

"Del progreso instrumental al progreso de la racionalidad", en Ambrosio Velasco (comp.), Progreso, pluralismo y racionalidad en la ciencia. Homenaje a Larry Laudan, IIFs- UNAM/FFyL-UNAM, México, 1999, pp. 81-102

Del progreso instrumental al progreso de la racionalidad*

SERGIO F. MARTÍNEZ
Universidad Nacional Autónoma de México

... commitment does involve, in a very real sense, counterpreferential choice, destroying the crucial assumption that a chosen alternative must be better than (or at least as good as) the others for the person choosing it, and this would certainly require that models be formulated in an essentially different way.

Amartya Sen

I

1. Introducción. En *El progreso y sus problemas*, publicado en 1977, Larry Laudan elabora un modelo del cambio científico basado en la idea de que todas las preguntas importantes respecto a la racionalidad de la ciencia pueden reducirse a evaluaciones del progreso en términos de una medida: la solución de problemas. Un ingrediente de este modelo era una cierta manera de entender "intuicionistamente" el registro histórico. Cualquier lector cuidadoso de ese registro podría llegar a decisiones incontrovertibles acerca de esos ejemplos de elección racional que constituirían la "base de datos" para evaluar las teorías rivales.

En sus trabajos posteriores Laudan ha abandonado este ingrediente intuicionista de su meta-metodología, y en su lugar pone énfasis en la necesidad de reconocer y explotar la separación tajante que según él existe entre el tema del progreso científico y el tema de la racionalidad. El tema de la racionalidad se refiere a la cuestión de qué tan eficientemente un agente individual pone sus creencias al servicio de sus fines a través de sus acciones; la racionalidad depende del contexto y el agente, es una racionalidad puramente instrumental. El progreso, por otra parte, no es dependiente del agente y el contexto. El progreso lo sigue caracterizando Laudan como en 1977, en términos de la

* Este trabajo se elaboró con apoyo del proyecto PAPIIT IN 402197.

efectividad en la solución de problemas con respecto a nuestros fines: decir que la ciencia ha progresado es decir que la ciencia puede resolver mejor los problemas con respecto a los estándares en relación con los cuales existe un problema y una solución.¹ Lo que nos permite establecer si ha habido o no progreso en un cierto campo de investigación es un examen de las maneras en las que la historia de la ciencia puede interpretarse como un registro que muestra que, con respecto a nuestros estándares, nuestros métodos son mejores para la solución de problemas. En sus trabajos posteriores (1984; véase nota 1), Laudan ha puesto énfasis en la importancia que él le otorga a la necesidad de entender de una manera diferente a la forma tradicional (basada en lo que llamo adelante la teoría estándar) cómo tienen lugar las decisiones en la ciencia. Laudan propone lo que él llama un modelo reticular de la racionalidad. Según este modelo, los fines, los métodos y las teorías se ajustan mutuamente.

Laudan en particular ha hecho ver que el supuesto implícito (heredado de la teoría estándar de la decisión racional) de que los fines no son racionalmente cuestionables es un supuesto que debemos abandonar. La idea de que los fines pueden ser objeto de discusión racional es uno de los temas centrales y distintivos de la filosofía de la ciencia de Laudan. Cuando en 1977 Laudan publicó *El progreso y sus problemas*, la moda filosófica consistía en elaborar las implicaciones mal entendidas del “historicismo” pregonado por Kuhn y Feyerabend, que utilizando este supuesto de la imposibilidad de poner a discusión racional los fines de la ciencia, llegaba rápidamente a la conclusión de que una crítica al positivismo nos conducía sin remedio a un relativismo epistémico radical. Según este relativismo, no podía justificarse un concepto de progreso científico que permitiera articular de manera filosóficamente satisfactoria el sentido en el que el conocimiento crece. *El progreso y sus problemas* es el inicio de un proyecto que sigue desarrollándose hoy en día, a través del cual Laudan muestra que es posible evitar el relativismo asociado usualmente con los “historicistas” por medio de una teoría de la ciencia que parte de reconocer que si bien la ciencia es progresiva en el mismo sentido

¹ Laudan dice en varios lugares que el progreso consiste en el desarrollo de una sucesión de teorías que se acercan cada vez más a lograr que un cierto fin se actualice (p. 65 de *Science and Values*, University of California Press, 1984; p. 78 de *Beyond Positivism and Relativism*, Westview Press, 1996).

que cualquier otra actividad, en el sentido que el progreso es simplemente una medida del éxito de ciertos medios para la realización de un fin, la ciencia es distintiva porque es una actividad con fines claramente definidos y distinguibles de los fines de cualquier otra actividad, fines cognitivos asociados con las virtudes predictivas de una teoría.

Comparto con Laudan la idea de que es sólo a través de una serie de confusiones y malinterpretaciones como se puede saltar del reconocimiento de los problemas con una concepción acumulacionista de la ciencia a una tesis “historicista” que niega que el origen de las normas epistémicas de la ciencia sean algo más que convenciones o negociaciones exitosas entre comunidades con diferentes intereses, una tesis que implícitamente niega la importancia que tiene en la ciencia la evaluación racional de los fines. Así, Laudan reconoce que la evaluación racional de los fines es importante, pero como veremos, este tipo de consideraciones sólo entra de manera muy rudimentaria en su teoría de la ciencia, y en particular, asume supuestos respaldados por una concepción instrumentalista de la racionalidad que no permiten elaborar una racionalidad de fines de una manera satisfactoria.

Una formulación cuidadosa del concepto de racionalidad instrumental es difícil, pero ni Laudan la hace ni tal formulación cuidadosa se necesita para nuestros propósitos. Para nuestros propósitos es suficiente hacer ver que la noción de racionalidad instrumental utilizada por Laudan asume varios supuestos de lo que llamaremos la teoría estándar de la decisión racional, una teoría desarrollada sobre todo en la economía, pero íntimamente ligada a importantes supuestos de la filosofía analítica. Esta teoría estándar de la decisión racional está ligada a una añeja tradición según la cual se tienden a modelar las decisiones como si estuvieran motivadas por preferencias (deseos) articulados por autointerés. La idea fundamental de la teoría estándar es que las decisiones racionales son cierto tipo de relación entre preferencias, acciones y consecuencias, y nada más. Esta idea toma cuerpo sobre todo a través de la tesis según la cual sólo es posible evaluar racionalmente una decisión con respecto a un conjunto completo, determinado y homogéneo de alternativas. Este supuesto se formula muchas veces (y Laudan en particular así lo hace) como el supuesto de que toda decisión racional tiene lugar con respecto al conjunto de utilidades ponderadas. Esto requiere que se asuma que las alternativas (y nuestras expectativas respecto de ellas) puedan determinarse por adelantado, con independencia del tipo de intervenciones en el

mundo que de manera usual requerimos de hecho para tener claridad respecto a las alternativas involucradas en una decisión. Este supuesto se manifiesta usualmente (y en Laudan así es) despreciando el problema de la caracterización de la estructura de los agentes-*cum*-contexto que hacen ciencia. Se piensa que la racionalidad que atañe a los filósofos no tiene por qué preocuparse por la manera como la estructura interna psicológica de los agentes y la estructura del medio ambiente (del contexto de instituciones y organizaciones) del que forman parte se constituye en un medio ambiente cognitivo con una “geografía de la razón” (un término que adelante explicaré muy brevemente y que en otros trabajos elaboro más a fondo).

Por ello, si bien Laudan es uno de los primeros filósofos de la ciencia que de manera sistemática reconoce la importancia que tiene el supuesto implícito de la imposibilidad de cuestionar racionalmente los fines de la ciencia en el planteamiento tradicional historicista, y las implicaciones que este supuesto tiene en la manera de plantear los problemas, Laudan retiene otro supuesto de la teoría estándar de la decisión racional que yo considero que es importante abandonar para realmente poder formular una teoría de la racionalidad científica. Según este supuesto, los juicios acerca de la racionalidad de una decisión en la ciencia pueden basarse exclusivamente en un análisis de las consecuencias esperadas de las posibles alternativas o estrategias que podemos seguir para alcanzar aquello que nos proponemos. Es sobre la base de este supuesto que Laudan hace la distinción tajante entre racionalidad y progreso, y elabora un modelo de la racionalidad científica, de manera simple y elegante (en términos de su modelo reticular de la racionalidad).²

En este trabajo quiero hacer dos cosas. En la parte I, a continuación de la presente introducción, muestro cómo una serie de críticas a

² Laudan a veces sugiere que el tipo de supuestos implícitos en la concepción instrumentalista de la racionalidad deben verse nada más como condiciones necesarias y no suficientes de la racionalidad. Como veremos, una concepción de la racionalidad mínimamente satisfactoria para modelar apropiadamente el tipo de decisiones que tiene lugar en la ciencia va bastante más allá de esas condiciones. Así, incluso asumiendo que esas condiciones son nada más necesarias, el punto sería que esos supuestos son prácticamente nunca suficientes. En este sentido, los elementos del concepto de racionalidad instrumental que Laudan toma como punto de partida para su análisis del proceso de decisiones en la ciencia son insuficientes para delinear una teoría satisfactoria.

la teoría estándar de la decisión racional en economía son extensibles a la teoría de la decisión de Laudan (implícita sobre todo en su “modelo reticular”). En la parte II sugiero una manera de elaborar una concepción no instrumental de la racionalidad que tiene en cuenta las críticas a la concepción estándar y que de manera natural no permite hacer la distinción tajante entre progreso y racionalidad que es lo que ha permitido a Laudan modelar el cambio científico sin enmarañarse en el tema de las “reconstrucciones racionales”, que según él es el pecado de los historicistas.

Estas críticas a la concepción estándar de la decisión racional, que son extensibles a la teoría de la decisión racional que Laudan promueve como modelo de la decisión en la ciencia, tienden a hacer que se ignore el papel constitutivo que desempeñan los valores, sobre todo a través de sus expresiones en instituciones y organizaciones de diverso tipo, en procesos de decisión.³ La primera crítica, vagamente formulada, es que no es posible entender el tema de la decisión racional a partir de una generalización de una teoría subjetivista de la decisión racional. En particular quiero criticar el supuesto implícito de que no hay diferencias significativas en la manera como se plantea el problema de una decisión racional para agentes que son parte de organizaciones de diferente tipo. Más adelante elaboraremos un poco este supuesto, pero la idea es que el supuesto cumple un papel importante en justificar la tendencia a ignorar el problema de cómo las personas individuales y las instituciones individualizan a los medios que consideran alternativas legítimas para alcanzar un fin, como parte de un contexto de compromisos y valores enmarcados en un contexto de instituciones. Más en general, tiende a ignorarse el problema de cómo la caracterización de los medios y los fines de las personas individuales moldea o le da forma a lo que se reconoce a nivel social como medios y fines legítimos, y viceversa.

El reconocimiento de esta dependencia de la caracterización de los medios y fines del nivel o tipo de organización lleva de manera natural a cuestionar la posibilidad de analizar la racionalidad de una acción por intermedio de un análisis de las consecuencias esperables. El otro

³ El sentido en el cual ese papel constitutivo no es aceptable para Laudan tiene que ver con el hecho que tal y como los valores desempeñan un papel constitutivo en el proceso de una decisión no puede reducirse a un análisis de las consecuencias esperadas en relación con los fines buscados. Más sobre esto adelante.

supuesto de la teoría estándar que me interesa destacar y criticar es aquel según el cual una decisión racional puede hacerse sin tomar en cuenta que la evaluación de las diferentes alternativas, por lo menos en cierto tipo de decisiones que son pertinentes en la ciencia, requieren un involucramiento de los agentes en las alternativas, de una manera que imposibilita evaluar las diferentes alternativas involucradas en una decisión en términos de las consecuencias esperadas de las diferentes alternativas. Las consecuencias sólo se constituyen en el proceso de la interacción del agente con la alternativa elegida.

En la sección 2 presento algunos aspectos pertinentes de la teoría estándar de la decisión, y examino críticamente el primer supuesto. En la sección 3 examino el segundo supuesto. En ambas secciones mi interés central es mostrar cómo el reconocimiento de la inviabilidad de esos supuestos incide en lo que debe ser una teoría de la racionalidad en la ciencia, y que en particular este reconocimiento puede reformularse como un cuestionamiento de la distinción tajante que Laudan hace entre metodología y progreso.

2. La teoría de la decisión como una generalización de la teoría subjetivista de la acción. De acuerdo con la teoría estándar de la decisión, un agente racional tiene preferencias que están suficientemente determinadas como para que puedan considerarse como dadas.⁴ Se asume además

⁴ De acuerdo con lo que se conoce como “teoría de la mejor acción”, la elección de una acción es racional cuando es el resultado de un procedimiento de decisión confiable. Un procedimiento de decisión es un procedimiento de deliberación, el procedimiento es confiable cuando típicamente resulta en la elección de una acción que, de ser hecha en el tiempo apropiado y en las circunstancias que se cree son el caso, permitirían alcanzar los fines que el agente tiene la intención de alcanzar. Esta teoría de la mejor acción dice algo acerca del tema de la elección racional bajo el supuesto de que el proceso que lleva a la elección en cuestión es de un cierto tipo. Si por ejemplo seleccionamos al azar la mejor acción de entre una serie de alternativas, no quisiéramos decir que en ese caso la elección es racional. Este hecho apunta a un supuesto importante implícito en la teoría estándar, el supuesto de que es parte de nuestra naturaleza el actuar, por lo menos en condiciones ideales, de manera cercana a como predice la teoría estándar. Exactamente cómo entendemos esas “condiciones ideales” es problemático, pero en todo caso es claro que la teoría estándar tendría que tomar en cuenta dificultades que tienen que ver con el hecho de que, como muestran una serie de experimentos en la psicología de la inferencia, los seres humanos tienden a cometer errores sistemáticos en el razonamiento, por lo menos en ciertas condiciones, que los alejan de lo que prescribe la teoría estándar. Esta brecha entre la mejor acción y la acción racional es una cuestión importante para el tema de este trabajo, pero aquí lo dejaremos de lado.

que estas preferencias son homogéneas y constituyen un conjunto completo y consistente de alternativas. Usualmente se tiende a asumir que esta teoría no se compromete con ningún supuesto adicional. En particular, se piensa que esta teoría no requiere adjudicarle contenido a esas preferencias, ni requiere que identifiquemos al agente como un cierto tipo de agente. Si el agente es una persona individual, o una comunidad o un país, esto no es importante para la aplicación de la teoría. En la medida que podamos identificar el conjunto consistente y completo de preferencias (determinadas y homogéneas), el problema de la racionalidad de una decisión o acción puede modelarse, y la teoría servir como norma de racionalidad.

Este supuesto puede criticarse desde diferentes perspectivas. En primer lugar hay buenas razones para pensar que las normas cognitivas y las normas morales tienen una base común, y que no es posible entender la estructura de las normas cognitivas sin tomar en cuenta que las normas y las decisiones muchas veces se toman por organizaciones o instituciones.⁵ Una crítica clásica de la concepción estándar en esa dirección es el trabajo “Rational Fools” de Amartya Sen. Sen hace ver que los compromisos con un principio o valor generan una variante de la idea del autointerés que permite distinguir dos tipos de decisiones: aquellas que pueden modelarse como “elecciones personales” en el sentido estricto, y aquellas que involucran el “bienestar personal”.⁶ La idea es que si bien las elecciones personales pueden modelarse por la teoría de las preferencias en un sentido estricto, las elecciones que involucran el concepto de bienestar personal requieren tomar en cuenta el concepto de metapreferencia (o preferencia de segundo orden). Sen y otros economistas y científicos sociales han hecho ver que estas distinciones son importantes en economía y otras

⁵ Véase por ejemplo *How Institutions Think*, de Mary Douglas, Syracuse University Press, 1986. Mary Douglas formula la idea como sigue: “The whole approach to individual cognition can only benefit from recognizing the individual person’s involment with institution-building from the very start of the cognitive enterprise. Even the simple acts of classifying and remembering are institutionalized” (1986, p. 67). Como Douglas lo reconoce, la idea de que el conocimiento debe entenderse como un tipo de bien público proviene de Ludwig Fleck. Hay varios intentos recientes en ética y epistemología de recobrar esta idea desde una perspectiva kantiana. Véase por ejemplo el trabajo de C. Korsgaard, *The Sources of Normativity*, y los trabajos de John McDowell que mencionaremos después.

⁶ Una reimpresión reciente de ese artículo de Sen se encuentra en la antología compilada por Jane J. Mansbridge, *Beyond Self Interest*, Chicago, 1990.

áreas clásicas en las que se plantea el problema de la decisión racional, y que el concepto de racionalidad no se agota con un modelo de preferencias. Albert Hirschman, por ejemplo, a partir de estas distinciones retoma un tema que ha desempeñado un papel importante en la filosofía de la ciencia de Fleck y Polanyi, entre otros, la importancia que tiene el despliegue de habilidades en las acciones como parte de un proceso de decisión.⁷ En la medida en que ese despliegue se toma en cuenta, muchas veces un proceso de decisión no puede reducirse a la consideración de las consecuencias esperadas de acciones consideradas como medios. Diremos más al respecto en la próxima sección.

Por otra parte, una serie de estudios empíricos han hecho ver que la percepción de lo que es justo en relaciones económicas desempeña un papel importante en entender el comportamiento y los cambios de patrones de consumo de una manera que no es compatible con la teoría estándar. Estudios acerca de actitudes hacia la colaboración han hecho ver que un porcentaje significativo de personas está dispuesta a cooperar como parte de un compromiso con un esfuerzo de ser justos con los demás. Hay estudios que muestran que el nivel de cooperación puede aumentarse significativamente en un grupo en donde se fomentan discusiones que tiendan a generar confianza mutua y sentimientos de identidad de grupo (véase, por ejemplo, varios de los trabajos en Mansbridge 1990). Un concepto como el de preferencia de segundo orden de Sen permite modelar las decisiones que toman este tipo de factores en cuenta como racionales, y por lo tanto no tenemos que contentarnos con la tendencia usual en economía y en filosofía de la ciencia de considerar “irracionales” estos factores.

Es posible entender a Laudan como reconociendo la importancia de las preferencias de segundo orden en la evaluación racional de fines; sin embargo, es importante hacer notar que en la medida que este concepto se toma en serio, el supuesto consecuencialista de la teoría estándar que Laudan comparte se torna poco creíble. No sólo las consecuencias esperadas, sino la percepción de lo que es justo, por ejemplo, desempeñan un papel en la evaluación de la racionalidad de una acción.

⁷ Uno de los muchos trabajos en los que esta distinción de Sen se elabora como parte de un reconocimiento de la importancia de cambios racionales en preferencias, y de la importancia de distinguir entre cambios en gustos y cambios en valores, es “Against Parsimony”, de Albert Hirschman, *Economics and Philosophy*, 1, 1985, pp. 7–21.

Muchas veces se piensa que es posible reformular (o extender) la teoría estándar de una manera en que se supere el tipo de dificultades que hemos mencionado, sin cuestionar la importancia de ese núcleo normativo al que Simon y otros llaman “la teoría sustantiva”, y que podemos identificar con lo que llamo aquí la teoría estándar. Voy a concentrarme en examinar una de estas pretendidas reformulaciones de la teoría estándar, la teoría satisfaccionista de Herbert Simon. Ésta es una teoría que puede considerarse muy exitosa en cuanto que permite superar muchas de las dificultades mencionadas, pero como veremos, no puede considerarse como una extensión o reformulación satisfactoria de la teoría estándar.

La *teoría satisfaccionista*, inicialmente propuesta por Herbert Simon en los años cincuenta, y que ahora es el punto de partida de una serie de modelos muy importantes en economía y otras áreas de la ciencia, reconoce la importancia que tienen las limitaciones de las capacidades cognitivas (producto de una evolución biológica) de los agentes humanos en la manera de plantearse y tomar una decisión. Así, se reconoce que los agentes tienen niveles de aspiración o “satisfacción” que sirven para determinar las alternativas que deben considerarse. Se evalúan las alternativas que están dentro del nivel de aspiración que pragmáticamente se elija. En esta teoría el nivel de aspiración, así como las alternativas que han de considerarse, se plantean y se modifican según nuestra experiencia. Las decisiones se hacen siguiendo una regla heurística, por ejemplo, la regla “selecciona la primera alternativa que satisfaga tu nivel de aspiración”. La regla heurística sólo nos dice qué hacer en cierto contexto, y tiene un “dominio de aplicación” dentro del cual, comparativamente hablando, es una manera eficiente de resolver problemas.

Es posible pensar esta propuesta como una mera versión de la teoría estándar y que se maximiza tomando en cuenta las alternativas que son factibles para un tipo de agente, y reconociendo las limitaciones en recursos, en tiempo y en capacidad de computabilidad. Sin embargo, y éste es el punto que me interesa destacar, hay un sentido importante en el cual la teoría de Simon no encaja con la teoría estándar. *En esta teoría no se asume que las alternativas estén dadas por adelantado, más bien, se construyen o se adquieren secuencialmente, y estas secuencias tienen un impacto en el proceso de decisión.* Éste no es un aspecto menor de la teoría de Simon. Para ver esto, quizás la manera más fácil sea hablar un poco sobre el tipo de problemas que inicialmente motivaron la teoría.

Simon propuso su modelo de decisión reflexionando acerca de la necesidad de incorporar los niveles de organización en una teoría de la decisión. Simon considera que un modelo de las decisiones racionales basado en el uso de reglas heurísticas permite superar una “aparente paradoja”. En palabras de él mismo:

The broader aim, however, in constructing these definitions of “approximate” rationality is to provide some materials for the construction of a theory of the behavior of a human individual or of groups of individuals who are making decisions in an organizational context. The apparent paradox to be faced is that the economic theory of the firm and the theory of administration attempt to deal with human behavior is at least “intendedly” rational; while, at the same time, it can be shown that if we assume the global kind of rationality of the classical theory, the problems of internal structure of the firm or other organization largely disappear. (En “A Behavioral Model of Rational Choice”, 1955. Reimpreso en *Models of Thought*, Yale University Press, 1977.)

Así, la teoría de Simon reconoce la importancia de relacionar una teoría de las organizaciones con una teoría de la racionalidad. Con ello Simon responde al desafío que formulamos en nuestra primera crítica, pero hace más serio todavía el problema para una teoría consecuencialista asociado con el segundo supuesto (la segunda crítica).

Simon y muchos de sus seguidores han sugerido que la teoría satisfaccionista puede entenderse como una teoría complementaria de la estándar. Esto requiere, como el mismo Simon sugiere, que se distinga entre dos tipos de racionalidad que desempeñan un papel en la teoría de las decisiones. Por una parte está el concepto “*substantivo*” de racionalidad (no psicologista) que captura lo que aquí llamamos teoría estándar, un concepto que, en la medida que entiende a la racionalidad como una evaluación de medios en relación con fines (en términos de consecuencias esperadas) puede hacer abstracción de la estructura psicológica y sociológica de la relación que se da entre agente y contexto, y por tanto servir de ideal normativo. Por otra parte, está el concepto “*procedimental*” de racionalidad (psicologista), un concepto que toma en cuenta las restricciones en el acceso a la información y en la capacidad de computación de los agentes, y con respecto a la cual la teoría substantiva tendría un papel normativo. Pero esta mane-

ra de ver la relación entre una teoría como la de Simon y lo que aquí llamamos la teoría estándar es problemática.

En primer lugar, la tesis de que hay tal núcleo teórico abstraible de la estructura contextual-contingente (psicológica) de los agentes *cum* medio ambiente es bastante cuestionable. Como Martin Hollis ha hecho ver en *The Cunning of Reason*, la teoría estándar de la racionalidad tiene mucho de psicología, el tipo de psicología asociada con una teoría humeana de la acción.⁸ En segundo lugar, la teoría de Simon tiene implicaciones que no son compatibles con la teoría estándar. Después de todo, *la diferencia entre un proceso de la decisión que trata de maximizar una elección y un proceso que trata de satisfacer una regla heurística tiene implicaciones para la manera como implícitamente entendemos la relación entre predicción y evaluación de alternativas*. En un modelo como el de Simon la evaluación de las alternativas no puede hacerse sólo examinando las expectativas de las posibles consecuencias de una decisión. En esta teoría no tiene por qué asumirse que las alternativas están dadas y determinadas por adelantado. La teoría de Simon permite pensar las alternativas construyéndose o adquiriéndose secuencialmente, y *permite tomar en cuenta esas secuencias con su impacto en el proceso de decisión*. De esta manera la historia de la organización de una investigación, en un laboratorio, por ejemplo, desempeña de manera natural un papel en la determinación de las alternativas y en la manera como se evalúan esas alternativas.

En la medida que las decisiones que de hecho se toman en la ciencia siguen los lineamientos de la racionalidad procedimental, la idea de racionalidad substantiva no psicologista quedaría nada más como un caso límite, pero no se justificaría considerar ese caso límite como normativo (algo que Simon parece sugerir). En todo caso tendría que mostrarse en qué sentido podemos entender los procesos de decisión modelados por una racionalidad procedimental como sujetos a

⁸ Esto es una consecuencia que ha sido muy explotada en contra de la teoría estándar como resultado de los trabajos de Amartya Sen. La tesis central del artículo de Sen (de donde también viene el epígrafe) "Is the Idea of Purely Internal Consistency of Choice Bizarre?" es precisamente que un conjunto de alternativas para elegir no puede ser interpretado como tal, hasta el punto que su consistencia o inconsistencia no es decidible, en tanto que no se incorpora algo externo a la elección; por ejemplo, los objetivos o valores que la elección busca reconocer y reforzar. El libro de Hollis fue publicado por Cambridge University Press en 1987. El artículo de Sen ha sido publicado en *World, Mind, and Ethics*, editado por J. Atham y Ross Harrison, Cambridge University Press, 1995.

la norma substantiva.⁹ De no poder mostrarse esto, la idea de racionalidad substantiva asociada con el modelo estándar de las decisiones simplemente perdería toda importancia filosófica. Para los propósitos de este trabajo todo lo que me interesa destacar es que *si las decisiones se evalúan tomando en cuenta la manera en la que se construyen secuencialmente las alternativas, entonces la pretensión de que hay un núcleo a-psicologista de la teoría estándar que puede entenderse como un ideal normativo se cuestiona de raíz.*

3. *Los problemas de la teoría estándar de la decisión con la relación entre experiencia, valor e identidad de los sucesos que tienen que considerarse en el proceso de una elección racional.* Muchas decisiones involucran a una persona, o a una comunidad, de una manera irreversible. Alguien que desea dedicarse a la filosofía en lugar de ser médico no puede ver la decisión como una elección entre dos sucesos o maneras de vivir que pueden ser evaluadas sin involucrarse de manera irreversible, y que va a “marcarlo” para toda la vida, incluso si en algún momento revertimos la decisión.

Un problema relacionado que plantean muchas situaciones tanto en la economía como en la historia de la ciencia y la tecnología es que, puesto que no están claramente definidos los estados finales que van a ser objeto de una decisión, ni son claramente distinguibles de los medios por los que pueden alcanzarse, tienen que incorporarse valores y maneras de plantearse la decisión que provienen de normas socialmente establecidas. Peter Galison, por ejemplo, ha hecho ver cómo el complejo mecanismo sociológico involucrado en el proceso de definir los estados finales que van a ser objeto de una decisión, así como la caracterización y evaluación de los medios, es crucial en la decisión de considerar terminado un experimento.¹⁰ Modelar estos complejos procesos sociológicos como procesos racionales, desde el punto de vista de la teoría estándar, es muy problemático, lo que más de una vez ha hecho pensar a los sociólogos en la “irracionalidad” del proceso. Sin embargo, si estamos dispuestos a aceptar que las decisiones racionales no excluyen el hecho que esas decisiones estén sujetas a

⁹ Entre los argumentos que muestran la dificultad que tendría que confrontar alguien que pretendiera llevar a cabo esta tarea están los de Cherniak en *Minimal Rationality*, MIT Press, 1985.

¹⁰ *How Experiments End*, The University of Chicago Press, 1987.

normas sociales, entonces sería posible hablar de ese tipo de decisiones como sujetas a criterios de racionalidad. Esto se hace plausible una vez que se reconoce que la ciencia, como la economía, no está en ningún estado definido, y que, por lo tanto, las decisiones tienen que tomarse en un ambiente en el cual las expectativas y los sucesos y alternativas que desempeñan un papel en la decisión se afectan mutuamente. Lo que dice Amartya Sen respecto de la economía es también cierto de la ciencia:

Observation is not a one-way process. Just as the decision theorist “reads” what people choose, people also “read” what is being offered. A variation of the “menu” can give us good reasons to re-evaluate the alternatives. (p. 25 de “Is the Idea of Purely Internal...”).

La teoría estándar tiene dificultades para asimilar este tipo de retroalimentación del proceso de “observación” porque no parece ser posible aceptarla y a la vez aceptar el supuesto de que todas las elecciones racionales pueden ser explicadas en término de alguna relación de preferencias que asume que esas preferencias son algo dado por adelantado.¹¹

Más de fondo, la teoría estándar (aplicada para modelar las decisiones en la ciencia) asume que hay una relación entre las probabilidades subjetivas que miden nuestras expectativas y un cierto estado del mundo; pero en la medida que ese estado del mundo es un estado que se ve afectado por las expectativas que tenemos respecto a él —de la misma manera que la economía puede afectarse por las expectativas de una devaluación—, el supuesto estándar de que las expectativas racionales pueden caracterizarse con independencia de la manera como las alternativas se constituyen como tales no es viable.

Así, entender un proceso racional de decisión, por lo menos en muchos casos importantes, requiere abandonar la tesis según la cual las alternativas pueden evaluarse en términos de sus consecuencias. Las consecuencias no preexisten a la construcción de las alternativas que tiene lugar a través de las acciones del agente. Este tipo de dificultad se exagera si tomamos en cuenta la interacción entre varios agentes. En este caso, la interacción entre las expectativas hace muy difícil

¹¹ Véanse varios de los artículos recopilados en *Elección colectiva y bienestar social*, de A. Sen, Alianza Universidad, Madrid, 1976.

esperar que un resultado sea racional en el sentido de la teoría estándar. La idea de que las decisiones políticas, o económicas, o metodológicas, son racionales en el sentido de la teoría estándar es todavía más inaceptable una vez que se toman en cuenta resultados como el famoso teorema de Arrow, que muestran que no es posible encontrar una fórmula que garantice resultados democráticos y eficientes en todos los casos, y que por lo tanto el supuesto de que es posible pensar en una teoría de las decisiones que sea aplicable tanto a las decisiones individuales como a decisiones de grupos sociales no tiene base. A diferencia de lo que piensa Laudan, no parece posible ignorar la complejidad de los procesos de decisión en la ciencia, el papel que desempeñan los procesos históricos a través del cual se plantean las bases para una decisión, ni es posible ignorar las diferencias entre decisiones individuales y decisiones de grupos.

Concretamente, en relación con la teoría de la decisión de Laudan, si bien me parece posible reconocer en la propuesta del modelo reticular de Laudan el tipo de retroalimentación entre medios y fines del que habla Sen, como Sen muestra, el reconocimiento de esta retroalimentación es incompatible con el supuesto central de la teoría estándar que Laudan comparte, el supuesto de que las decisiones pueden modelarse en términos de una postulada relación de preferencia. Es incompatible además con el supuesto consecuencialista según el cual las decisiones pueden modelarse como elecciones hechas a partir de la evaluación de las posibles consecuencias esperadas de las alternativas (que se asume generan un conjunto completo de utilidades ponderadas). En todo caso, Laudan tendría que mostrarnos cómo es posible reconciliar estas dos ideas que Sen y muchos otros economistas consideran haber mostrado que son incompatibles. Podría argüirse que esa incompatibilidad no surge en la ciencia, pero además de que tal proyecto me parece sumamente inviable, requeriría incorporar algún criterio de demarcación entre la ciencia y la no-ciencia con el que Laudan ciertamente no comulgaría.

En la parte II desarrollaré mínimamente la idea de “razón externa” y esbozaré el papel que desempeña en una teoría no instrumentalista de la racionalidad. No me interesa desarrollar aquí esta idea más allá de lo que es mínimamente necesario para hacer creíble la tesis defendida en este trabajo, que la distinción tajante entre el tema de la racionalidad y el progreso, tal y como Laudan lo propone para resolver el desafío historicista, no es aceptable. Un modelo de las decisio-

nes racionales en la ciencia no puede evitar considerar esos puntos de anclaje en la historia de los procesos por medio de los cuales se llega a una decisión, y por lo tanto no puede partir de una supuesta separación tajante entre el problema de evaluar la racionalidad de un agente individual (en términos de un supuesto conjunto completo de utilidades cognitivas ponderadas) y la racionalidad de un proceso histórico que nos interesa evaluar.

II

4. *La teoría estándar en la filosofía de la ciencia.* En la filosofía el concepto de racionalidad instrumental implícito en la teoría estándar de las decisiones está ampliamente difundido. Como Martin Hollis ha recalcado, este concepto está íntimamente ligado a toda una larga tradición de pensar la racionalidad en el contexto de una tradición humeana que tiende a modelar las decisiones como si estuvieran motivadas por deseos articulados por autointerés. La idea fundamental es que las decisiones racionales son un cierto tipo de relación entre preferencias, acciones y consecuencias, y nada más. Los fines del agente tienen que ser “internos”, puesto que de otra manera las expectativas racionales no pueden ser calculadas. Por ello, la tesis de que hay algo más que racionalidad instrumental está íntimamente ligado con la discusión en ética y epistemología respecto a si es posible elaborar el concepto de “razón externa”. Razones externas serían aquellas que “nos permiten decidir racionalmente en contra del balance de nuestras razones internas” (Hollis, *The Cunning of Reason*, p. 91). Una razón es interna si es parte de las razones que uno puede esgrimir como razones para hacer algo a través de un ejercicio del razonamiento práctico. Dependiendo de cómo entendamos razonamiento práctico esta noción de razón interna puede ser más o menos restringida. Bernard Williams ha hecho ver que esta noción de razonamiento práctico puede incorporar el papel de la imaginación en el proceso de deliberación, y con ello ha hecho ver que la noción de razón interna es bastante más amplia de como se entiende usualmente. Lo distintivo de una razón interna, sin embargo, es que es una razón que incide en el proceso de pensamiento por medio del cual un agente individual pondera (pone en la balanza) razones para determinar aquello para lo que tenemos una razón para actuar. Una

razón externa sería una motivación racional para actuar que estaría sujeta a normas de corrección, pero que no sería “controlable” por motivaciones existentes, como sucede con las razones internas. Una razón externa tendría que estar sujeta a criterios de corrección que no serían reducibles a criterios de corrección para procesos de deliberación (de un agente con respecto a cierto fin). Esto a primera vista parece imposible de concebir, y Williams sobre esta base concluye que simplemente no hay tal cosa.¹²

Sin embargo, Martin Hollis ha hecho ver que este tipo de razones “externas” entran en juego a la hora de tomar decisiones que involucran toda una vida, son el tipo de razones que se requieren cuando tenemos que evaluar racionalmente un “proyecto” de vida.

Como Hollis arguye con detalle en su libro *The Cunning of Reason*, es posible incorporar normas morales, no solamente normas consensuales o coercitivas (que pueden ser acomodadas por la teoría estándar) en una teoría de la racionalidad una vez que se reconoce la existencia de este tipo de razones.

Esta idea ha sido desarrollada también por varios epistemólogos contemporáneos y en particular por John McDowell. Él formula un concepto de “razón externa” a partir de un concepto de experiencia según el cual el mundo externo directamente nos da razones para creer o para actuar.¹³ Aceptar la existencia de razones externas en una decisión requiere aceptar que existen procesos de razonamiento o de deliberación correctos que dan lugar a una motivación para actuar, pero que no están “controlados” por las motivaciones existentes, y que por lo tanto escapan al balance de nuestras razones “internas”.

McDowell arguye que en la transición de no estar motivado a estarlo por una razón externa el agente no tiene por qué pasar por un proceso de deliberación correcto, que le haga ver las cosas correctamente. La transición no tiene por qué tener lugar a través de un proceso de deliberación correcto. Esta manera de ver las cosas —nos dice

¹² Williams hace ver en su famoso trabajo “Internal and External Reasons” (recopilado en *La fortuna moral*, UNAM, 1993), que no parece ser posible entender la transición entre un estado en el que no se está motivado por una razón externa a un estado en el que sí se está motivado, por medio de una deliberación correcta. Porque de ser esta transición el resultado de una deliberación correcta, esto parecería requerir razones internas que guiaran el proceso.

¹³ Esta idea está desarrollada sobre todo en el capítulo cuatro de *Mind and World*, Harvard University Press, 1994.

McDowell— se hace plausible si generalizamos de la ética a la epistemología la idea aristotélica de “educación” como un proceso de habituación a modos apropiados de comportamiento que están inextricablemente ligados a la inculcación de modos de pensamiento apropiadamente relacionados. Esta manera de plantearse el concepto de razón externa abre la posibilidad de entender el tipo de “conversiones” y de “reconstrucciones” de las que hablan los historicistas como un cambio *inteligible* en la orientación de las motivaciones que no es explicable como un proceso de hacer ver, por medio de procesos de razonamiento práctico controlado por motivaciones existentes, razones internas de las que el agente no era consciente

Así, en la medida que ser educado de la manera apropiada no es un camino racional a esa manera de ser, el reconocimiento de la existencia de razones externas tiene que incorporarse en una explicación de cómo llegamos a considerar las cosas correctamente que no puede reducirse a localizar razones internas, y esto nos obliga a tener que lidiar en una teoría de la racionalidad científica con el tema de las “reconstrucciones racionales” de procesos históricos.

En la medida que un razonamiento correcto responde a estándares o normas que en el sentido anterior son externos, estándares con respecto a los cuales una educación correcta consiste en “sintonizarse” respecto a ellos, una acción racional no puede ser neutral con respecto a su contenido (y por lo tanto no puede ser evaluada con criterios puramente instrumentalistas que se apliquen a relaciones medio-fin), ni puede simplemente entenderse como lo entiende un teórico estándar de la decisión, y Laudan en particular, como un procedimiento que simplemente impone coherencia y determinación práctica en un conjunto de motivaciones o intereses.

Laudan explícitamente reconoce que determinar si un agente actúa de manera tal que ese agente cree que sus fines van a ser satisfechos es una condición necesaria para adscribirle racionalidad a la acción, aunque no suficiente. Esto podría sugerir que mi crítica a Laudan sería inapropiada, que Laudan no tiene por qué considerarse aludido por una crítica a la teoría estándar de la decisión antes elaborada. Sin embargo, lo que he mostrado es que cualquier concepción viable de racionalidad tiene que tomar en cuenta la existencia de esas “anclas” en un proceso de decisión que (por lo menos en ciertos casos) pueden entenderse como “razones externas”. Estas “anclas” son puntos de referencia que toda explicación del progreso tiene también que tomar en cuenta. Son pre-

cisamente el tipo de motivaciones racionales que cumplen un papel cuando pretendemos evaluar si un proceso es o no progresivo.

Laudan implícitamente asume que no hay tal cosa como razones externas que establecerían una relación entre “nuestros fines” y los fines de Newton, por ejemplo. En todo caso, en la medida que una racionalidad no instrumental involucra el reconocimiento de razones externas, la pretensión de Laudan de que “las acciones de un agente pueden ser juzgadas como racionales sólo con respecto al conjunto ponderado de sus utilidades cognitivas” es defectuosa.

Laudan ha insistido mucho en la tesis de que, al contrario del supuesto implícito de los historicistas, la metodología no tiene implicaciones para la reconstrucción de normas racionales. Para él la pregunta crucial es si los científicos en el pasado tienen una estructura de utilidades (cognitivas) substancialmente similares a la nuestra. En la medida que esto no es el caso, la pretensión de utilizar nuestra metodología para evaluar la racionalidad de decisiones en el pasado se vuelve muy problemática. Esta tesis de Laudan proviene directamente del supuesto de que no hay razones externas que permitan explicar cómo una metodología tiene implicaciones para la construcción de normas racionales. En la medida que no es posible restringirse a un modelo que tome en cuenta razones internas para entender nuestra racionalidad, la pretensión de Laudan de cortar la conexión entre metodología y racionalidad falla.

Las dificultades con una teoría de la decisión como la de Laudan y en general las deficiencias de una teoría que no incorpora razones externas no son tan aparentes por la manera como Laudan habla de los métodos de la ciencia. Para Laudan los métodos de la ciencia son “tácticas y estrategias”, que él entiende como reglas para la elección de teorías. Con esta manera de hablar, Laudan oculta la importancia que tienen en la ciencia muchos métodos que no pueden verse meramente como reglas que guían la elección de teorías. Las técnicas para la purificación del DNA de diferentes especies no pueden entenderse como meras reglas “tácticas o estratégicas” para decidir entre teorías. Hay una gran pluralidad de fines a los que podemos asociarlas, y muchos de esos fines no tienen nada que ver con un problema de elección entre teorías. Las maneras en las cuales históricamente se van diversificando e individualizando las técnicas muchas veces desempeñan un papel importante en decisiones respecto a qué programa de investigación seguir.

Una evaluación de la confiabilidad de una técnica, por ejemplo, puede desempeñar un papel determinante en la decisión de seguir, evitar o concluir un programa de investigación, y este tipo de evaluaciones no se prestan a ser modeladas como exige la teoría estándar, como una decisión entre alternativas que depende de una evaluación de las utilidades cognitivas que se consideran posibles. La confianza en las habilidades de los posibles participantes es muchas veces más importante que una decisión entre teorías. Es precisamente en este tipo de procesos de decisión donde más claramente la noción de razón externa puede contribuir a evaluar la racionalidad de una acción.

Como ya dijimos, tomarse en serio el papel de las razones externas en un proceso de decisión involucra reconocer que la evaluación racional de una decisión requiere tomar en cuenta motivaciones racionales que dan inicio al proceso de deliberación. De esa manera puede reconocerse el papel constitutivo de los valores como “maneras de ser” históricamente distinguibles, sin que sea posible hacer una disección de esos valores en términos que sean accesibles a la teoría estándar de la decisión. No tiene por qué pensarse que debe ser posible, ni siquiera “en principio” (como lo asumiría un creyente de una racionalidad sustentada sólo en razones internas), separar la contribución de los procesos de procesamiento de motivaciones para actuar de la contribución del mundo a través de nuestros valores en la toma de decisiones. Como nos dice McDowell, el mundo no es “motivacionalmente inerte”, y esto fundamenta epistemológicamente la necesidad de incorporar en una teoría de la racionalidad científica a la historia de la ciencia, exactamente de la manera que Laudan sostiene que no puede hacerse. La historia de la ciencia puede verse no sólo como la materia prima para un catálogo de las decisiones entre teorías en el pasado, sino como una manera de aproximarnos a esa “geografía de la razón” que es necesario incorporar para entender los puntos de partida de esos procesos de deliberación que muchas veces se extienden a lo largo de siglos, y que resultan en normas y capacidades racionales.¹⁴

¹⁴ He hecho ver en otros trabajos cómo esta idea puede incorporarse en una filosofía de la ciencia que se toma muy en serio el concepto de práctica científica como un punto de partida para entender la estructura y la dinámica del conocimiento científico. Desde esta perspectiva, los conocimientos son recursos para la deliberación y la acción, y las razones externas, como las razones internas, son parte de esos recursos.

Otro problema con la teoría estándar de la decisión y que Laudan hereda es el supuesto implícito de la teoría estándar (y de Laudan) según el cual los diferentes objetivos dignos de consideración en una decisión son coherentes o compatibles de manera tal que el modelo estándar puede utilizarse. Este supuesto ha sido criticado por Sen y muchos otros filósofos en años recientes. Versiones de esta crítica provienen de trabajos muy famosos de Kenneth Arrow y Maurice Allais. Allais muestra que, al contrario de los supuestos de la teoría estándar, tendemos a forjar compromisos con ciertos resultados de una elección, y que, por ejemplo, un cambio en las probabilidades de las alternativas puede llevar a un cambio de preferencias que de acuerdo con la teoría estándar no deberían verse afectadas. Kahneman y Tversky han utilizado los trabajos de Allais para elaborar un modelo diferente de las decisiones en situaciones de incertidumbre. De acuerdo con esta teoría, hay un proceso continuo de “edición” a la manera en que las alternativas de una elección se expresan, y además reconoce que cuando hay una dependencia entre sucesos, la elección entre diferentes prospectos no está únicamente determinada por las probabilidades de los estados finales, sino que está también influida por referencias a ciertas situaciones que se consideran puntos de referencia.

De tomarse en serio un modelo como éste, el supuesto básico de la teoría estándar (al que nos hemos referido como “el segundo supuesto”) que comparte Laudan, y que le permite reducir la racionalidad a una mera relación entre medios y fines cognitivos queda en entredicho en un nivel muy básico. De acuerdo con la teoría prospectiva desarrollada inicialmente por Kahneman y Tversky *nuestras elecciones en situaciones de incertidumbre están determinadas conjuntamente por valores y probabilidades, no exclusivamente por una función de utilidad*. Mary Douglas ha extendido esta teoría, y las críticas en esta dirección a la teoría estándar de la decisión racional haciendo ver las fuertes limitaciones que tiene esta teoría cuando se trata de utilizar como guía en la evaluación de los riesgos, y por lo tanto como guía en el desarrollo de políticas de prevención, en las complejas sociedades industriales del presente:

Nonetheless, for good reasons the pure theory of rational choice has little guidance for the contemporary questions about industrial risk. The first reason is that these concerns have to do with the ends or objectives of rational behavior, and the theory, strictly speaking,

has nothing to say about ends. One could suppose it is absurd to turn to this body of theory for guidance about the public acceptability of technological risks. It can say a lot about the coherence between different subsidiary levels of choice once the major goal has been given. But we should recognize that there is a limit to the possibility of goals being made coherent with one another. Living in society forces rational beings to tolerate a lot of incoherence (en *Risk Acceptability According to the Social Sciences*, Nueva York, 1985, p. 49).

En la ciencia de hoy día, como en las sociedades industriales de las que habla Douglas, hay serios impedimentos para hacer coherentes los diferentes fines que guían la empresa científica. Decir, como dice Laudan, que el fin de la ciencia es la búsqueda de teorías confiables no puede ser más que la expresión de una convicción, que en última instancia no parece tener otra base que el supuesto de la teoría estándar acerca de la posibilidad de hacer coherentes toda esa complejidad de fines que de hecho guían el trabajo de científicos y comunidades concretas, y formular esa coherencia de una manera simple y elegante.

Más bien, parece que en la ciencia como en cualquier actividad social, mínimamente compleja, es necesario estudiar la manera como los medios y los fines se individualizan y se modifican a través de la historia de la ciencia, y cómo esos fines y esos medios se contraponen y refuerzan mutuamente, más allá de aquello que podemos modelar como un balance de razones internas, en los complejos procesos de decisión que van haciendo la historia de la ciencia.

5. *Conclusión.* Larry Laudan hace una distinción tajante entre metodología y progreso. El progreso tiene que ver con una evaluación de la historia respecto a la manera como nuestros fines se promueven. La metodología tiene que ver nada más con la racionalidad instrumental involucrada en la búsqueda de satisfacer de la mejor manera posible los fines de un agente. Como he hecho ver, ni la metodología ni el progreso pueden verse de esta manera. En la medida que una evaluación racional de un proceso involucra razones externas, la relación entre la historia de la metodología y nuestra evaluación del progreso tienen que verse más íntimamente ligadas de lo que Laudan pretende. Y si se niega la existencia de razones externas, las críticas que hemos reseñado a la teoría estándar de la decisión racional que-

dan en el camino de cualquier intento de reducir el problema de la racionalidad y el progreso en la ciencia a una concepción instrumentalista. Una parte importante del problema que plantea la caracterización del concepto de progreso científico tiene que ver con la intuición de que la racionalidad también progresa. Laudan tiene razón en que la teoría estándar de la decisión no tiene cabida para tal intuición, pero una teoría de la decisión que incorpore la noción de razón externa sí lo tiene.