

CIENCIA  
PENSAMIENTO  
Y CULTURA

# *arbor*

VOLUMEN CLXXXII

Nº 718

marzo-abril [2006]

MADRID [ESPAÑA]

ISSN: 0210-1963



MINISTERIO  
DE EDUCACIÓN  
Y CIENCIA



Consejo Superior  
de Investigaciones Científicas

**CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS**

# ¿ES POSIBLE SER MORALMENTE RESPONSABLE?

## Notas para una nueva definición del concepto de sujeto\*

Vicente Raga Rosaleny

*Departament de Metafísica i Teoria del Coneixement. Universitat de València. Phronesis-analytic philosophy group  
vicente.raga@uv.es*

**ABSTRACT:** *The Strawson's Basic Argument is the stronger against moral responsibility in Philosophy of action. One should be responsible of his identity to be moral responsible of his actions, but then nobody could be never responsible.*

*In this article I criticize orthodox solutions to Strawson's sceptical challenge and show how they share with the Argument the same theological notion of monadical agent. A new solution needs a new conception of agent.*

**KEY WORDS:** *Philosophy of Action. Moral Responsibility. Agent. Compatibilism. Incompatibilism. Determinism. Free Will. Social Identity.*

El Argumento Básico de Galen Strawson es uno de los más potentes que se hayan formulado en el campo de la filosofía de la acción en contra de la responsabilidad moral. Este autor exige una responsabilidad por los rasgos constitutivos del agente, para poder dar cuenta de la responsabilidad moral por las propias acciones, que convierte tal atribución en algo imposible de justificar. En este artículo discutiré el argumento strawsoniano, mediante tres pasos preliminares que me permitirán, como conclusión, afirmar la posibilidad de tal responsabilidad siempre que se dote al sujeto de una nueva definición, diversa de lo que tienen en común todas las otras.

En primer lugar, intentaré establecer una posible génesis de la noción de individuo que está en la base del Argumento Básico y da toda su potencia a la argumentación escéptica strawsoniana. Luego expondré ese Argumento para, en tercer lugar, exponer los intentos de solución que desde el compatibilismo e incompatibilismo se han venido formulando. El fracaso de todas las soluciones ortodoxas me permitirá mostrar los supuestos compartidos tanto por el argumento escéptico, como por sus intentos de respuesta, a saber, la noción de individuo expuesta en el primer paso de nuestra exposición preliminar.

**RESUMEN:** Una de las objeciones más fuertes contra la responsabilidad moral en filosofía de la acción sería el Argumento Básico de Galen Strawson, que plantea la necesidad de que uno sea responsable de ser quien es para poder ser responsable moralmente de sus acciones, exigencia ésta harto difícil o imposible de cumplir por lo que, en consecuencia, nadie sería nunca responsable.

Los diversas respuestas ortodoxas al reto escéptico de Strawson comparten con el Argumento Básico una misma noción teológica heredada de agente-mónada y, según trato de mostrar en mi artículo, sería necesaria una nueva concepción del sujeto para dar cuenta correctamente de este Argumento.

**PALABRAS CLAVE:** Filosofía de la acción. Responsabilidad moral. Agente. Compatibilismo. Incompatibilismo. Determinismo. Libre albedrío. Identidad social.

Llegados a este punto, mi exposición cambiará de rumbo e intentará, consciente de la historia que hay detrás de la noción de agente-mónada que subyace a los intentos de respuesta y al propio problema expuesto, proponer una concepción de sujeto que dé cuenta de todos los problemas del anterior y, en tanto que histórica y socialmente configurada, polémica con el esquema de la propuesta strawsoniana. No sería necesario que el agente fuera *causa sui*, para poder ser moralmente responsable de sus acciones, porque tal noción de agente es incoherente y no se compadece con una definición del ser humano, como encrucijada donde sociedad e historia se unen productivamente.

### I

La imagen que del agente se desprende de la tradición dominante en el campo de la filosofía de la acción actual no es la única posible. Es más, las distinciones que tal imagen comporta, los compromisos metafísicos que mantiene, tornan sumamente implausibles las atribuciones de responsabilidad moral y la defensa de la libertad que

habitualmente van ligadas a ella, ante argumentos que cuestionen las distinciones externo/ interno, heterónimo/ autónomo, así como sobre los aspectos sociales, históricos y de contenido, que constituyen al agente.

La definición de individuo, de su libertad y responsabilidad moral asociadas, que maneja esta tradición, podría tener sus orígenes más remotos en un pensador medieval Pedro Abelardo. Como paradigma de un cambio sociocultural relevante que tuvo lugar en su época ha de entenderse la exhortación al conocimiento de uno mismo que aparece en sus escritos éticos. La idea principal, que caracterizaría a Abelardo como representante del movimiento espiritual que en el siglo XII habría coadyuvado a ese temprano "surgimiento del individuo" era: que la moralidad de un acto se define sólo por la intención de quien lo realiza, lo cual enlaza bien con la noción teológica cristiana que anima su concepto de voluntariedad.

El pecado como consentimiento interno a la inclinación a realizar una acción mala, implicaría siempre, según se desprende de la doctrina abelardiana, un suceso psicológico específico, una decisión y, en tanto que rechazo consciente de la voluntad divina, sería un acto de plena afirmación de la propia individualidad.

Este carácter interno y aislado de la decisión, el *consensus*, implicaría, alejándose de otros modelos como el aristotélico, por un lado la separación entre el carácter moral o *ethos* y la decisión moral: el *consensus* no es necesitado por ninguna inclinación antecedente. Y, por otro lado, el corte entre la decisión moral y la acción externa<sup>1</sup>. En ese espacio vacío entre el *carácter* y la *acción*, estaría la decisión moral y, por tanto, el individuo.

## II

Con esta breve exploración por la historia de la filosofía querría, pues, haber mostrado los vínculos históricos observables entre la concepción abelardiana del sujeto y el individuo atomizado, reducido a mera estructura decisoria con creencias, deseos y valores ("creseos"), desconectados éstos (y vaciados de contenido), por mor, en principio, de una apropiada atribución de responsabilidad, de la sociedad y de la historia (entendidos éstos más

bien como amenazas condicionantes que como condiciones posibilitantes), en la filosofía de la acción contemporánea.

Sin embargo, también hay diferencias notables entre ambas definiciones del sujeto. Los avances de la ciencia moderna en ámbitos diversos, y los cambios ideológicos variados propios de la Modernidad han llevado a una visión naturalizada del mundo y del ser humano, con implicaciones metafísicas, ontológicas y antropológicas varias, que separan nitidamente al modelo abelardiano de aquellos que imperan en el ámbito que nos ocupa. Responsabilidad moral y libre albedrío o voluntad, siguen ligadas, y permanecen todos los elementos indicados en la caracterización del *consensus*, pero la divinidad, contra cuyo trasfondo se dibujaba la libertad humana plena, parece haberse ido emborronando hasta desaparecer de la composición pictórica.

Pero, ¿verdaderamente se ha ido? En un mundo en el que, en virtud de los avances científicos, no hay proposición general más ampliamente compartida, con mayor apoyo inductivo y empírico, que la de que todos los sucesos, entendiéndolos en sentido amplio, tienen explicación, cabe preguntarse por el lugar del agente, instancia de la voluntariedad, decisión y responsabilidad. Si por explicación se entiende la explicitación de las circunstancias que hicieron necesario el suceso acaecido, y si tales sucesos, en un mundo naturalizado, se insertarían en cadenas o secuencias causales (epistemológicas pero, de modo especial, también ontológicas), donde un estado de cosas causa otro, que a su vez es causado por uno anterior, y así sucesivamente, ¿en qué se diferenciaría el mero suceder, o causalidad transeúnte, del causar la acción del agente, o causalidad inmanente?<sup>2</sup>

Y la respuesta a esta pregunta sería que, precisamente, el agente se diferenciaría del mero suceder en el cumplimiento de las condiciones de la acción y voluntad libres, que darían cuenta de la subsiguiente atribución de responsabilidad, en la estela de la tradición dominante, y que serían esencialmente dos: que esté en poder del agente el realizar o no la acción; que esa "capacidad" resalte sobre un amplio trasfondo de deseos, de posibilidades no actualizadas, pero actualizables (el Principio de Posibilidades Alternativas) y que el agente constituya el origen de esa acción, deseo, volición, y no sea otra cosa ajena a él la que

tenga el "control" (el Principio de la Originación o, como veremos luego, de la Responsabilidad Última).

En ese mundo tal y como lo he caracterizado, y centrándome en el requisito de la originación, por mor del argumento que voy a desarrollar, parecería necesario que uno fuera causa incausada de aquellas acciones que propiamente pueden atribuírsele, responsabilizándose por ellas, en tanto que acciones libres, que le pertenecerían, que controlaría, sin que otra cosa las causara, sin otra fuente de originación. Pero (y aquí retorna el dios emborronado), motor inmóvil, causa incausada, y demás epítetos se han venido atribuyendo, esta vez sí desde Aristóteles, a la divinidad, que se habría encarnado, pues, ahora en el hombre aparentemente naturalizado.

Chocan aquí dos intuiciones ampliamente extendidas, de sentido común, dos fuentes diversas de conocimiento y tradición, dos conjuntos de instituciones sociales, macroestructuras, etc. Por un lado, el punto de vista práctico, subjetivo, que da cuenta de una capacidad humana central, la de influir en el mundo y ser origen de las propias acciones, voliciones, y que se acompañaría de las más diversas actitudes reactivas y prácticas asociadas, en la interacción común (sustentando, además, no sólo importantes instituciones, sino el propio sentido de una vida valiosa). Por otro, el punto de vista objetivo o teórico, antes apuntado, ligado al avance de las ciencias, que dibujaría un panorama donde tal capacidad se diluiría o encontraría seriamente amenazada.

Más allá del rotulo "concepción heredada del determinismo", que se desprendería de la caracterización que del mundo naturalizado he hecho, donde las acciones y elecciones humanas se entenderían como efectos de las secuencias causales en que, más ampliamente, se integran todos los sucesos del mundo<sup>3</sup>, el problema de ese segundo punto de vista, tal como lo plantean los sucesores del determinismo fuerte clásico (aleccionados por el propio avance de la ciencia, por ejemplo en el campo de la física, con los avances de las diversas formulaciones de la física cuántica) sería el de una cuestión de incompatibilidad generalizada. La libre voluntad implicaría verdadera responsabilidad y tan incompatible sería con el determinismo como con el indeterminismo, siendo por tanto inviable su defensa, pese a nuestro primer punto de vista, si cupiera, como estos autores defienden, aplicar a tal dilema el tercio excluso.

No sería, pues, en realidad, necesario decidirse por uno de ellos (lo que supondría eludir las cuestiones relativas a la plausibilidad de una solución de síntesis, entre determinismo e indeterminismo<sup>4</sup>, así como las argumentaciones relativas a la relevancia y sentido de las reflexiones derivadas de los nuevos campos abiertos por la física, la neurología, entre otros), como muestra uno de los argumentos más relevantes de los últimos tiempos en esta línea "dura", y contra el que trato de articular este ensayo: el Argumento Básico de Galen Strawson.

Según se explicita en el Argumento Básico<sup>5</sup>, pues, no seríamos responsables moralmente de nuestras acciones, sea el determinismo verdadero o falso, dado que nadie podría cumplir el requisito de Responsabilidad Última (que postulamos como una de las dos posibles condiciones de la responsabilidad moral). Las acciones libres, que impliquen verdadera responsabilidad, ya lo he apuntado, lo serían en virtud de una razón del agente que las causa; esto es, en función de cómo es uno mentalmente. Pero no querriamos, ésta es la baza de Strawson, que esas cadenas causales que llevan de nuestros deseos, creencias y valores a nuestras voliciones y acciones, fueran nuestras de modo azaroso, sin bases, por suerte (entendida como *lack of control*, no como *chance*, y del tipo de la que podríamos denominar suerte constitutiva), ni en virtud de algún tipo de control externo al agente<sup>6</sup>. En suma, el requisito para que las acciones y razones fueran propiamente del agente es que éstas se eligiesen en virtud de unos principios, de los que, a su vez debería ser responsable, mediante otros principios de elección, y así *ad infinitum*.

Si lo enuncio primero de un modo menos riguroso cabría decir que no estamos ante una cuestión epistemológica, esto es, que no cabría bloquear el Argumento, y la cuestión planteada de la imposibilidad de la responsabilidad moral, apelando a un desconocimiento, o al posible estatuto inalcanzable de una completa explicación del comportamiento humano<sup>7</sup>, de sus razones para actuar (pues sobre ese eje gira el Argumento de Strawson, y a ese punto dirigen sus miradas todos aquellos que atienden a la condición del Origen). Ni tampoco atendiendo a una posible salida por vía indeterminista, entendida ésta como puro azar<sup>8</sup>, ya que la acción y decisión impredecibles, carentes de bases, parecen en primera instancia sí libres, pero difícilmente responsables (y dotados de una libertad valiosa), ya que ésta, la responsabilidad por las acciones y decisiones, es algo que responde a mis razones, guías de mi elección, y es fruto de los rasgos

(creencias, deseos, motivaciones) que estrechamente me constituyen, bases de mi decisión.

Y ésto cuestiona la tercera posibilidad de escape, la apelación a razones, de las que requeriríamos poder causal, y cuyas explicaciones, la explicación por razones, equipararíamos, siguiendo a Davidson<sup>9</sup>, con la explicación causal (con lo que esta argumentación se inclinaría del lado del determinismo, pretendidamente compatible con la libertad y responsabilidad). Así pues, si hubiese una explicación causal de mi elección, con base en mis razones, educidas de mis rasgos constitutivos, y nos centramos en la condición del Origen o de Responsabilidad Última, el problema sería el de la responsabilidad por el propio carácter, el de si somos fuente de esas cadenas causales en que consisten nuestras razones.

Por supuesto que cabe, como arguye Dennett<sup>10</sup>, frente a las exigencias de ultimidad o control absoluto, de autodeterminación como autocreación de corte teológico, apelar a la "complejidad", del agente, a su ser sensible a razones, a su capacidad para la revisión y alteración del carácter basándose en la reflexión<sup>11</sup>. Pero ésto sería quizá tomarse las cosas con demasiada ligereza, o pretender recluir en el campo de las meras abstracciones teóricas lo que constituye un problema metafísico que sí afectaría a nuestras conductas e instituciones, y que los avances de la genética, neurología, sociología, etc, plantean como cercano y cotidiano.

Sin ultimidad, por un lado, ni el mecanismo sensible a razones mencionado por Dennett (y argumentado por los semicompatibilistas Fischer y Ravizza)<sup>12</sup> habría sido adquirido por medios controlables por el agente, ni su efectiva aplicación de tales capacidades en circunstancias específicas se ejercería bajo su control. Y, por otro lado, los intentos de alcanzar la ultimidad por vía indeterminista chocarían con el requisito de la racionalidad, de la elección conforme a rasgos propios (aunque las acciones autoconfigurativas de Kane traten de dar una respuesta incompatible bastante elaborada a esta objeción).

### III

En cuanto a las posibles respuestas al Argumento: compatibilistas, incompatibilistas libertaristas, o pragmáticas (y revisionistas), dentro de la tradición filosófica que nos concierne,

todas resultan insatisfactorias por lo que respecta a la pregunta por el origen, condición de la atribución de responsabilidad de los agentes individuales y, en gran medida, por la razón ya apuntada del tipo de caracterización de éstos, que los deja inermes ante planteamientos como el de Strawson (por cuanto el requisito de autocreación divina, que contrasta con la finitud humana y la abstracción sociohistórica operada, se deja sentir al resituar al sujeto en su contexto mundano).

Compatibilistas clásicos (como Hume o Ayer), o de nuevo cuño, con su compatibilismo refinado (jerárquicos, como Frankfurt, evaluativos o de fuente múltiple, como Watson, de la razón, como Wolf), todos ellos se caracterizan por un postulado de libertad implicando verdadera responsabilidad en un mundo determinista y, excepto quizá Wolf (que, frente a los otros acentuará el componente normativo), por la cada vez más afinada búsqueda de un yo real o profundo, origen de las cadenas causales, de las razones, que concluyen en acciones de las que pueden responsabilizarse.

Frente a los que optarían por enfrentarse al Argumento Básico mediante una remisión a una instancia decisoria separada de los estados mentales, razones del agente, afectadas por los argumentos que cuestionan su control último y exigen una autodeterminación amenazada por las cadenas causales de sucesos, que parecerían arrastrar al sujeto en sus encadenamientos; la opción compatibilista, ya desde Hume<sup>13</sup>, propone la identificación del agente con sus creencias, deseos y motivos.

Libre y responsable sería aquel que hiciese lo que quisiese hacer, esto es, que actuase conforme a las razones que tuviera<sup>14</sup>, y aceptando que la relación entre razón, decisión y acción pudiese ser de necesidad causal; así, de lo que se trataría sería de distinguir entre la causalidad que necesita mis acciones, pero con la que me identificaría, y aquel control que me constreñiría, coaccionaría, bien fuera mediante una fuerza externa, bien mediante estados psicológicos de carácter patológico<sup>15</sup>.

El problema (amén de los más generales, que afectan al compatibilismo en su conjunto) es que, como se encargaría de mostrar Frankfurt<sup>16</sup>, si bien el compatibilismo clásico acertaría a distinguir, según los compatibilistas, en qué consiste la libertad de acción, descuidaría el libre albedrío, cayendo en complicaciones varias y víctima de diversos

contraejemplos claros, como la dificultad para evitar atribuir responsabilidad moral a los animales, que al menos en algunas de sus especies más desarrolladas parecen, en tanto que poseedores de deseos y creencias, capaces de desarrollar acciones libres (al modo compatibilista); o, aún más grave, la incapacidad en que se encontraría para explicar casos de experiencia de falta de libertad en virtud de deseos no compulsivos (y, en la misma línea, pero ahondando en ello, los ejemplos de condicionamiento social, que serían de control no constrictivo, y que, de nuevo, no podríamos aceptar como situaciones en las que al agente pudiera atribuírsele adecuadamente responsabilidad, aunque el compatibilista clásico no tuviese otro remedio).

La concepción estructural y jerárquica de Frankfurt sortearía esas dificultades, atendiendo a una libertad, con la atribución de responsabilidad subsiguiente, entendida como un querer lo que se quiere querer. El agente frankfurtiano poseería deseos de hacer cosas, o "deseos de primer orden", y los deseos que le movieran efectivamente a la acción constituirían su voluntad. Sin embargo, y aquí estaría el grado de complejidad que lo distinguiría de niños y animales, además poseería un segundo nivel de deseos, "de segundo orden", o actitudes reflexivas y evaluativas, cuyo objeto serían los deseos de primer orden, constituyendo su "volición de segundo orden", deseos de que ciertos deseos nos conduzcan a actuar, una subclase de éstos.

Libertad y responsabilidad se darían en las acciones en que hubiera armonía entre niveles, sin atender a la condición del Principio de Posibilidades Alternativas, como ya indiqué, quedando la experiencia de la falta de libertad no compulsiva para los casos en que tal coherencia de niveles no se diese y el sujeto actuase, por ejemplo, por un deseo de primer orden que no querría, desde su segundo nivel reflexivo, que constituyese su voluntad. Con ello también se daría cuenta de una diferencia entre agentes, con una idea de gradualidad interesante, existiendo una categoría de ellos, la del *wanton*, en la que se incluirían todos los que careciesen de esa complejidad y se dejasen guiar por la mera economía de los deseos de primer orden.

La propuesta frankfurtiana, sin embargo, ha sido objeto de importantes críticas, como la que Watson realiza<sup>17</sup>, en el seno de su propia tradición. Y es que, entre otras observaciones críticas posibles (dentro de las cuales se incluiría la cuestión más general de la originación, sobre la que he de volver al tratar del compatibilismo en general, tal como lo

enuncian estos autores, sus exponentes más destacados), se puede remedando el Argumento Básico, cuestionar que la adición de niveles del mismo tipo de estados mentales —deseos— logre otra cosa que desplazar jerárquicamente la *wantonidad*. Cabría, pues, preguntar en qué sentido obtendría el agente un control volitivo de sus estados mentales mediante la simple adición de contendientes del mismo tipo al conflicto de deseos, y si no sería necesario, a su vez, postular nuevos niveles de deseabilidad que tan sólo extenderían *ad infinitum* el problema (lo que resultaría claramente inaceptable y, más aún, abocaría a un corte arbitrario del ascenso de niveles por mor de la finitud psicológica de los agentes).

La respuesta frankfurtiana<sup>18</sup>, que pasaría por la identificación con los deseos del agente, que constituirían su yo profundo, mediante un compromiso decisivo, resulta cuando menos, pese a sus intentos de afirmar lo contrario, oscura y, probablemente arbitraria, además de no escapar de un cuestionamiento general, relativo a las fuentes de los deseos o más altos niveles de volición<sup>19</sup>. Críticas a las que tampoco escapa su intérprete, Watson, pese a su planteamiento de múltiples fuentes, que pretendería sortear la dificultad de la jerarquía frankfurtiana.

En el agente, según Watson, una fuente es la de los deseos, parte apetitiva del alma platónica (en la interpretación de la doctrina de éste, del que Watson se reivindica seguidor, aunque esto debería tomarse con cierta distancia), y otra la de los valores, o esquema evaluativo, parte racional. De nuevo la armonía sería el signo de la libertad y responsabilidad, aunque la duplicidad de orígenes y carácter de los estados mentales, con la evaluación de lo deseado, parecería mejorar el planteamiento estructural anterior. Sin embargo, y aun cuando se plantease, por ejemplo, una evaluación profunda, involucrando al yo (identificado esta vez con sus valores) hondamente, una evaluación radical ya que la elección radical, dado el determinismo, no sería posible<sup>20</sup>, ¿qué habríamos ganado en realidad respecto a Frankfurt? En ambos casos, cabría plantearse el carácter externo, heterogéneo de los elementos constituyentes del agente, y la ausencia de control por relación al origen, en los términos en los que lo plantea el Argumento Básico.

En ese sentido resulta interesante la propuesta de Wolf<sup>21</sup>, que caracteriza en estos términos críticos las propuestas compatibilistas observadas, denominándolas punto de vista

del "yo real", y propone la suya propia, asimismo compatibilista, el "punto de vista de la razón". A las argumentaciones vistas, opone pues Wolf algunas observaciones en la línea apuntada (y que, asimismo recoge el propio Watson)<sup>22</sup>, como la importancia de la historia y del entorno o, como se plantea de modo más general en los denominados casos del "Mundo feliz", la posibilidad de un control no constrictivo encubierto, por el que los deseos del agente, mediante por ejemplo un condicionamiento social refinadamente oculto, provendrían de una fuente ajena al mismo (y cabría entonces preguntarse, como sucede en otro ejemplo facilitado por Wolf, el del niño que vive una infancia extremadamente desgraciada y lleva a cabo luego actos punibles, en qué sentido podría el agente ser considerado justamente responsable por ser quién es y, en última instancia, dada la conexión causal aseverada entre rasgos de la personalidad y acciones del individuo, de sus actos).

El problema es que la perspectiva wolfiana, pese a plantear la libertad y responsabilidad como funciones de la capacidad para actuar de acuerdo con la razón, esto es, como poder de reconocer lo Verdadero y lo Bueno, recaería en problemas semejantes a los que pretende criticar y superar. Pues si bien se optaría en su perspectiva por una renuncia a la responsabilidad metafísica de la propia constitución del yo, decantándose por una responsabilidad moral, con la capacidad de autorrevisión y corrección como corolarios (y la introducción de la suerte moral, como elemento relevante de su propuesta), de nuevo el control normativo, de lo correcto, aun cuando no se hubiera reducido a una cuestión de complejidad psicológica, se vería aquejado por el problema del origen en el sentido en que, por ejemplo, los casos del "Mundo feliz" lo plantean<sup>23</sup>.

Aparentemente cerrada la vía compatibilista, la opción incompatibilista, en una de sus más sutiles formulaciones, la de Kane, se presenta como alternativa deseable al intento de responder al reto del Argumento de Strawson. La perspectiva incompatibilista rechaza la posibilidad de defender libertad y responsabilidad al tiempo que se afirma el determinismo, tal como lo hemos expuesto, apostando por la indeterminación y, en muchos casos, por una instancia decisoria, desligada de los estados mentales del agente, amenazados por ese determinismo invasor, en tanto que seres en el mundo natural, y dadas las cadenas causales en que se enlazan tales estados.

Sucintamente expuesta, por razones de espacio, cabe decir de la posición de Kane, que sorteando algunas de las objeciones clásicas que se le hacen al incompatibilismo, y consciente de los límites expuestos respecto de la anterior línea argumentativa, propone contemplar la exigencia de ultimidad en su carácter radical, así como la condición finita humana, planteando la exigencia, si es que ha de existir algún tipo de responsabilidad y libertad valiosa, de que el agente realizase algunas acciones libres indeterminadas, en el sentido de una posibilidad categórica, entre alternativas igualmente asequibles dadas unas mismas circunstancias.

Pero si las razones o motivos nos inclinan a elegir A en lugar de B, siendo A el resultado razonable de la deliberación, parecería que la elección de B, en exactamente las mismas circunstancias, si posible sería irracional (y caería dentro de las objeciones que por lo que respecta al incompatibilismo libertarista, rápidamente, hemos señalado). Para evitarlo y defender esta posibilidad, que algunas veces debería tener el agente en su mano, esto es, la de poder ir configurándose mediante ese tipo de decisiones y acciones, postula Kane la racionalidad plural, la posibilidad de que existan razones diversas, por ejemplo morales y no morales, que no obliguen a escoger, dada la historia psicológica del sujeto, una opción antes que otra. Es decir, que tales razones no puedan ordenarse según una escala bayesiana de preferencia, de racionalidad; que sean inconmensurables, dadas las mismas circunstancias.

Teniendo en cuenta esta salvedad, lo que Kane propone es que en determinados momentos el agente elige, en virtud de una intención general, no *a priori*, no determinada por rasgos y circunstancias anteriores (lo que de nuevo recuerda la caracterización abelardiana del *consensus*), de un modo u otro, con libertad, reforzando con ello sus rasgos morales u otros. El control, la responsabilidad, última es posible, está al alcance del sujeto y no mediante actos indeterminados irracionales, o postulaciones metafísicas problemáticas, eso al menos pretende Kane, sino a través de decisiones sobre las que el agente tiene un control no antecedentemente determinante (lo que reintroduciría las amenazantes cadenas causales), pero sí en el mismo momento de darse.

Sin embargo, y dejando de lado la compleja discusión relativa a las múltiples críticas que el libertarismo de Kane ha



recibido<sup>24</sup>, baste hacer hincapié en las dudas que suscita una posición que, en última instancia se entiende como elección sin bases, y trata así de evitar el regreso al infinito del Argumento Básico, que acepta en su formulación como reto. Cabe preguntarse si resulta plausible una autodeterminación hecha con independencia de las propias creencias y razones, que se reduce a constatar en el momento de la elección, que ésa, cualquiera que fuese la decisión tomada, era la opción preferida, porque, de hecho, tautológicamente fue la preferida y realizada.

Así pues, los ya tradicionales argumentos compatibilistas e incompatibilistas, se encuentran en dificultades al tratar de responder al Argumento Básico. Dificultades compartidas, pese a la diferencia de posiciones en el espectro del campo de la acción, que permiten especular sobre la común condición que les aboca al fracaso ante retos como el que plantea Galen Strawson.

Y es que, ciertamente, mantienen todas estas corrientes supuestos comunes que, ya hemos parcialmente adelantado y denunciado, en tanto que no asumidos como tales serían pre-supuestos o prejuicios. La concepción puramente estructural del agente, desentendiéndose de contenidos (en todo caso, de nuevo en general, moralmente neutros), la atención y el intento de dar cuenta de libertad en términos de procedimientos, antes que de actitudes, el formalismo pues, unido a, en general, una desatención radical a la dimensión histórica y social del agente (en todo equiparable así al pecador abelardiano), caracterizan a estas diversas corrientes.

Existen, sin embargo, otras alternativas, que rompen con este paradigma fallido del libre albedrío y de la responsabilidad, como la que plantea P. F. Strawson<sup>25</sup>. Éste, frente a las opciones hiperintelectualizadas de compatibilistas e incompatibilistas, plantea la responsabilidad en términos sociales, de actitudes y trato que la gente se da, unos a otros. Considerar a la gente responsable es tratarla de determinados modos, realizar demandas generales, y reaccionar según las respuestas del otro. Las actitudes reactivas, tales como la ira o el resentimiento, entre muchas otras, serían constitutivas de la responsabilidad moral, responsabilidad que se iría adquiriendo en la sociedad a medida que el agente se adiestrara en la práctica de tales actitudes.

No habría una justificación teórica ulterior, en la propuesta de P. F. Strawson a la pregunta metafísica o la validación de un cierto contenido de creencias en torno a la responsabilidad. De este modo, la dimensión ontológica en que se mueve el reto del Argumento Básico, no encontraría espacio para plantearse en el seno de las prácticas reactivas, en tanto que éstas descansarían en aquello que constituye el marco de los juegos de lenguaje básico en nuestra concepción del ser humano. La excusa o la exención de actitudes reactivas<sup>26</sup>, en el límite el hecho de que a determinados seres no les otorguemos responsabilidad alguna, reservándoles un trato objetivo, descansaría no en un descubrimiento metafísico, sino en un simple test de la capacidad para entrar en la dinámica de las actitudes reactivas.

Sin embargo, pese al atractivo de esta propuesta, pretendidamente conciliatoria, la consideración meramente pragmática de la responsabilidad, de las actitudes reactivas que, según Strawson no se modificarían en virtud de descubrimientos metafísicos como la verdad del determinismo, resulta poco plausible<sup>27</sup>. Se podría, pues, poner en cuestión la supuesta inescapabilidad de esas reacciones humanas, o que no pudiesen elevarse interrogantes relativos al marco de las prácticas, ya que éste presenta diversos puntos desde los que cabría solicitar una revisión de las actitudes reactivas, por ejemplo focalizando el carácter cognitivo de la responsabilidad, la posibilidad de que nuestras actitudes dependieran de juicios falsos. Parecería que un más completo informe de las condiciones de la responsabilidad podría suscitar problemáticas metafísicas que podrían socavar la integridad normativa de las prácticas. De este modo, la perspectiva ontológica, rasgo también compartido por las reseñadas corrientes compatibilista e incompatibilista, así como por el reto del Argumento Básico, continuaría siendo una posición mejor asentada (teórica e intuitivamente).

Es por ello que parece necesario, sin olvidar lo aprendido con P. F. Strawson, ni las críticas realizadas a los intentos de solución tradicionales, atender al Argumento Básico desde una perspectiva metafísica, ontológica, enfrentándose desde esa posición a la cuestión de la condición de Origen, o Responsabilidad Última. En lugar de no tomar completamente en serio la exigencia de ultimidad, o creer que ésta está al alcance del agente con relativa celeridad, en lugar de despreocuparse por la condición ontológica del



agente, nuestra propuesta, retomando las sugerencias de diversos autores<sup>28</sup>, sería la de cuestionar su estatuto individualista, abstraído de la urdidumbre de interlocución del espacio social en el que surge, y aunque de manera ya secundaria en nuestra atención (por imperativos de espacio), ahistóricamente concebido, desustancializado, sin contenidos, reducido a mera estructura, formalizado y procedimentalizado, ajeno a la dimensión actitudinal de la responsabilidad (y a la cuestión de las demandas y expectativas normativas sociales a ella ligadas, como hemos visto en P. F. Strawson).

#### IV

Diversos factores, aunque sólo haya reseñado uno muy claro al inicio de mi texto, han coadyuvado a la concepción del yo originador como portador de la responsabilidad moral, en el sentido en que el Argumento Básico plantea la cuestión, pero cabría preguntarse en qué sentido lo que constituyen condiciones y puntos de partida en la constitución del sujeto. Los rasgos heredados del agente y el entorno social, entre otros, deberían convertirse directamente en condicionantes *tout court* y elementos por los que cabe responsabilizarse de un modo absoluto, según el modelo de la autocreación (máxime cuando, sin tales condiciones el agente, directamente, no existiría).

La educación, que cualifica al agente para participar en el sistema social, en los juegos de lenguaje del elogio y la culpa, que está destinada a hacer del niño origen de sus propias acciones, puede realmente cualificarse gradualmente, como pueden hacerlo responsabilidad y libertad, en función del tipo de sociedad, de factores variables y flexibles, de tipo sociohistórico, y no parece claro el paso de esa gradualidad y amplitud a una exigencia rasa y extrahumana de autocreación *ex nihilo*, según una noción de individuo y agencia que no se compece con la de nuestras sociedades e historia, con nuestra condición finita.

Vivimos en horizontes sociales que incluyen discriminaciones valorativas, y nuestra identidad se define por los compromisos e identificaciones que asumimos en ese marco. Identidad pues en el seno de la interlocución, siempre como respuesta a la pregunta ¿quién soy yo? (y el uso del lenguaje sólo existe y se mantiene en el seno de una comunidad

lingüística), dentro de la conversación social, de las prácticas y demandas normativas sociales. Por ellas, evaluándonos mediante juicios valorativos, regulamos nuestras conductas y en esta dimensión normativa y social de la responsabilidad moral es donde caben elementos de la agencia no contemplados en las perspectivas tradicionales, como el sentido del propio valor o la capacidad como agente.

Y así, tal como expone Benson, podría suceder que un agente cumpliera con todos los requisitos y capacidades procedimentales que se suponen suficientes como para garantizar su libertad y responsabilidad (dejando por ahora entre paréntesis el reto strawsoniano o el del determinismo), esto es, la conexión causal apropiada, sin obstáculos, entre voliciones y acciones y que, sin embargo, como sucede en diversos casos que revelan la dimensión social de la libre agencia, no confiase en su competencia como agente y, por lo tanto, no pudiese disfrutar de esa libertad aparentemente inobjetable.

Sentimientos como el de la vergüenza<sup>29</sup>, o situaciones en las que se manifiesta la influencia de otros agentes, por ejemplo en los denominados casos de "luz de gas", o de instituciones históricas, como la de la esclavitud; todos ellos vinculados a la dimensión social del agente, por cuanto degradan la confianza en su capacidad para responder por la propia conducta, a la luz de las demandas y expectativas normativas fruto de la interacción social, sin que necesariamente las capacidades procedimentales se vean afectadas, revela la artificiosa dicotomía que entre interioridad/ exterioridad y autonomía/ heteronomía, plantean argumentos como el de Galen Strawson (y respetan, en tanto que asumen la misma figura de agencia, los argumentos contemplados).

En la medida en que, a la luz de esta argumentación, se revela la mayor complejidad de las relaciones entre individuo y sociedad, la posibilidad de que factores externos al cuerpo individual tanto limiten, cuanto contribuyan al desarrollo de la propia autodeterminación, y a la inversa, mostrándose además una variabilidad historicosocial que, de nuevo, los argumentos tradicionales y especialmente el de Strawson, no contemplan, cabría retar a la exigencia de autodeterminación o autocreación en los términos en que lo formula el Argumento Básico. Así pues, el agente sí que podría ser responsable moralmente de sus

actos, aunque no fuese *causa sui*, porque los condicionantes sociales, históricos, los estados mentales que escapan a su control personal, son las condiciones de la constitución del individuo y posibilidades abiertas a la decisión y acciones del sujeto sociohistórico, pero no de modo absoluto, como propiedades de un agente-mónada sino como cualidades historicosociales, que posibilitan un mayor o menor grado de desarrollo moral.

Quizá no seamos dioses, pero tampoco, necesariamente, simples eslabones de cadenas causales inacabables. Para que sea posible dar cuenta de un concepto humano de responsabilidad sería necesario abandonar el modelo teológico-metafísico de antaño, en un sentido más radical que aquel en que lo planteaba Dennett, pues, como reza el título de un libro de Rafael Sánchez Ferlosio, mientras no cambien los dioses, nada ha cambiado.

### Notas

\* Este trabajo ha sido llevado a cabo dentro del proyecto de investigación "Creencia, motivación y verdad" (BFF 2003-08335-C03-01) y gracias al disfrute de una beca FPU del Ministerio de Educación y Ciencia. Agradezco algunas conversaciones mantenidas con Juan José Colomina, Bernat Martí, David Mateu y Pascual Raga, que de algún modo han contribuido a la elaboración de este texto y, especialmente, el apoyo de mi pareja, Mara, quien me ha brindado la calidez y comprensión necesarias para completar este trabajo.

1 "Ni en el mundo, ni, en general, tampoco fuera del mundo, es posible pensar nada que pueda considerarse como bueno sin restricción, a no ser tan sólo una *buena voluntad*" (Cf. Kant, 1992, 21).

2 Cf. Chisholm, 1982, 28-29.

3 Cf. Honderich, 2002, 461-476.

4 Véase Double, 1996, 847-854, que critica a Honderich en tanto que al entender ambos en sentido actitudinal, como asimismo el propio Double, no podría apelar a su verdad o falsedad.

5 Puede seguirse en Strawson, 1986, 28-29; Strawson, 1994, 5-24.

6 Luego refinaremos ésto, con las distinciones relativas al control constrictivo, no constrictivo y los casos del "Mundo Feliz" (control no constrictivo encubierto).

7 Cf. Ayer, 1982, 16-17, ya que si no de facto sería teóricamente alcanzable.

8 Si bien la cuestión, como veremos es más compleja y, tanto esta objeción,

de entrada más simple, como la ya más severa de la ininteligibilidad o de la postulación de misteriosas entidades metafísicas, en pro de una responsabilidad y libertad incompatibilistas, será susceptible de una reformulación más interesante en autores como Kane.

9 Cf. Davidson, 1980, 3-19.

10 Cf. Russell, 2002, 242-244.

11 También en breve volveré sobre este punto, con una sucinta, pero más detallada, exposición de las diversas posiciones que podrían encuadrarse en el seno de una actitud compatibilista.

12 Cf. Fischer, Ravizza, 1998, 170-206.

13 Y la ya mentada versión remozada ayeriana, que aquí sigo (Cf. Ayer, *op. cit.*; Moya, 1997, 117-119).

14 Lo que supone un primado de la autodeterminación, en detrimento de la otra condición indicada al inicio de mi artículo, el Principio de Posibilidades Alternativas, rasgo éste que aqueja a toda esta tradición (y no simplemente como una opción expositivo-metodológica, lo que daría pie al establecimiento de una crítica, como de hecho se ha llevado a cabo en reflexiones en algún punto cercanas a las expuestas en este artículo, al criticar la noción de un agente concebido en términos casi divinos, *causa sui* teológica, que es en el extremo lo que requieren todas las posiciones que, acercándose al problema del libre albedrío y la responsabilidad moral desde la posición tradicional, optan por la vía de la autodeterminación, quieran ser o no radicales).

15 Y aquí se introducen los bien conocidos casos de comportamiento obsesivo, la

**Recibido:** 27 de octubre de 2005

**Aceptado:** 14 de diciembre de 2005

cleptomanía, entre otros, cuya especial modulación causal se distinguiría del control del agente, de su causalidad propia.

- 16 Cf. Frankfurt, 1982, 89-90.  
 17 Cf. Watson, 1982, 107-110.  
 18 Cf. Frankfurt, 1988, 159-176.  
 19 Como se deduce además de su reiterado rechazo y, en el artículo antes citado, explícito, de la dimensión histórica, de la importancia del origen causal de deseos y creencias.  
 20 Cf. Taylor, 1982, 111-126.  
 21 Véanse Wolf, 1990, 67-93; Wolf, 1987, 46-62.  
 22 Cf. Watson, 1987a, 151-153.  
 23 De hecho, de nuevo Wolf rechaza, en el caso de la acción buena, la condición del Principio de Posibilidades Alternativas y aprueba un dominio incluso irresistible de lo Bueno, la primacía de la armonía, por delante de cualquier otra consideración, con lo que quedaría preso de las posibles manipulaciones (no resulta imposible imaginar un Mundo Feliz donde los

agentes hubiesen sido manipulados para actuar, desear y evaluar reflexivamente conforme a lo Verdadero y lo Bueno) o cadenas causales de estados mentales que escaparían al control agente.

- 24 Véanse Strawson, 1986, 25ss; Wolf, 1990, 46ss, para críticas al libertarismo en general.  
 25 Véase su Strawson, 1995, 37-67.  
 26 Véase para tal distinción Watson, 1987b, 259-261.  
 27 Véanse críticas en Watson, *op. cit.*; Vargas, 2004, 220-222.  
 28 Entre ellos Dent, 1997, 171-183; Taylor, 1996, 41-68, y después Benson, 1994, 650-668; Russell, 2002, 237-239; Richardson, 2001, 287-303.  
 29 Véanse James, 1952, 189-191; Taylor, 1985, 53-84; Deigh, 1983, 225-245; Williams, 1993, 50-102; Goffman, 1970. Esta interesante discusión se centraría en un sentimiento moral, de autoprotección, relacionado con la idea de degradación de la propia imagen del yo, y con su estructura de

interacción yo-mundo, que reafirmaría así el entrelazamiento individuo-sociedad, la dimensión social de las actitudes morales, de la responsabilidad y la libertad, más allá del reductivo y solipsista análisis tradicional (ya que como indiqué siguiendo a Benson, la imagen que uno tiene de sí mismo, resultado de la interacción con los otros miembros de la propia comunidad, esto es, producto social y no del individuo, afectaría a la respuesta y competencias como agente que puede desarrollar el sujeto que el grupo evalúa. La identidad y roles así socialmente constituidos, de modo parcial, mediante la vergüenza, en tanto que interiorización de un juicio social, metafóricamente plasmado mediante la imagen del tribunal o del observador externo, la sensación de ser observado, afectarían a las capacidades y posibilidades de actuación del agente, sin que los modelos tradicionales en filosofía de la acción pudiesen dar cuenta de ello).

## BIBLIOGRAFÍA

- Ayer, A. J. (1982): "Freedom and Necessity", en G. Watson (ed.), *Free Will*. New York, Oxford U. P.
- Benson, P. (1994): "Free agency and Self-Worth", *The Journal of Philosophy*, 91, 650-668.
- Chisholm, R. M.: "Human Freedom and the Self", en G. Watson (ed.), *op. cit.*
- Davidson, D. (1980): *Essays on Actions and Events*, Oxford, Clarendon Press.
- Deigh, J. (1983): "Shame and Self-Esteem", *Ethics*, 93, 225-245.
- Dent, N. J. H. (1997): "La responsabilidad de ser lo que se es", en J. Corbí; C. J. Moya (eds.), *Ensayos sobre libertad y necesidad*, Valencia, Pre-textos.
- Double, R. (1996): "Honderich on the Consequences of Determinism", *Philosophy and Phenomenological Research*, 61, 847-854.
- Fischer, J. M.; Ravizza, M. (1998): *Responsibility and Control*, Cambridge, Cambridge U. P.
- Frankfurt, H. G.: "Freedom of the Will and the Concept of a Person", en G. Watson (ed.), *op. cit.*
- Frankfurt, H. G. (1988): "Identification and Wholeheartedness", en H. G. Frankfurt, *The Importance of what we care about*, Cambridge, Cambridge U. P.
- Goffman, E. (1970): *Estigma*, Buenos Aires, Amorrortu.
- Honderich, T.: "Determinism as True, both Compatibilism and Incompatibilism as False, and the Real Problem", en R. Kane (ed.), *op. cit.*
- James, W. (1952): *The Principles of Psychology*, Chicago; Londres, et. alii.: *Encyclopaedia Britannica*.
- Kant, I. (1992): *Fundamentación de la metafísica de las costumbres*, Madrid, Real Sociedad Económica Matritense de Amigos del País.
- Moya, C. J. : "Libertad y responsabilidad moral", en J. Corbí; C. J. Moya, (ed.), *op. cit.*
- Richardson, H. S. (2001): "Autonomy's many normative Presuppositions", *American Philosophical Quarterly*, 38, 287-303.
- Russell, P.: "Pessimists, Pollyannas and the New Compatibilism", en R. Kane (ed.), *op. cit.*
- Strawson, G. (1986): *Freedom and Belief*, Oxford, Clarendon Press.
- Strawson, G. (1994): "The Impossibility of Moral Responsibility", *Philosophical Studies*, 75, 5-24.
- Strawson, P. F. (1995): *Libertad y resentimiento*, Barcelona, Paidós.
- Taylor, Ch.: "Responsibility for Self", en G. Watson (ed.), *op. cit.*
- Taylor, Ch. (1985): *Philosophy and the Human Sciences: Philosophical Papers*, Vol. II, Cambridge, Cambridge U. P.
- Taylor, Ch. (1996): *Fuentes del yo*, Barcelona, Paidós.
- Vargas, M. (2004): "Responsibility and the Aims of Theory: Strawson and Revisionism", *Pacific Philosophical Quarterly*, 85, 218-241.
- Watson, G.: "Free Agency" en G. Watson (ed.), *op. cit.*

Watson, G. (1987a): "Free Action and Free Will", *Mind*, 96, 145-172.

Watson, G. (1987b): "Responsibility and the Limits of Evil", en F. Schoeman (ed.): *Responsibility, Character and*

*the Emotions*, New York, Cambridge U. P.

Williams, B. (1993): *Shame and Necessity*, Berkeley; Los Angeles, California U. P.

Wolf, S. "Sanity and the Metaphysics of Responsibility", en F. Schoeman (ed.), *op. cit.*

Wolf, S. (1990): *Freedom within Reason*, Oxford, Oxford U. P.