entre disciplinas ontológicamente heterogéneas como un problema de comunicación entre los tipos disciplinares –las categorías centrales del vocabulario teórico- que emplea cada una de ellas. En ese sentido elucidar el carácter de los tipos institucionales supone elucidar y precisar el problema de la vinculación entre disciplinas neurológicas y comportamentales.

§2.

El problema de los tipos institucionales entonces el siguiente: ¿Cómo se puede ofrecer un encuadre conceptual que preserven la noción de categoría o tipo institucional y de enlace etiológico que incluyan categorías físicas –i.e. neurológicas? Examinaremos un encuadre posible de los tipos institucionales que permita dicho desarrollo. Pero ello exige una elucidación mínima de la noción de tipo disciplinar. En un sentido lato los tipos disciplinares designan clases naturales/sociales, pero esto no explica mucho pues se exige la correspondiente aclaración de estas últimas. Respecto de ellas se han dado varias teorías que se pueden entender desde un nominalismo radical hasta formas del realismo modal – estos últimos presentes en algunas teorías filosóficas sobre las leyes físicas.

Algunos filósofos nominalistas sostienen que toda clasificación depende de intereses epistémicos y que la única existencia real es la de individuos que se pueden ordenar de cualesquiera manera y por cualquier interés; esta cuestión la expuso en un modo muy gráfico Ian Hacking: “When I have tried to indicate that there are interesting differences between 'human kinds' -- kinds of people and their behaviour -- and what are commonly called natural kinds, I found myself assaulted from left and right. The nominalist 'left' says that all kinds are human, or at any rate, there are no kinds in nature. The realist 'right' says that there are indeed natural kinds, and that human kinds -- at any rate those susceptible of systematic study - - are among them.” (Hacking, 109).

Incluso algunos filósofos con vena postmodernista han citado la archisabida clasificación que ofreciera Jorge Luis Borges y que inspirara uno de los textos cumbres del postmodernismo continental: “En sus remotas páginas está escrito que los animales se dividen en (a) pertenecientes al Emperador, (b) embalsamados, (c) amaestrados, (d) lechones, (e) sirenas, (f) fabulosos, (g) perros sueltos, (h) incluidos en esta clasificación, (i) que se agitan como locos, (j) innumerables, (k) dibujados con un pincel finísimo de pelo de camello, (l) etcétera, (m) que acaban de romper el jarrón, (n) que de lejos parecen moscas.” («El idioma analítico de John Wilkins», *Otras inquisiciones*, en *Obras completas*, Buenos Aires: Emecé, p. 708). Dejando de lado la retórica de derechas e izquierdas, el nominalismo reduce la discusión sobre categorías científicas

a lexicografía de diccionario, y equipara la discusión de la Real Academia Española en la revisión de su diccionario, con las discusiones en la elaboración del DSMV –el sistema de nosografía psiquiátrica en discusión hacia el año 2012. De ese modo el nominalista equipara una discusión sobre convenciones con una discusión sobre causalidad, inducción, capacidades disposicionales, propensiones y predicción y sobre los complejos problemas de diseño de muestras, análisis de correlación y teoría estadística. Pero no hay ninguna posibilidad de analizar siquiera rudimentariamente el papel de las clasificaciones en los libros de ciencia realmente existente con ninguna idea que surja de tal ejercicio literario o de convenciones lingüísticas; es más no hay ningún libro de ciencias que pueda interpretarse al modo de los nominalistas, en la medida que los tipos disciplinares tienen roles explicativos y predictivos y que efectivamente las clasificaciones muchas veces son centrales en el descubrimiento científico. Ni los libros de química que emplean la taxonomía química de Mendeleev, ni las discusiones entre cladistas y feneticistas, ni el análisis de roles de distintos tipos de sociedades económicas, ni la interpretación de los teoremas del cálculo de probabilidad exigidos en cualquier sistema taxonómico categorial, son meramente casos de agrupamientos de individuos sino que son discusiones y aplicaciones acerca de propiedades y roles etiológicos **objetivos**, seleccionados por sus capacidades para la explicación y la predicción. Otra cosa diferente es sostener que dichas propiedades son esenciales, o que son disposicionales o que deban entenderse en términos de poderes causales; esta sí es una discusión abierta.

Rechazando el nominalismo, es posible dar una cierta caracterización mínima de clase natural/social que parta de considerar el carácter independiente y objetivo de las propiedades centrales que caracterizan las clases naturales y que están involucradas en forma central en los roles explicativos y predictivos. Como se señala en un reciente estado de la cuestión “Assuming that one is a naturalist about (some of) our classifications, what makes a classification into one of natural kinds? The following are often suggested criteria or characteristics of natural kind classification: ...*Members of a candidate natural kind should have some (natural) properties in common....* Natural properties here are often taken to be *intrinsic properties... Members of a natural kind... should form a hierarchy ...<y> should be categorically distinct...*” (Bird & Tobin).

John Locke enunció correctamente este problema: solo cuando las categorías disciplinares se definen en términos objetivos e

independientes, y no en términos de intereses o convenciones es posible entender su éxito argumentativo, explicativo y predictivo; en terminología lockeana solo cuando se considera a las categorías disciplinares como esencias reales i.e. '[T]he real internal, but generally (in substances) unknown constitution of things, whereon their discoverable qualities depend, may be called their [real] essence ... and in this sense we speak of the essence of particular things...<lo que determina que todo proceso de investigación atienda > to have a real, but unknown, constitution of their insensible parts...<de> 'on all natural things; from which flow those sensible qualities which serve us to distinguish them one from another, according as we have occasion to rank them into sorts, under common denominations.' (Locke, *An Essay concerning Human Understanding*, III, 3, §15, § 17).

Si bien la concepción aristotélica ha sido tradicionalmente señalada como esencialista, Julius Moravcsik ofreció una interpretación de las categorías de una disciplina a partir de sus roles etiológicos; en esa dirección Moravcsik interpretó la noción aristotélica de *aitia* como **factor generativo**; y de ese modo ofreció un esquema analítico que preserva tales roles sin comprometerse con el esencialismo aristotélico v.g. "...what Aristotle proposed as a metaphysical scheme will be interpreted here - without prejudice- as a psychological and semantic claim. According to this thesis humans form concepts basically along the stuff + structure + function + causal power scheme. We shall label these structures aitiational schemes <o esquemas generativos>." (Moravcsik, 18), un discípulo de J. Moravcsik caracteriza tales esquemas como "... (1) *formal*: <structure> the basic category of which distinguishes the meaning of a word within a larger domain; (2) *constitutive* <stuff>: the relation between an object and its constituent parts; (3) *telic* <formal> the purpose or function of the object, if there is one; (4) *agentive* <causal power>: the factors involved in the object's origins or coming into being." (cf. James Pustejovsky, 383).

El enfoque de Julius Moravcsik es central para la superación del enfoque esencialista; en ese sentido el elemento basal para analizar un tipo disciplinar será el análisis del complejo **trama inferencial/roles etiológicos**; esto es equivalente a analizar la capacidad explicativa y predictiva de una ciencia. El examen de los roles explicativos y predictivos de las clases naturales/sociales es una consecuencia de la tesis naturalista mínima que está en la base de toda filosofía de las clases naturales/sociales: las propiedades intrínsecas involucradas en ellas no

son meramente propiedades aisladas o relaciones diádicas, sino que están enlazadas en sistemas y estructuras reales n-ádicas, lo que tendrá efectos en la consideración que se haga sobre la inducción, la explicación y las leyes físicas/sociales.

Ese carácter reticular de trama inferencial/roles etiológicos es analizado por el propio Moravcsik al afirmar "Philosophers are mesmerized by the identifying and referential roles of nouns and noun phrases. Thus they tend to overlook what is distinctive in certain predicative uses. The construal of natural-kind terms as rigid designators seems intuitively plausible when we consider the roles these terms play in subject position. They do, however, play equally important roles in predicative position. Especially significant is their role in "because"-clauses. For example, in using the phrase 'because it is made of iron' to explain the solidity of a structure, we rely on what will be construed as the essential properties of iron (or, in different words, associated criteria) to have our claim interpreted as having explanatory value. Without the link between natural-kind terms and associated properties, it is difficult to assign to these expressions the role they play in typical everyday or scientific explanations. In this respect, as we shall see, they differ from such paradigms of rigid designators as proper names." (Moravcsik, 9).

Podemos generalizar el ejemplo del autor y considerar aquellas esquemas oracionales "X /Pred/ Y" en los que /Pred/ no solo es "... a causa de" sino que puede ser reemplazado por toda una variedad de clausulas de esquemas generadores que ofrezcan explicaciones parciales o totales, deterministas o estocásticas, predicciones certeras o ajustadas, fundadas en modelos o en trayectorias temporales, etc. En este caso el tipo disciplinar/clase natural-social no solo aparecerá en funciones nominales entre un término y una clase determinada por una propiedad, sino que nos encontramos con una red de niveles de abstracción y de clausulas etiológicas.

§3.

El desarrollo de la trama inferencial/sistema de roles etiológicos supone la conmensurabilidad y la homología etiológica de las clases naturales/sociales involucradas, pero no supone una relación de cada-

uno-de-los-eventos-X-a-cada-uno-de-los-eventos-Y. Efectivamente fragmentos de dicha trama inferencial/sistema de roles etiológicos serán las causas próximas de determinados sesgos o tendencias o trayectorias temporales de comportamientos, y cada uno de los parámetros involucrados deberá poder medirse en algún sistema cuantitativo de representación de la información que provea la investigación sobre dichos fragmentos de la trama inferencial/sistema de roles etiológicos. Pero la exigencia de co-extensionalidad puede ser rechazada. Específicamente el hecho que un comportamiento tenga varios efectores materiales –v.g. diferentes roles del sistema endócrino o de localizaciones cerebrales- no impide la formulación de leyes estrictas, en contra de la asunción que plantea la tesis de Fodor.

Esto supone una revisión de la doctrina de las clases naturales/sociales en la dirección que señalara Julius Moravcsik o sea considerando que estos términos no solo nombras agregados reales de un único nivel mereológico-funcional que se pueden representar con la teoría de conjuntos, sino que suponen propiedades complejas de sistemas de relaciones n-ádicas y de relaciones entre estructuras y diferentes niveles mereológico-funcionales. O sea que debemos salir de cualquier consideración de las categorías disciplinares a partir de la función nominal exclusivamente. Richard Boyd ha presentado una teoría ontosemántica de las clases naturales/sociales, que permite resolver los problemas de conmensurabilidad y congruencia, al tiempo que rechazar la exigencia de coextensionalidad estricta: la **teoría de los racimos homeostáticos de propiedades** (conocida técnicamente como **homeostatic property clusters** o HPC)⁴-. Dicha teoría se ha empleado en el contexto de la filosofía de la biología para dar cuenta de la noción de especie biológica para negar el esencialismo biológico i.e. "...la tesis de que éstas son géneros naturales, pero sin que esto lleve necesariamente a aceptar que tienen una esencia en el sentido descrito¹⁸. Según R. Boyd y R. Wilson,

⁴ La teoría de los agregados homeostáticos de propiedades fue propuesta originalmente por Richard Boyd en el contexto de una discusión sobre el realismo moral (en su "How to be a Moral Realist," in G. Sayre-McCord (ed.), *Essays on Moral Realism*, Ithaca, NY: Cornell University Press 1988), y posteriormente desarrollada por el mismo autor y otros filósofos como Philip Kitcher, Paul Griffiths y Robert Wilson cfr Dieguez 2009. Es interesante señalar que la teoría de Richard Boyd parece haberse desarrollado con entera independencia de las tesis de J. Moravcsik.

las especies serían géneros naturales entendidos como *agrupaciones de propiedades mantenidas homeostáticamente*. Lo que esto significa es que las especies presentan propiedades que tienden a darse juntas porque hay mecanismos causales subyacentes que hacen que se refuercen unas a otras, aunque son mecanismos sujetos a excepciones. Es decir, la posesión de alguna de ellas hace más probable la posesión del resto, si bien pueden darse casos de individuos en los que falten algunas de tales propiedades. En tales casos, es una cuestión teórica, para *a posteriori*, y no meramente conceptual, determinar si estamos ante un individuo que cae bajo ese género natural o no. Incluso podrían darse casos racionalmente indecidibles.” (Dieguez, 97).

Boyd da una presentación exhaustiva de dicha teoría en el marco de una discusión sobre el realismo moral, en ella las tesis ontológicas fundamentales son las siguientes⁵:

1. There is a family *F* of properties which are 'contingently clustered' in nature in the sense that they co-occur in an important number of cases.
2. Their co-occurrence is... a sort of *homeostasis*. Either the presence of some of the properties in *F* tends (under appropriate conditions) to favor the presence of the others, or there are underlying mechanisms or processes which tend to maintain the presence of the properties in *F*, or both.
3. The homeostatic clustering of the properties in *F* is causally important: there are (theoretically or practically) important effects which are produced by a conjoint occurrence of (many of) the properties in *F* together with (some or all of) the underlying mechanisms in question.
-
6. Imperfectly homeostasis is nomologically possible or actual: something may display some but not all of the properties in *F*:

⁵ La teoría que propone Richard Boyd no solo caracteriza la ontología de las clases naturales/sociales sino que ofrece una teoría del significado de los términos de clase natural/social. Por razones de argumentación hemos limitado la presentación de sus tesis ontológicas.

-some but not all of the relevant underlying homeostatic mechanisms may be present⁶.

La determinación del racimo específico que opera en una situación no será una cuestión conceptual sino empírica, y cuando dicha determinación se pueda lograr "...<8>will typically depend upon quite particular facts about the actual operation of the relevant homeostatic mechanisms, about the relevant background conditional and about the causal efficacy or the partial cluster or properties from F."; esto conlleva que el vocabulario categorial de clases-racimo pueda exhibir "...<9> many cases of extensional vagueness which are such that they are not resolvable, even given all the relevant facts and all the true theories." (Boyd, 197).

De acuerdo con la teoría HPC en un esquema oracional teórico "X /Pred/ Y" tanto X como Y designan racimos homeostáticos de propiedades + mecanismos enracimadores. Consideremos el caso mencionado. En ese caso *comportamiento inversor pro-portafolios domésticos (home bias investment)* y *miedo* pueden considerarse tipos disciplinares. En su estudio sobre el sesgo hacia carteras domésticas en inversores, Peter Kenning *et al* hallaron que "Differential brain activity was detected between decisions involving funds of different national origins. In situations where participants had to decide between mutual funds from different countries, we found increased activity in the precentral gyrus, the fusiform gyrus and the inferior occipital gyrus. Moreover, during home-biased decisions we found a correlation between activity within the amygdalahippocampal regions of the brain and the investor's general risk aversion. This region has been found to be involved in negative emotional processing such as fear, so one interpretation is that home-biased financial decision making is modulated by negative emotions associated with risk aversion." (op. cit.).

Los autores sostienen la identificación *Miedo = Actividad específica de regiones de la amígdala y el hipocampo del sistema límbico*; pero no afirman que haya una dependencia estricta (correlación constante) entre actividad límbica y decisiones financieras sesgadas hacia activos domésticos sino que su hipótesis es "Against this background, we hypothesize that homebiased financial decisions might correlate with

⁶ Mantenemos la numeración original de las tesis del autor, centrándonos en las tesis ontológicas.

neural activity within the amígdala” (ídem). La salvedad *Aganist this backraund* es precisamente la asunción del tipo disciplinar *Miedo = Actividad específica de regiones de la amígdala y el hipocampo del sistema límbico*. Por consiguiente los autores no sólo postulan una correlación sino una correlación más un mecanismo casual específico que opera como causa próxima del comportamiento económico de cierto tipo. Pero dicho mecanismo causal no opera en cada caso de inversión –o sea que la proposición “*Miedo = Actividad específica de regiones de la amígdala y el hipocampo del sistema límbico genera un comportamiento inversor pro-portafolios domésticos*” no es una ley física tal como la caracteriza Fodor. Pero tampoco es una tendencia estadística sometida a una clausula *caeteris paribus* no determinada; por lo contrario dicha proposición supone una genuina ley de tipos institucionales.

La noción central para caracterizar un tipo disciplinar/clase natural-social es la de racimo homeostático. Una serie de propiedades están vinculadas entre sí por relaciones materiales de co-ocurrencia y co-variancia v.g. el sistema límbico, el sistema endócrino y el sistema cognitivo tienen determinadas una cierta trama de relaciones materiales entre sus rasgos centrales; el sistema límbico opera de determinada manera y tiene tales efectos asociados con ciertos estados de los sistemas endócrinos y cognitivo, pero no con cualquiera, dicha asociación es contingente si se considera cada rasgo por separado –i.e. el rasgo R1 del sistema límbico se puede asociar con el Rasgo RC1 o RC2 del sistema cognitivo- pero cada ensamble de rasgos de cada sistema operan en una red que opera como efector nomológico de ciertos comportamientos.

Mientras que la concepción estándar de las clases naturales/sociales las limita a un único nivel mereológico-funcional de realidad –como condición de la conexión cada-caso-de-X-con-cada-caso-de-Y- en la concepción institucional las clases pueden cortar niveles jerárquicos: así *estado de temor/miedo* y comportamiento reductor de temor –del cual el comportamiento inversor pro-portafolios domésticos es un caso- pueden aparecer en genuinos enlaces nomológicos, en los que no habrá una conexión constante de caso-de-actividad-amígdala-hipocampo-límbica a caso-de-variación-de-miedo y caso-de-inversión-tal-y-cual, pero si habrá una conexión constante entre racimos-homeostáticos-de-rasgos-del-sistema-límbico/subagencial y rasgos-de-comportamientos-intencional/agencial. Cada clase involucrará un enlace homeostático

común pero diferentes rasgos operando en cada caso dentro de un conjunto determinado por dicho enlace. Dicho de otra manera *miedo* en tanto actividad del área admigdalo-hipocampo o subagencial será un recorte de rasgos de niveles distinto que *miedo* en tanto comportamiento agencial, pero en cada caso la relación n-ádica de homeostasis que opere delimitará racimos equivalentes. Una estructura real –por ejemplo el mecanismo bioquímico y neural del miedo- es al mismo tiempo un elemento comportamental que enracima distintas propiedades en distintos agentes, pero que preserva un núcleo homeostático –de variación conjunta.

§4.

La propuesta de R. Boyd permite sostener la existencia de genuinos enlaces nomológicos institucionales entre diferentes niveles de escala y organización mereológica, y el rol explicativo próximo de los mecanismos involucrados en dichos enlaces nomológicos, a partir de interpretar a los tipos disciplinares como racimos homeostáticos. Nuestro supuesto filosófico es que esa teoría permite elucidar los tipos institucionales; en ese sentido el programa de la neuroeconomía puede superar las objeciones estándares contra las leyes psico-físicas y desarrollar proyectos empíricos de reconocimiento y examen de las causas próximas –subagenciales- del comportamiento económico.

Bibliografía

- Bird, Alexander & Tobin, Emma (2008), "Natural Kinds", *The Stanford Encyclopedia of Philosophy (Summer 2010 Edition)*, Edward N. Zalta (ed.), disponible en: <http://plato.stanford.edu/archives/sum2010/entries/natural-kinds>. Revisada el 05/10/10.
- Diéguez, Antonio (2009), "¿Es la vida un género natural? Dificultades para lograr una definición del concepto de vida". *ArtefaCToS*, 1, pp.81-100, ISSN 1989-3612.
- Boyd, Richard (1988), "How to Be a Moral Realist". En: G. Sayre-McCord (ed.) *Essays on Moral Realism*, Ithaca, Cornell University Press, pp.181–228.

- Charland, Louis (2002), "The Natural Kind Status of Emotion" Brit. J. Phil. Sci. 53, pp.511-537.
- Fodor, J. (1974): "Special sciences and the disunity of science as a working hypothesis". Synthese, 28, pp.77-115, ISSN 0039-7857.
- Hacking, Ian (1991) "A tradition of natural kinds". Philosophical Studies, 61, pp.109-126, ISSN: 0031-8116.
- Hanson, Norwood Russell (1959), "On the Symmetry between Explanation and Prediction". The Philosophical Review, 68, pp.349-358.
- Hempel, Carl & Oppenheim Paul (1948), "Studies in the Logic of Explanation". Philosophy of Science, 15, pp.135-175; (version castellana en *La explicación científica*, Paidós, Buenos Aires, 1979).
- Kenning, Peter et al. (2007), "The role of fear in home-biased decision making -first insights from neuroeconomics". Westfälische Wilhelms-Universität Münster; Neuroeconomic Research Report No. 2. Disponible en: <http://mpra.ub.uni-muenchen.de/1076/> ; consultado el 21/09/2010.
- Locke, John (1690), *An Essay concerning Human Understanding*; edición electrónica de Jonathan Bennett. Disponible en <http://www.earlymoderntexts.com/loess.html>.
- Miroli, Alejandro (2008), "De la neurología a la decisión racional: el camino del funcionalismo homuncular"; en. Gustavo Marques (ed) *Racionalidad, Economía e Interdisciplinariedad* 1a ed., Buenos Aires: Ciece. FCE-UBA.
- Moravcsik; Julius (1981), "How Do Words Get Their Meanings?". The Journal of Philosophy, 78 (1981), pp.5-24
- Pustejovsky, James (1991, 2004), "The Generative Lexicon". En: Steven Davis & Brendan S. Gillon (eds.), *Semantics: A Reader* Nueva York, Oxford University Press USA, 2004, pp.369-393, ISBN 0195136977
- Ryle, Gilbert (1938/1971), "Categories". En: *Collected Papers, Volume II: Collected Essays*. New York: Barnes and Noble
- Zak, Paul. J. (2004), "Neuroeconomics" Philosophical Transactions of the Royal Society B (Biology), 359(1451), pp.1737-1748.