

## Responsabilidad, confianza y modelos humanos \*

J. Francisco Álvarez. UNED. Madrid

Las tensiones que aparecen entre los rasgos descriptivos y normativos de diferentes modelos sobre la estructura de la conducta de los seres humanos, pueden encontrar cierta complementariedad resolutive en una perspectiva que entiende a la racionalidad como tejer dinámico en el que intervienen explícitamente los valores y las normas.

La responsabilidad personal puede entenderse como un valor cuya satisfacción (no optimización) tratan de alcanzar los individuos. Al incorporar este valor en nuestros modelos humanos se puede lograr una comprensión mejor de las prácticas sociales y de ciertas formas de génesis de la acción colectiva; para las ciencias sociales resulta interesante estudiar hasta qué punto la confianza puede ser la expresión racional de un compromiso.

La responsabilidad y la confianza pueden entenderse tanto en sentido predictivo como normativo. Dilucidar el vínculo entre un ejercicio responsable de la racionalidad y la confianza indispensable para diversas prácticas sociales, puede ayudarnos a elaborar modelos complementaristas de los individuos, que contribuyan a resolver ciertos dilemas comunes en filosofía de las ciencias sociales.

Intento recuperar aquí algunas ideas filosóficas tradicionales (por ejemplo, vinculadas al Adam Smith de la *Teoría de los sentimientos morales* –1759-) poniéndolas en conexión con otras propuestas, como las que avanzaba M. Hollis en *La confianza interna a la razón* (1987), último libro que publicó antes de su muerte, o con las que ha planteado A. Sen en relación con la inclusión de los valores y las normas en el análisis económico (Sen, A. K., 1977, 1985, 1987, 1993, 1999).

Como espero que quede claro, se trata de precisiones que están muy alejadas de distinciones como la weberiana entre ética de la convicción o ética de la responsabilidad, y a la que considero como una más de las falsas dicotomías sin

---

\*El presente trabajo se inscribe en un análisis más general sobre Capacidades potenciales, racionalidad acotada y evaluación tecnocientífica que vengo desarrollando como proyecto conjunto de investigación con México subvencionado por el Programa de Cooperación Científica con Iberoamérica de la AECL. Responsable equipo español: J. Francisco Álvarez. Responsable equipo Iberoamericano: León Olivé. Mi trabajo y discusiones en el seno del proyecto La cultura de la tecnociencia, dirigido Javier Echeverría y financiado por el Ministerio de Ciencia y Tecnología BFF2002-04454-C10-01, han contribuido a precisar muchas de las nociones que aquí se adelantan (En prensa, Isegoría número 29, diciembre 2003)

contenido filosófico serio. El breve análisis de tres aparentes dicotomías, señaladas por Sen -selección reflexiva/evolutiva, valoración directa/ indirecta y razonamiento moral/prudencial- pretenderá mostrar también cómo es posible desarrollar una visión complementarista que responda mejor a nuestras intuiciones morales, a la par que destaca algunos componentes que deberíamos incorporar a los modelos humanos que utilizamos en nuestra reflexión político-moral.

### **Emociones: de Aristóteles a la neuropsicología**

Con las palabras siguientes se abre el libro III de la *Ética Nicomáquea*: «Por lo tanto la virtud tiene que ver con las emociones y las acciones, y solamente las acciones voluntarias son las que merecen elogio o recriminación. Las acciones involuntarias se perdonan y a veces incluso sentimos pena por ellas. Por tanto parece conveniente que al estudiar la virtud definamos lo voluntario y lo involuntario; y esto también le prestará servicio al legislador a la hora de asignar recompensas y castigos». Aristóteles, *Ética Nicomáquea* III,i (1109b 30-35).

En este fragmento se nos ofrece una estupenda síntesis sobre el tema central que quiero abordar, y que considero de «enorme actualidad» para el pensamiento social. No es extraño que se hayan propuesto tantas y tan diversas traducciones de ese texto clásico, porque las virtudes, las acciones y las emociones se anudan en el centro de un problema que, permítaseme decirlo, me parece uno de los más importantes para la ciencia social: elaborar una teoría de la acción que nos permita sustentar una noción de racionalidad suficientemente abarcadora como para comprender la acción social. Para ese proyecto se han sugerido en diferentes épocas modelos muy diversos, que han propuesto variadas matizaciones sobre la articulación de los sentimientos, las pasiones y las emociones con la racionalidad humana.

Quizás sea por ello por lo que me parece que yerra cierto traductor al español (Julio Pallí Bonet, Ed. Gredos) cuando afirma que, en el texto aristotélico citado, aparece «una noción todavía confusa de la libertad psicológica, [que] se mezcla [con] una noción, no menos confusa, de la responsabilidad moral, y el autor [Aristóteles] pasa sin cesar de un aspecto al otro». Antes al contrario, defenderé que en esa supuesta ambigüedad y confusión, en esa imprecisión, se encuentra buena parte de lo más sugerente de la reflexión aristotélica para el pensamiento social contemporáneo.

Vale la pena citar, una vez más, otro texto de la *Ética Nicomáquea*, al que directa o indirectamente se ha referido en diversas ocasiones el economista A. Sen, y en

el que se nos dice que cada ámbito de investigación tiene su criterio de precisión y rigor; querer tratar un tema de manera más precisa de lo debido es una forma de incorrección del análisis. Dice Aristóteles: «Nuestra exposición será suficientemente satisfactoria, si es presentada tan claramente como lo permite la materia; porque no se ha de buscar el mismo rigor en todos los razonamientos, como tampoco en todos los trabajos manuales (...) es propio del hombre instruido buscar la exactitud en cada materia en la medida en que la admite la naturaleza del asunto» 1094b 10-25

Algo podemos avanzar en nuestro asunto, aunque no pretendamos dar demostraciones, porque, de nuevo Aristóteles: «evidentemente, tan absurdo sería aceptar que un matemático empleara la persuasión como exigir de un retórico demostraciones» (24-27)

¿Cómo definir lo voluntario y lo involuntario? Parece muy difícil precisar esa distinción, sin remitirnos a alguna suerte de dicotomía entre lo que consideramos emocional y lo que consideramos producto del logos, ya que al parecer algo se nos presenta como voluntario cuando se puede razonar sobre ello. El problema aparece una y otra vez en la historia del pensamiento occidental, ahora nos encontramos con nuevas aportaciones que proceden de las recientes teorías neuropsicológicas sobre la conexión entre emociones y razonamientos.

En Elster (2001) se resumen algunos de los estudios pertinentes para el análisis de los mecanismos causales que intervienen en las emociones y se clasifican esos trabajos en cuatro categorías: Los análisis que proceden del estudio experimental de sujetos humanos normales. El estudio de los pacientes con lesiones cerebrales que inducen patologías conductuales, emocionales y cognitivas. Los trabajos experimentales sobre la conducta animal y la observación de animales en libertad o en condiciones seminaturales.

Considerándolo dentro del segundo grupo, Elster menciona y resume el trabajo de Antonio Damasio quien, «en *El error de Descartes* ... utiliza los trabajos sobre tales pacientes [con lesiones cerebrales] para enfrentarse a la vieja concepción de que las emociones son un obstáculo para la adopción racional de decisiones. En su teoría, las emociones emergen más bien como una precondition esencial para la racionalidad. En pacientes con lesiones prefrontales ha observado cierta pasividad o neutralidad emocional y una tendencia a diferir o a no adoptar las decisiones más triviales. Concluye diciendo que nuestra capacidad para adoptar decisiones rápidas, cuando el tiempo es lo decisivo, se debe a nuestra dependencia o confianza en ciertos

‘marcadores somáticos’ o, en sentido tradicional, ‘a una cuestión de entrañas’ o ‘sentimientos viscerales’». Más adelante, en un capítulo dedicado al estudio de las relaciones entre elección emoción y adicción, sigue diciendo Elster: «Cierta número de autores han defendido recientemente que las emociones son indispensables para la adopción racional de decisiones, porque capacitan a los individuos para preparar su mente ante situaciones que son demasiado complejas para ser manejadas por el análisis racional de opciones y sus consecuencias. Por un lado, porque las emociones nos permiten evitar la falta de resolución: tomar *alguna* decisión cuando eso es lo que importa en vez de intentar adoptar la decisión óptima. Y, por otro, porque en algunos casos las emociones pueden efectivamente ayudarnos a adoptar la *mejor* decisión ... En conjunto, podemos concluir diciendo que no está probada la tesis de las emociones como elementos que mejoran la racionalidad. Puede que termine siendo cierta, pero hacen falta más datos», Elster (2001).

Efectivamente, hay mucho trabajo por hacer para mostrar la adecuación de la propuesta de Damasio, pero en conjunto me parece un enfoque más relevante que lo que señala Elster. Muy a pesar de la crítica de Elster, me encuentro muy predispuesto en favor de Damasio por algunas razones que trataré de señalar. Todas ellas muy alejadas de las correspondientes investigaciones neurológicas. Hay que decir de paso que tampoco las críticas de Elster se producen desde el interior de la neurología. Pero, para situar con precisión hasta dónde llegan las discusiones actuales, habría que señalar que el mismo Elster en un trabajo aún más reciente, en su libro *Ulysses Unbound*, Cambridge, 2000, insiste en que al menos desde Descartes se ha venido manteniendo que la prudencia o el interés propio a largo plazo puede simular a la moralidad y, ya que la moralidad se ha pensado que era más frágil que la prudencia, muchos han celebrado que está última, la prudencia, aparezca como suficiente para garantizar el orden social. Por el contrario, si uno considera que el interés propio tiene miras más cortas que la procedente de una visión a largo plazo, puede que las emociones morales sean las necesarias para simular la prudencia. Una perspectiva tal (sea la de Damasio, sea la de Elster) pondría en una seria dificultad a la dicotomía weberiana a la que me he referido anteriormente.

La habitual lectura teleológica de la ética aristotélica como opuesta a la versión deontológica de la moral kantiana (que, en definitiva, también está detrás de la oposición weberiana), no parece aceptable sin más. Es frecuente olvidar el importante papel que, para Aristóteles, tiene el carácter del individuo en la determinación de la

responsabilidad de la acción y, además, un carácter -el nuestro- cuya conformación puede estar en el seno de nuestra propia responsabilidad como agentes (tal como aparecerá claramente en la obra de Stuart Mill cuando critica al necessitarismo por confundir el determinismo con la ausencia de libertad).

### **Confianza y responsabilidad en la acción social**

Lo que pretendo sugerir es que para analizar mejor la responsabilidad vinculada a nuestros actos voluntarios resulta interesante atender a otro valor, la confianza, entendida en sus aspectos no consecuencialistas; en todo caso a una confianza adquirida en un proceso social concreto que nos exige refinar el modelo de ser humano que está siendo soporte de dichas responsabilidad y confianza: nos aparece un agente posible que no es un simple articulado racional de medios y fines.

Como ha señalado David Miller: «Un mundo en el que las personas interactúen exclusivamente mediante transacciones en el mercado no solamente es moralmente estéril sino que probablemente resulte insostenible. La cuestión es cómo prevenir que el mercado invada todas aquellas esferas de la vida social que de hecho actúan como un contrapeso de él. Recientemente diversos filósofos han comenzado a investigar las razones por las cuales ciertos bienes y relaciones humanas deben mantenerse aisladas de los mecanismos de mercado: Por ejemplo que las partes del cuerpo no deberían venderse y comprarse, o por qué los matrimonios no deberían transformarse en contratos negociados entre marido y mujer (Anderson 1993; Andre 1995; Walzer 1983). Las razones que dan son muy variadas pero una de ellas se refiere a como la exposición a los mecanismos de mercado puede cambiar el carácter de las relaciones para lo peor, por ejemplo, minando la confianza» Miller, D. (1999).

Dicotomías como las weberianas (fuese o no defendida articuladamente por Weber), colocan los temas en términos de «dar al Cesar lo que es del Cesar» y conducen inevitablemente al hiperracionalismo y a una suerte de aceptación inevitable del «fetichismo de la mercancía». Sin embargo, la conexión entre responsabilidad y confianza, y sus satisfacciones (no optimizaciones) respectivas aparecen como componentes de un posible modelo de individuo que consideramos soporte de emociones y objetivos. Si no nos consideramos responsables es muy probable que no tengamos confianza en los demás. Y si tenemos confianza es muy probable que nos consideremos responsables de nuestras acciones. Por tanto, aunque resulta

indispensable disponer de una teoría de la racionalidad, ésta haría bien en incorporar adecuadamente la cuestión de los valores y las emociones.

Una vez más Aristóteles: «Cuando ellas [las acciones] se hacen involuntariamente no se actúa ni justa ni injustamente excepto por accidente; pues entonces se hace algo que resulta accidentalmente justo o injusto». Las tensiones, que se dan y que aparecen entre los aspectos descriptivos y los componentes normativos de diferentes modelos propuestos sobre la estructura de la conducta de los seres humanos, pueden encontrar cierta complementariedad resolutive en una perspectiva que entienda a la racionalidad como un entretejer dinámico en el que intervienen explícitamente los valores y las normas, además de los correspondientes criterios consecuencialistas.

El modelo del homo economicus suele incorporar el egoísmo racional como elemento dominante de la conducta humana. Pero es conveniente recordar que incluso el mismo Adam Smith, a quien se suele pintar como una especie de profeta del capitalismo que propone entender al individuo orientado exclusivamente por el más puro egoísmo, proponía un modelo más complejo que incorporase otros valores. Planteaba Adam Smith, en el mismo inicio de *La teoría de los sentimientos morales*: «Por más egoísta que se pueda suponer al hombre, existen evidentemente en su naturaleza algunos principios que le hacen interesarse por la suerte de otros, y hacen que la felicidad de estos le resulte necesaria, aunque no derive de ella nada más que el placer de contemplarla».

Una teoría complementarista de la racionalidad, como la que pretendo proponer, incorpora elementos como la responsabilidad y la confianza (también otros valores y emociones como la envidia, el miedo, los celos, la prudencia y otras) en las diversas fases de modelización que podríamos considerar como aproximaciones sucesivas a un sujeto agente (a quien «le preocupan» sus valores), y creemos que muestra así una capacidad operativa y empírica mayor, que consideramos básica para una comprensión más adecuada de múltiples actividades (un caso particular es la misma actividad científica). «Los valores se imponen a la ciencia porque tienen mucha importancia para la ciencia», decía el recientemente fallecido W. v. O. Quine.

### **Algunos cambios en la tradición**

En conexión con estas cuestiones se perciben diversos síntomas que parecen señalar cambios en el ámbito de los estudios de la ciencia económica y que requieren especial reflexión. Uno de tales síntomas es el retorno de un problema que podríamos calificar

como el de la economía y la felicidad. Tras largo tiempo en que parecía que se había producido una crítica profunda a todo tipo de utilitarismo y a la economía del bienestar, vemos señales que indican que (como suele decirse) una vez más se había tirado el niño con el agua de la bañera. Determinados problemas, vinculados cuando no derivados de la globalización, como puedan ser la miseria y pobreza extrema, las brechas entre nacionales y emigrantes, los flujos migratorios y la conformación de espacios económicos supranacionales privilegiados, nos fuerzan a recuperar la preocupación por ese objetivo general de la felicidad. El componente consecuencialista del utilitarismo no parece que deba ser abandonado tan rápidamente. Se fortalece teóricamente lo que podríamos llamar un consecuencialismo de segundo orden, o consecuencialismo de segunda generación que, a partir de un modelo de ser humano más complejo, reformula y reorganiza buena parte de la economía clásica y la economía del bienestar.

En 1749 Ludovico Antonio Muratori publicó *Della Pubblica Felicità*, formulaba allí el concepto de felicidad pública como uno de los objetivos de la política pública que debería consistir en conseguir los mejores medios económicos para alcanzar esa felicidad (véase referencia en H. D. Dixon, «Economics and Happiness», *The Economics Journal*, 107, pp. 1812-1814, 1997). Un economista italiano contemporáneo de Muratori, Antonio Genovesi, estableció conexiones entre los cálculos individuales sobre el dolor y el placer y esa teoría de la política que buscaba la felicidad pública. Hay un matiz en las obras de los italianos de ese periodo, que ha sido puesto de manifiesto en Padgen, A. (1994) y que me interesa destacar sobremanera para matizar mi noción de confianza. Se trata de la idea de que «la inercia moral, al reemplazar el concepto de justicia por el de orden y el de razón por el de sentimiento, permite que algunos hombres, que disfrutaban del poder, impidan que otros satisfagan sus necesidades. La estructura de la costumbre sostiene el orden de clases. Una relación que se sustenta en el poder, pero un poder que no se explota mediante la fuerza, sino mediante lo que Pagano en *Saggi politici (Ensayos políticos)*, llama "confianza", *fede* (aunque distinguiéndola de "la buona fede", que para Pagano como para Genovesi, era la única base del comercio)»Padgen, A. (1994).

Vemos ahí una idea en la que se ha insistido, a mi parecer, con excesiva frecuencia y con poco rigor: la aparente vinculación ineludible entre la noción de confianza y, por así decirlo, la vida comunitaria no tamizada por la racionalidad ilustrada, por la política representativa y por los derechos individuales. Sin embargo, esa es una lectura muy sesgada que no entra a analizar el importante papel de diversos

elementos, incluso emocionales y expresivos, en la conformación de nuestra racionalidad.

Sin duda que hay muchos elementos negativos en ciertas formas de confianza, una de ellas, bastante interesante, es la que suele denominarse «trampa de las creencias». En parte se analiza en Elster, J. (2001, pág. 148)

Ante un tema de tal envergadura como el de la responsabilidad, se pone a prueba la propia manera de hacer filosofía. Desde luego hay más de un filósofo que ha considerado este tema, el de la responsabilidad, como lugar especial para mostrar su manera peculiar de entender el oficio. Este es el caso de John L. Austin, quien en sus trabajos "Tres modos de derramar tinta" y "Un alegato en pro de las excusas" Austin, J. L. (1989, caps. 12 y 8), trata de abordar este tema y lo presenta como el lugar ideal para mostrar su forma de entender la actividad filosófica. En su caso se trata de análisis del lenguaje, o como le gustaba mejor decir «fenomenología lingüística» Austin, J. L. (1989, pág. 175). Solamente decir ahora sobre Austin que su manera de interpretar a Aristóteles, planteando la responsabilidad, y por tanto las excusas, como previa a la libertad (y a la consiguiente justificación), es un camino interesante a explorar, que además le conecta con algunas de las contribuciones recientes de Hilary Putnam -véase Putnam (2001).

A pesar de la importancia de su análisis del lenguaje, no deberíamos olvidar que Austin nos remite a algo más. No sería justo rechazarle diciendo que simplemente aborda una cuestión de lenguaje ordinario, porque lo cierto es que hace algo más (y en mi opinión bastante más importante): no defiende una teoría de interfaz entre mente y mundo, sino que está hablando del mundo de la única manera que podemos hacerlo que es siendo parte de él, como ha puesto de manifiesto recientemente Putnam en su obra *La cuerda de tres cabos*. Me parece que de hecho podemos interpretar los análisis de Austin como precisiones para elaborar modelos de los individuos que actúan en el mundo.

Me parece encontrar también material conexo en las reflexiones de J. S. Mill sobre la libertad y el determinismo, en particular las que desarrolla en el segundo capítulo de *La lógica de las ciencias morales*. Aunque no pueda tratarlo aquí con todo el detalle que se merece, veamos una cita un poco larga de J. Stuart Mill:

«Y, efectivamente, si lo examinamos de cerca, encontraremos que este sentimiento de ser personalmente capaces de modificar nuestro propio carácter si **nos lo proponemos**, es en sí mismo el sentimiento de libertad moral del que somos



conscientes. Una persona se siente moralmente libre si siente que sus hábitos o sus tentaciones no son sus amos, sino que ella es quien manda sobre ellos[Se percibe cierto parecido con Aristóteles-1105a20-, un estado que podemos llegar a ser o alcanzar, llegar a ser responsables]: quien aún siguiéndoles sabe que podría resistirles; que caso de darse el deseo de separarse de ellos, para ese propósito, no se requeriría un deseo más potente que el que se sabe capaz de sentir. Desde luego es necesario, para dotar a nuestra consciencia de completa libertad, que hayamos tenido éxito en la conformación de nuestro carácter en todo lo que le hayamos intentado configurar hasta ese momento; pues si lo hemos deseado y no lo hemos conseguido, en esa misma medida no tenemos poder sobre nuestro propio carácter - no somos libres. O por lo menos, debemos sentir que nuestro deseo, aunque no sea lo bastante fuerte como para alterar nuestro carácter, es lo suficientemente fuerte como para dominar nuestro carácter cuando los dos [carácter y deseo] se enfrentan en cualquier particular caso de conducta. Y, por tanto, se dice con justeza, que nadie salvo una persona firmemente virtuosa es completamente libre.

La aplicación de un término tan impropio como el de Necesidad a la doctrina de la causa y el efecto en lo que concierne al carácter humano me parece uno de los ejemplos más destacados en filosofía del abuso de los términos, y sus consecuencias prácticas me parecen un ejemplo de los más manifiestos del poder del lenguaje sobre nuestras asociaciones. El asunto nunca será comprendido de manera general hasta que el término objetable sea eliminado. La doctrina del libre albedrío, o de la libre voluntad, al tener en cuenta esa parte de la verdad que la palabra necesidad elimina de nuestra vista, es decir, el poder de la mente para cooperar en la formación de su propio carácter, ha dado a sus adherentes un sentimiento práctico mucho más cercano a la verdad que el que se ha dado (así lo creo) generalmente en las mentes de los necesitaristas. Estos últimos puede que hayan visto mejor la importancia de lo que los seres humanos pueden hacer para conformar los caracteres de los otros; sin embargo me parece que la doctrina del libre albedrío ha alimentado entre sus defensores un espíritu mucho más fuerte de autoformación.» (J. Stuart Mill. *La lógica de las ciencias morales*, capítulo 2.3, final).

Sin ese tipo de sentimientos difícilmente aceptaríamos nuestra responsabilidad. Pero correspondientemente ese mismo sentimiento nos hace fiar de los otros, establece

nuestra confianza en las acciones de los demás. Saber que son capaces de cumplir sus promesas porque son personas responsables. Por poner un ejemplo, al plantear demandas y exigencias a los demás suele creerse que conformamos causalmente su conducta, sin embargo, es posible que eso también «se logre» como resultado no directo mediante nuestra «confianza» en los demás. Simplemente no les consideramos simples soportes de deseos y aspiraciones egoístas. Les consideramos «fines en sí mismos», «confiamos en ellos». No es una simple idea conexas con una especie de bondad absoluta del género humano, sino un mecanismo de segundo orden para la búsqueda de acciones sociales responsables en una sociedad de seres humanos libres o al menos satisfactoriamente responsables.

Interesa señalar que esa combinación de responsabilidad y confianza resulta necesaria para el funcionamiento de toda una serie de instituciones sociales, incluso para el funcionamiento del sistema de mercado. O para cualquier otro proyecto de vida en común. Sin cierto grado de confianza no podríamos funcionar en la sociedad. Incluso, el mismo tema tan de moda del riesgo y la incertidumbre, en cuanto lo analizamos un poco nos conduce a recoger parte de esta idea de la confianza -véase Giddens, A. (1990).

Un problema derivado del presente se conecta con el tema del «velo de ignorancia» propuesto por Rawls para analizar aquella situación original en la que supuestamente se consigue plantear una «interpretación procedimental del concepto kantiano de autonomía y del imperativo categórico» (Rawls, 1972). Sin embargo, un teórico de juegos como Ken Binmore, si bien acepta el procedimiento, nos dice: «El interés de la posición original descansa en que representa una versión estilizada [un modelo humano en mis términos] de los principios del compórtate como se comportan contigo, que ya están firmemente establecidos como criterios de toma de decisión conjunta dentro del sistema de convenciones comúnmente aceptadas y que mantienen a la sociedad unida»(Binmore, K. G. 1998, pág. 112). Evidentemente es una formulación más o menos elaborada de la confianza, que establece una posibilidad de definir una comparación interpersonal de utilidades.

De esta manera aparece como inevitable tejer de manera más fina el velo de ignorancia, permitirle tener componentes culturales, evolutivos y destacar la conexión con normas, valores, emociones y pasiones. Un grueso velo, mejor dicho un rudo y basto velo, detrás del cual la gente olvida las preferencias que tiene en la vida real y que los convierte a todos en iguales resuelve aparentemente el problema de las

comparaciones interpersonales. Pero lo hace porque da el problema por resuelto. Posiblemente podemos plantear que la responsabilidad como protovalor se encuentra en todas las sociedades, pero no en todas aparece como un valor conceptualizado y practicado conscientemente. Según aparezca o no los resultados de la acción serán diferentes. En este aspecto se conecta con la confianza, la confianza que tenemos en la responsabilidad de los otros. De hecho se podrían pensar así la génesis de la expectativas normativas -véase al respecto el importante trabajo de Sugden, R. (1998).

### **Hacia los tratamientos formales de la confianza y otras normas**

El conjunto de elementos morales de una sociedad constituyen buena parte de sus activos. Algunos de esos valores, por ejemplo, el rechazo de las formas de corrupción gubernamental, vienen a ser para la sociedad de mercado, utilizando una analogía propuesta por A. Sen, como el oxígeno que necesitamos para respirar: No lo echamos en falta sino cuando no está presente.

Hay un camino formalista para plantear estas cosas, mejor dicho más de un camino. Dejaré indicados dos de los que considero más prometedores. Uno se conecta con la propuesta de Sen sobre las capacidades potenciales y los desempeños, Sen, A. K. (1992), otra vía es la nueva teoría de los juegos «socialmente inmersos»(Burns, T. R. y A. Gomolinska (2000), Burns, T. R., A. Gomolinska y L. D. Meeker (2001)). Sin duda hay un tercero, que es el más desarrollado hasta ahora, y que podríamos caracterizar de manera general como teoría de juegos evolutivos –siendo Ken Binmore uno de sus más destacados cultivadores, aunque vale la pena revisar las notables contribuciones, con mayor sensibilidad filosófica, de otros autores como Skyrms (1990, 1996)-.

Como ha señalado Baron (1998), la confianza suele considerarse como un bien público, nadie suele quejarse porque haya demasiada y de su presencia se puede aprovechar incluso el desconfiado. Es necesaria para que funcionen bien las organizaciones y tiene importancia en el funcionamiento de las empresas y las instituciones. Aparece claramente en la relación de los ciudadanos con el gobierno, pues esto en último término significa confiar en los demás. También es importante en la esfera internacional, incluso para analizar problemas como los nuevos fenómenos relacionados de peligros antropogénicos de nivel mundial, los problemas de recursos internacionales, los problemas de la población y los medioambientales.

¿Qué es lo queremos promover con la confianza? La confianza en que la gente cooperará. En definitiva se trata de sugerir un modelo de acción social que se sustente

en otros criterios diferentes a los modelos que generan, de manera bastante simple, dilemas como el llamado dilema del prisionero.

La virtud de la confianza es la tendencia a comportarse como si uno creyera que los otros se comportarán de la manera correcta, y la tendencia a valorar esa conducta tanto en los demás como en uno mismo. La norma que compite con esta es la de la sospecha. Tener todo por escrito, contratar un abogado. Por ejemplo, tenemos confianza en «tener palabra», en que es una persona de palabra, algo que en cierta manera se opone al apotegma «No te fíes ni de tu padre».

No parece oportuno analizar ahora las diversas contribuciones formales que se proponen para tratar estos problemas. Pero si quisiera dejar indicado que recientemente algunos autores, básicamente procedentes del ámbito de la Sociología, han avanzado algo en lo que comienza a conocerse como teoría generalizada de juegos generalizada (TGJ). Esta teoría se apoya en la teoría matemática de las reglas y los *complex* de reglas para extender y generalizar la teoría de juegos. Tratan de afrontar el hecho evidente de que Von Neuman y Morgenstern definieron un juego como el conjunto de reglas que es capaz de describirlo pero no desarrollaron una teoría de las reglas. Desde luego resuenan las campanas del segundo Wittgenstein para «seguir una regla».

En la versión tradicional, una regla de un juego nos dice *quién* hace *qué* y *cuándo* puede hacerlo. Al mismo tiempo también nos dice *cuánto* se obtiene – planteamiento claro de la tradición en Binmore, K. G. (1992). La estructura que habitualmente se utiliza para presentar esta información es la conocida como un árbol. No es otra cosa que un caso especial de lo que se llama en matemáticas un grafo. Un conjunto de nodos o vértices, algunos de los cuales están unidos por caminos. Un árbol es un grafo conectado sin ciclos. Los modelos humanos que se han venido formulando en la teoría en esa forma tradicional, no parecían superar fácilmente a los «tontos racionales» de los que hablaba A. Sen hace ya más de treinta años.

Sin embargo, el tema de los modelos humanos, que me ha preocupado durante bastante tiempo, me parece que puede ahora ser formulado como una posibilidad más compleja en el seno de la teoría generalizada de juegos. Rompe el corsé de las dificultades que hasta ahora planteaba esa teoría, sobre todo le ofrece un mayor calado filosófico que es lo que principalmente quisiera comentar para terminar. Se relaciona expresamente con este tema de la responsabilidad y la confianza, además de con otros problemas diversos como puedan ser el problema de los agentes morales, los elementos

evolutivos o no evolutivos de esas conductas, los problemas de nucleación cultural, los espacios de relación y otros muchos. Por fin nos encontramos con una teoría formal suficientemente general que es capaz de incorporar desde la noción de Bourdieu de hábito pasando por la popperiano-boudiana lógica de la situación, además de admitir fácilmente las formulaciones procedentes de la teoría de la racionalidad acotada. En particular me parece que permite, sin ningún tipo de esfuerzos especiales, incorporar los recientes desarrollos en el campo de la racionalidad acotada, por ejemplo, los trabajos de Rubinstein, A. (1998).

Brevemente, y sin pretender desarrollar aquí planteamientos formales, me gustaría decir que la TGJ se conecta muy adecuadamente con una teoría general de los valores en la ciencia y su dinámica. La teoría generalizada de juegos trata de explicar las acciones humanas planteando que los actores realizan una acción o eligen entre diversos cursos de acción haciendo comparaciones, estableciendo las diferencias y las coincidencias entre las diversas acciones y las normas y valores que mantienen, y, «en general determinan si se satisface o no, y en que grado, un determinado valor, norma u objetivo» (Burns, T. R. y A. Gomolinska, 2000).

Por mi parte entiendo que en las prácticas tecnocientíficas (Álvarez Á., J. F. (2001), Echeverría, J. (2001)) están actuando un conjunto de matrices evaluativas que se corresponden con el conjunto de roles y modalidades que aparecen en el planteamiento de la TGJ. Me parece encontrar en la teoría generalizada de juegos, el espacio formal que puede resumir en buena medida nuestros planteamientos filosóficos al hacer un espacio razonable a nuestras demandas prácticas y morales.

Por ejemplo, para comprender la génesis y las prácticas de las mafias, aunque sean fenómenos muy complejos y para cuyo estudio debemos aprovechar resultados procedentes de muy diversos campos de los estudios sociales, no resulta ocioso que nos preocupemos por las formas de aceptación y articulación de la responsabilidad y la confianza. Pero resulta imprescindible disponer de un modelo o adoptar una perspectiva complementarista de la racionalidad, en línea con algunas sugerencias planteadas por Amartya Sen al analizar el papel substantivo de las normas y los valores en el proceso de elección, incluso teniendo en cuenta los resultados procedentes del análisis evolutivo sobre la génesis de las normas.

Como han indicado Avner Ben-Ner y Louis Putterman: «La necesidad de incorporar los valores en el análisis económico se ha manifestado con fuerza en un área

de la economía que ha trabajado con la forma más pura del homo oeconomicus, la teoría de juegos»(Ben-Ner, A. y L. Putterman, 1998).

Como ellos mismos recuerdan en la introducción a la importante compilación sobre estos temas, que han editado recientemente: «Afrontar el problema sólo desde la racionalidad se ha mostrado insuficiente, se ha visto como necesario tener en cuenta las consideraciones sobre las reglas del juego, sobre las instituciones que las producen y las motivaciones de los participantes que van más allá de su interés propio... Así los trabajos de Kahneman han tratado de incorporar en sus modelos conceptos tales como el sentido de equidad o justicia que los individuos incorporan en sus interacciones. Se ve que las reglas del juego tienen mucha importancia sobre los resultados y en la formación de las instituciones. Los economistas han avanzado poco en el análisis de muchos problemas agudos del presente que parecen tener un fuerte contenido económico, precisamente porque han excluido los valores del análisis económico. Esperamos romper viejos tabúes disciplinares y vincular la perspectiva económica de la elección racional con el análisis de la formación y cambio de los valores» (Introducción a *Economics, values and organization*).

### **Una defensa de la complementariedad**

¿Cómo incluir y acomodar los valores y las normas en el análisis económico? Al respecto ha hecho A. Sen algunas observaciones que me parecen muy importantes. Es frecuente plantear determinadas dicotomías que, en opinión de Sen, lo que realmente nos exigen es estudiar las *complementariedades* que puedan existir entre elecciones supuestamente exclusivas. Amartya Sen señala tres de esas supuestas dicotomías que es preciso superar. Se trata de analizar la conexión entre la selección reflexiva y la selección evolutiva, la complementariedad entre una valoración directa y otra indirecta y la importancia de tener en cuenta simultáneamente los aspectos éticos y los prudenciales. Veamos en líneas generales, siguiendo a Sen, estos tres tipos de complementariedad.

#### **a) Complementariedad entre selección reflexiva y selección evolutiva.**

En la moderna teoría de juegos evolutivos, la opción social aparece teniendo en cuenta su papel en la supervivencia y la selección se produce por la consideración posterior de los resultados que se dan en el mundo globalmente considerado. Sin embargo, es importante tener en cuenta que las normas y los valores están sujetos a la reflexión y a la selección racional, reflexionamos y seleccionamos racionalmente las normas y los

valores, por lo tanto los antecedentes que se den en la mente de cada individuo tienen efecto en el resultado de la acción social. La contribución de la filosofía moral ha sido precisamente la de sugerir cómo razonamos sobre las normas y cómo actuar de acuerdo con ellas. «Las alternativas organizativas e institucionales se seleccionan simultáneamente a y conjuntamente con las regularidades evaluativas» Sen, A. K. (1998).

Está claro que los resultados de la selección evolutiva son importantes, pero «¿el proceso de selección debe tener en cuenta solamente ese mecanismo (con la selección consciente reducida a un simple refrendo de la selección natural? ¿Por qué no pueden los dos medios de selección actuar activamente?». Algo muy cercano a esto lo ha planteado Elliot Sober. Los dos procesos pueden actuar cada uno por separado pero puede que incluso trabajen conjuntamente: «el egoísmo (psicológico) no merece ser considerado como la hipótesis por defecto que deberíamos adoptar en la medida en que resulta consistente con lo que observamos ... Aunque sea en un grado pequeño, el peso de la evidencia favorece al pluralismo ... Es menos probable que haya evolucionado un conjunto de motivos puramente egoístas que uno que incluya deseos últimos tanto egoístas como altruistas» Sober, E. (1998).

De manera que la situación en vez de apoyar el «dogma» de la incomunicación entre la reflexión crítica y la selección evolutiva, más bien señalaría a que si bien la reflexión crítica no es inmune a la selección evolutiva, «tampoco la selección evolutiva convierte a los seres reflexivos en meros autómatas carentes de pensamiento» (Sen).

Tradicionalmente la filosofía moral no ha tenido en cuenta los procesos evolutivos, quizás sea esa actitud la que ha llevado a algunos teóricos a proponer «deskantianizar» a los filósofos morales contemporáneos, en particular a John Rawls, con quien, por otra parte, suelen coincidir en diversos aspectos -véase el planteamiento que se hace en Binmore, K. G. (1994). Posiblemente esto indica más bien que resulta indispensable complementar la selección reflexiva mediante el análisis evolutivo. «Las reglas con las que vivimos no pueden ser inmunes a nuestro razonamiento crítico, lo mismo que no pueden dejar de verse influidas por la selección evolutiva. El reconocimiento de una influencia no elimina la otra» (Sen, A. K., 1998).

#### **b) Complementariedad entre la evaluación directa e indirecta**

Otra segunda dicotomía que es más aparente que real se refiere a si debemos tener en cuenta exclusivamente los efectos inmediatos o los no mediatos e indirectos. Los

efectos indirectos son los que principalmente han estudiado los teóricos de los juegos evolutivos, tienen que ver con ese conjunto de elementos que los sociólogos o políticos suelen llamar efectos no deseados de la acción o subproductos de la acción social. Sin embargo, la existencia de esos efectos indirectos, con toda la importancia que puedan tener a largo plazo, no excluyen que al actuar de acuerdo con normas y valores podamos tener interés además en efectos directos, ya sea por una posición prudencial o una moral.

Al actuar de acuerdo con normas y valores, podemos estar interesados en efectos que pueden ser indirectos y no inmediatos, además de los resultados inmediatos (que pueden incluir la satisfacción moral directa, o méritos prudenciales conexos como el placer de ser bien considerado por los otros). «Por ejemplo, podemos no comer la última manzana de la fuente porque pensamos que esa restricción es la manera correcta de comportarse (nos apoyamos en un argumento moral directo) o porque no nos guste la manera en que nos mira la gente si nos apresuramos a comernos la última manzana (una preocupación inmediata prudencial)».

Como dice Sen: «Incluso si la respuesta a la pregunta ¿Por qué nos preocupamos por lo que otros piensan de nosotros? Se puede explicar en términos de las ganancias materiales que obtenemos al ser bien visto por los demás, esto no disminuye ni impide que sea razonable preocuparse directamente por lo que otros piensan de nosotros (de hecho es confortable ser bien visto e incluso puede darnos alguna confianza moral de que no nos estamos comportando mal). Los efectos indirectos complementan más que suplantando los directos».

### **c) Complementariedad entre ética y prudencia**

Muchos de los trabajos sobre el rendimiento económico de los códigos empresariales, han mostrado que, en ciertos casos, aquellos con mejores resultados suelen ser los que no tienen como objetivo principal e inmediato la búsqueda simple del rendimiento económico; pueden servir de ejemplo de estudios que logran explicar el papel prudencial de muchas normas éticas. Vendría a suponer una explicación instrumental de las normas éticas. Pero, como señala Sen, esa consecuencia no se sigue de tales estudios: «Consideremos que aceptamos completamente una demostración de que incluso si la gente fuera completamente amoral (en todo lo que puede incluir el pensamiento intencional), todavía podrían emerger las reglas de conducta moral (a través de efectos indirectos, selección natural, etc.) Esto sería un descubrimiento



importante, pero no sería lo mismo que establecer que las personas son de hecho amorales en su pensamiento y elección. Un ejercicio hipotético no puede establecer una regularidad empírica. Los dos procesos pueden actuar cada uno por separado pero puede que incluso actúen conjuntamente. Es importante ver cómo y por qué el proceso prudencial puede operar con o sin razonamiento moral, pero esto no elimina el papel efectivo del razonamiento moral mismo». De nuevo aparecen claras conexiones con la idea defendida por Elliot Sober, que ya hemos comentado.

### **Racionalidad sintética y moralidad**

En mi opinión se trata de comprender la racionalidad como el resultado de un proceso conformado tanto por la trama de la racionalidad instrumental como por la urdimbre de la racionalidad expresiva, al menos por la urdimbre de algo que comparte muchos puntos con lo que algunos han llamado racionalidad expresiva (Hargreaves Heap, S. (1989)). En algunos casos los componentes expresivos pueden formularse como restricciones de la racionalidad instrumental, en otros casos la racionalidad expresiva debe entenderse como una ampliación del contenido informativo. La cuestión no consiste en utilizar una noción instrumental para unas cosas y una perspectiva expresiva para otras, más bien se trata de pensar en la complementariedad sistemática Álvarez Á., J. F. (2002).

Ambas formas de racionalidad se dan en el mismo individuo, no aparecen como modelos alternativos de individuo. Al esclarecer los elementos de integración y complementariedad tratamos de formular un agente integrado, más adecuado en sus aspectos descriptivos y más preciso predictivamente.

Ahora bien, tanto los elementos instrumentales como los expresivos parecen dejar toda la explicación a lo que podríamos llamar el "software" de nuestro modelo humano. Sin embargo, hay elementos que podríamos denominar "hardware" de nuestro modelo que deberíamos también tener en cuenta. No es sólo el tejido (la trama y la urdimbre) lo que hay que tener en cuenta, también importa el telar. Este lugar vendría a ocuparlo buena parte de lo que conocemos sobre las emociones, procesos causales que producen determinados resultados aunque no conozcamos bien los procesos de composición de esos mecanismos. En este aspecto las investigaciones más recientes de J. Elster me parecen significativas al vincular la acción con los procesos neurofisiológicos, con nuestra forma peculiar de procesar la información y nuestros mecanismos de reacción.

Es decir, en esta línea quisiera defender una noción que podríamos llamar racionalidad acotada, procedimental, por satisfacción y expresiva.

### **Amartya K. Sen entre la ética y la economía. Cultura y confianza**

En las palabras finales con las que cerraba su discurso de aceptación de uno de los al menos treinta doctorados *honoris causa* con los que ha sido homenajeado, el otorgado en 1995 por la Universidad de Valencia (España), plantea Amartya Sen un elemento clave para comprender sus análisis económicos: "Los códigos morales son parte integral del funcionamiento económico, y pertenecen de manera destacada a los recursos sociales de una comunidad. La economía moderna ha tendido a abandonar totalmente estos aspectos de los sistemas económicos. Hay buenas razones para intentar cambiar ese abandono y reintroducir en la corriente principal de la ciencia económica este componente crucial de la actividad de una economía. Efectivamente, queda mucho por hacer".

En sus trabajos Sen utiliza el análisis económico con una clara apertura a cuestiones más generales sobre los valores, la cultura y la libertad. Ya es suficientemente aceptada la importancia de sus trabajos empíricos sobre el hambre, su aportación al estudio de la agregación de los valores individuales en el estudio de las decisiones colectivas. Como el mismo ha dicho, ha empleado la mitad de su vida en el estudio de la teoría de la elección social pero sin olvidar que la cultura determina cómo las personas ganan y gastan su ingreso. "La cultura puede contribuir a nuestra comprensión de los patrones de conducta, el capital social y el éxito económico. Creo que la influencia de la cultura se puede transferir y que podemos aprender de ella (...) La cultura no es algo que está ahí fuera como las Pirámides, sino que es un proceso, dinámico y constructivo, con actividades de emulación e imitación. Es importante estudiar cómo se forman y se transforman los valores, como cambian e interactúan con otros valores y culturas" (Sen, A. K., 1999a).

Estas actividades, formativas, culturales en sentido bien amplio, «el respeto a las normas», «el desarrollo de la legalidad republicana», son las que en nuestra época pueden generar confianza ( desgraciadamente en otras pudieron complementarse con «pan y circo», y más recientemente pudiera ser el *american way life* «hamburguesas y Gran Hermano») y pueden ayudar a que efectivamente (y no sólo contrafácticamente) podamos utilizar el modelo de la posición original y el velo de ignorancia.

## Bibliografía citada

- ÁLVAREZ Á., J. F. (2001): "Capacidades potenciales y valores en la tecnología. Elementos para una axiología de la tecnología" en LÓPEZ CEREZO, J. A. y J. M. SÁNCHEZ RON (comps.) (2001), Biblioteca Nueva, Madrid, pp. 231-242.
- ÁLVAREZ Á., J. F. (2002 (en prensa)): "El tejido de la racionalidad acotada y expresiva" en BEAUMONT, M. (comp.) (2002 (en prensa)), Universidad de Campinas. Brasil, Campinas.
- AUSTIN, J. L. (1989) -edición de Urmson, J. O Warnock-: *Ensayos filosóficos*, Alianza, Madrid.[original 1961, v.cast. de GARCÍA SUÁREZ, A.].
- BARON, J. (1998): "Trust: beliefs and morality" en BEN-NER, A. y L. PUTTERMAN (comps.) (1998), Cambridge University Press, Cambridge.
- BEN-NER, A. y L. PUTTERMAN (comps.) (1998): *Economics, Values and Organization*, Cambridge University Press, Cambridge.
- BINMORE, K. G. (1992): *Fun and games. A text on Game Theory*, D.C. Heath and Company, Lexington, Mass.
- BINMORE, K. G. (1994): *Playing Fair: Game theory and the social contract*, MIT Press, Cambridge, Mass.
- BINMORE, K. G. (1998): "A utilitarian theory of political legitimacy" en BEN-NER, A. y L. PUTTERMAN (comps.) (1998), pp. 101-132.
- BURNS, T. R. y A. GOMOLINSKA (2000): "The Theory of Socially Embedded Games: The Mathematics of Social Relationships, Rule Complexes, and Action Modalities", *Quality & Quantity*, **34**, pp. 379-406.
- BURNS, T. R., A. GOMOLINSKA y L. D. MEEKER (2001): "The Theory of Socially Embedded Games: Applications and Extensions to Open and Closed Games", *Quality & Quantity*, **35**, pp. 1-32.
- ECHEVERRÍA, J. (2001): "Ciencia, tecnología y valores. Hacia un análisis axiológico de la actividad tecnocientífica" en IBARRA, A. y J. A. LÓPEZ CEREZO (comps.) (2001), pp. 137-148.
- ELSTER, J. (2001): *Las pasiones: emoción, adicción y conducta humana*, Paidós, Barcelona.[original 1999, v. cast. de ÁLVAREZ Á., J. F. y A. KICZKOWSKI].
- GIDDENS, A. (1990): *The Consequences of Modernity*, Polity Press, Cambridge.
- HARGREAVES HEAP, S. (1989): *Rationality in Economics*, Blackwell, Oxford.
- HARGREAVES HEAP, S. y Y. VAROUFAKIS (1995): *Game Theory. A Critical Introduction*, Routledge, Londres.

- HOLLIS, M. (1987): *The Cunning of Reason*, Cambridge University Press, Cambridge.
- IBARRA, A. y J. A. LÓPEZ CEREDO (comps.) (2001): *Desafíos y tensiones actuales en ciencia, tecnología y sociedad*, Biblioteca Nueva/O.E.I., Madrid.
- LÓPEZ CEREDO, J. A. y J. M. SÁNCHEZ RON (comps.) (2001): *Ciencia, tecnología, sociedad y cultura en el cambio de siglo*, Biblioteca Nueva/O.E.I., Madrid.
- MILLER, D. (1999): "Market, ethics of the" en CRAIG, E. y L. FLORIDI (comps.) (1999).
- PADGEN, A. (1994): "Francesco Maria Pagano's 'Republic of Virtue' Naples 1799" en FONTANA, B. M. (comp.) (1994), Cambridge University Press, Cambridge.
- PUTNAM, H. (2001): *La trenza de tres cabos. La mente, el cuerpo y el mundo*, Siglo XXI de España, Madrid.[original 1999, v. cast. de ÁLVAREZ Á., J. F.].
- RUBINSTEIN, A. (1998): *Modelling Bounded Rationality*, MIT Press, Cambridge.
- SEN, A. K. (1977): "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory" en HARRIS, H. (comp.) (1977), 1979. Apareció inicialmente en *Philosophy and Public Affairs*, 6.
- SEN, A. K. (1985): "Well-being, Agency and Freedom: The Dewey Lectures 1984", *The Journal of Philosophy*, **82**, n° 4, pp. 169-221.
- SEN, A. K. (1987): *On ethics and economics*, B. Blackwell, Oxford, UK ; Nueva York, NY.
- SEN, A. K. (1992): *Inequality reexamined*, Russell Sage Foundation ; Harvard University Press, Nueva York Cambridge, Mass.
- SEN, A. K. (1993): "Positional Objectivity", *Philosophy and Public Affairs*, **22**, 2, pp. 126-145.
- SEN, A. K. (1998): "Foreword to" en BEN-NER, A. y L. PUTTERMAN (comps.) (1998).
- SEN, A. K. (1999a): *Development as freedom*, Oxford University Press, Nueva York.
- SEN, A. K. (1999b): *Reason before Identity*, Oxford University Press, Oxford.
- SKYRMS, B. (1990): *The Dynamics of Rational Deliberation*, Harvard University Press, Cambridge, MA.
- SKYRMS, B. (1996): *Evolution of the Social Contract*, Cambridge University Press, Cambridge.
- SOBER, E. (1998): "El egoísmo psicológico", *Isegoría*, **18**, pp. 47-70.
- SUGDEN, R. (1998): "Normative expectations: the simultaneous evolution of institutions and norms" en BEN-NER, A. y L. PUTTERMAN (comps.) (1998), pp. 73-100.

