

**ALGUNAS OBSERVACIONES SOBRE LA
ATRIBUCIÓN DE LA RESPONSABILIDAD
(MORAL) DE NUESTRAS ACCIONES**

*SOME REMARKS ON THE ATTRIBUTION OF (MORAL)
RESPONSIBILITY TO OUR ACTIONS*

Juan José Colomina Albiñana

Vicente Raga Rosaleny

Universidad de Valencia
Valencia-España

Sentimos que nuestras acciones están sujetas a nuestra voluntad en la mayoría de las ocasiones, e imaginamos sentir que la voluntad misma está sujeta a nada; porque cuando al negarla, somos llevados a intentarlo, sentimos que se mueve fácilmente por todas partes, y produce una imagen de sí misma incluso en ese lado donde no se estableció (Hume, 1988, p. 408).

RESUMEN

Nuestra intuición nos dice que a pesar de pertenecer a un mundo determinado físicamente somos libres. Nosotros argumentaremos que tras este postulado se encuentra una concepción errónea de sujeto, que ilusoriamente nos hace creer que nosotros elegimos hacer lo que acabamos haciendo. Contra este argumento, Galen Strawson plantea una objeción basada precisamente en la imposibilidad del agente de configurarse a sí mismo, pero es una

objección que recae en la misma equivocada noción de sujeto. A partir de aquí analizaremos los argumentos dennettianos a favor del compatibilismo, lo que nos permitirá concluir apuntando a la correcta dirección que debería recorrer todo intento de estipulación de una noción de individuo plausible.

Palabras Clave: Responsabilidad Moral, Argumento Básico, Sujeto, Principio de Posibilidades Alternativas (PPA), Compatibilismo.

ABSTRACT

Our intuition says us that in spite of belonging to a determinate world we are physically free. We will argue that after this postulate is an erroneous conception of subject that very falsely makes us think that we chose to do what we ended up doing. Against this argument, Galen Strawson raises an objection based indeed on the impossibility of the agent to form itself, but it is an objection that falls to the same one mistaken subject notion. From this point, we will analyze the dennettian arguments in favour of the compatibilism which will allow us to conclude pointing at the correct direction that would have to cross all attempt of stipulation of a reasonable notion of individual.

Key Words: Moral Responsibility, Basic Argument, Subject, Principle of Alternative Possibilities (PAP), Compatibilism.

INTRODUCCIÓN

Uno de los argumentos escépticos más fuertes, por lo que respecta a la cuestión de la atribución de responsabilidad moral, que se han formulado en el campo de la filosofía de la acción, sería el Argumento Básico de Galen Strawson. La exigencia strawsoniana de una responsabilidad por ser lo que se es, por los rasgos constitutivos del agente, para poder dar cuenta de la responsabilidad moral por las propias acciones, convierte tal atribución en algo, aparentemente, injustificable. Nuestra intención en este ensayo será rebatir el argumento strawsoniano, o al menos ponerlo en entredicho, y, para ello, daremos primero unos pasos preliminares, que nos permitirán, como conclusión, afirmar la posibilidad de tal responsabilidad.

En un primer momento, pues, mediante una breve ojeada a la historia de la filosofía, trataremos de establecer cuál puede ser el origen de la noción de individuo que está en la base del Argumento Básico, y que es la que dota de toda su fuerza a la argumentación escéptica strawsoniana. En segunda instancia estableceremos ese mismo Argumento, que se pretende válido independientemente de compromisos con distinciones entre determinismo e indeterminismo para, en tercer lugar, exponer sucintamente los intentos de respuesta que desde posiciones compatibilistas e incompatibilistas se han venido, o podrían venirse formulando, haciendo especial hincapié en la propuesta de Daniel C. Dennett, que nos permitirá plantear una concepción de sujeto que dé cuenta de todas las limitaciones anteriores y, en tanto, que histórica y socialmente configurada, polémica con el esquema de la propuesta strawsoniana.

I

La imagen que del agente se desprende de la tradición dominante en el campo de la filosofía de la acción actual, sobre el que se vierten las cuestiones relativas a la libertad (entendida como voluntad libre) y a la responsabilidad moral (que tiene a la libertad como su presupuesto), no es la única posible.

La noción de individuo, de su libertad y responsabilidad moral asociadas, que subyace a esta tradición, podría tener sus orígenes más remotos en el pensamiento de un autor escasamente citado y conocido por ella, el pensador medieval Pedro Abelardo. Éste tendría una tesis fundamental para el temprano “surgimiento del individuo” que caracterizó al siglo XII: 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. En ese espacio vacío entre el *carácter* y la *acción* estaría la decisión moral y, por tanto, el individuo.

Fáciles de establecer son los vínculos genealógicos 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.

Así pues, 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, entendiendo suceso 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?

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 de la Responsabilidad Última).

Sin embargo, más allá del rótulo “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,¹ el problema tal como lo plantean los sucesores del determinismo fuerte clásico 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.

Así pues, según se explicita en uno de los argumentos más relevantes en este sentido, el Argumento Básico de Galen Strawson (1986 y 1994), 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 lo serían en virtud de una razón del agente que las causa; esto es, en función de cómo es uno mentalmente. Pero no querríamos, 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

¹ Honderich, T. (2002). Determinism as True, both Compatibilism and Incompatibilism as False, and the Real Problem. En R. Kane (ed.), *Oxford Handbook of Free Will* (pp. 461-476). Oxford: Oxford U. P.

como *chance*, y del tipo de la que podríamos denominar suerte constitutiva), ni en virtud de algún tipo de control externo al agente (Williams, 1993).

Por supuesto que cabe, como arguye Dennett, 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. Pero esto 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 semi-compatibilistas Fischer & Ravizza (1998)) 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.

II

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.

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.

El problema, sin embargo, es que, como se encargaría de mostrar Frankfurt, H. G. (1982), 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.

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.

La propuesta frankfurtiana, sin embargo, ha sido objeto de importantes críticas, como la que Watson (1982) realiza, en el seno de su propia tradición. Y es que, entre otras observaciones críticas posibles, 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 *wantonicidad*.

En ese sentido resulta interesante la propuesta de Wolf (1990), 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 Wolf algunas observaciones en la línea apuntada (y que, asimismo recoge el propio Watson (1987)), como la importancia de la historia y del entorno.

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, 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.

Aparentemente cerrada la vía compatibilista, la opción incompatibilista, en una de sus más sutiles formulaciones, la de Kane (1996), 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. 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.

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.

Sin embargo, y dejando de lado la compleja discusión relativa a las múltiples críticas que el libertarismo de Kane ha recibido, 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 presupuestos o prejuicios. La concepción puramente estructural del agente, desentendiéndose de contenidos, 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.

III

Llegados a este punto, queremos detenernos en los intentos de solución que el filósofo estadounidense Daniel C. Dennett ofrece de las dificultades apuntadas en la primera parte de este escrito. Como hemos tenido ocasión de mostrar, la respuesta adecuada a los problemas de la autoconfiguración personal y del Principio de Posibilidades Alternativas, los dos pilares sobre los que (parece) se sustenta la posibilidad de atribución de la responsabilidad (moral) de nuestras acciones, pasaría por el rechazo o, al menos, la reformulación de la anterior concepción del sujeto como un agente-mónada marcado teo-teleológicamente y la formulación de una nueva noción de yo en la que se tenga en cuenta tanto la inclusión del individuo dentro del mundo físico como el respeto de su libre voluntad.

Dentro de tal intento, que podemos denominar compatibilista, podemos encuadrar la teoría dennettiana del sujeto. Dicha tentativa consta de dos momentos: por una parte, una vertiente

filogenética en la que Dennett intenta incluir al ser humano dentro de la cadena evolutiva; por otra, una vertiente ontogenética en la que ofrece una nueva visión del sujeto como capaz de auto-atribuirse a sí mismo razones para actuar.

Así pues, desde el primer punto de vista, Dennett propone un marco en el que colocar las diversas opciones de diseño: es la llamada torre de la generación y la prueba.² La evolución organiza un conjunto de células en diversos modelos según lo adecuado de sus respuestas: una buena respuesta (lo que no significa que sea la única, o la mejor, o la más correcta) es aquella que funciona, y una conducta que funciona se verá potenciada y también se verá reforzada la estructura capaz de ocasionarla, maximizando por ello sus posibilidades de desarrollo. Por ello, en cada estrato de la torre podemos colocar multitud de clases de organismos que pueden desenvolverse en su entorno, lo que permitirá la expansión de su población (crecimiento horizontal); pero sólo aquellos organismos cuyos movimientos son más eficaces verán potenciada su estructura de tal modo que podrán sobrevivir y evolucionar (crecimiento vertical).

La defensa de tal punto de vista obvia cualquier tipo de teleología, porque el requisito para que una determinada distribución fenotípica se potencie no es la consecución de ningún fin, sino que está sujeto al ciego proceso de selección por funcionalidad. Llegados a un cierto punto de la cadena evolutiva, aparecen cierto tipo de criaturas capaces de fabricar herramientas conceptuales que les permitirían ir más allá de la mera determinación natural, pero además el uso de dicha herramienta otorga la misma inteligencia a aquellos que la emplean sin haberla creado. Cuanta más información requiera el diseño de una herramienta, mayor inteligencia potencial confiere al usuario. Las herramientas más notables son aquellas denominadas mentales: las palabras. Los organismos humanos son capaces de beneficiarse de la experiencia encarnada en las herramientas que la comunidad social en conjunto ha ido elaborando y

² Dennett, D. C., 1996, pp. 101 y ss.

transmitiendo a lo largo de los años, lo que les permite reflexionar mejor acerca de lo que deben hacer o pensar a continuación (les permite aprender "... a pensar mejor sobre lo que deben pensar a propósito del siguiente movimiento" (p. 624).³

El organismo humano, así, puede modificar su propia estructura fenotípica de modo que le permita crear, identificar y generalizar relaciones no existentes en el mundo externo. La diferencia entre las criaturas humanas y las criaturas no-humanas consistiría, pues, en el manejo más eficaz que tienen las primeras de pensamientos concretos referentes a objetos del mundo que se constituyen de forma abstracta como representaciones.

Los organismos humanos, así, son privilegiados porque la evolución les ha dotado del poder de construir hipótesis y representaciones generales acerca de la identidad de los objetos e individuos externos; podemos rastrear objetos según su relevancia en el entorno cambiante y percatarnos de nuestras acciones porque todo ello lo asociamos a nuestro esquema intencional.

Pero, si desde el punto de vista filogenético es la aparición del yo, como soporte de agencialidad y marcada por las herramientas cognitivas y lingüísticas que él mismo se ha construido, la que permite al individuo evolucionado dotarse de sus propias razones de acción (lo que le permite auparse hasta la libertad), desde el punto de vista ontogenético lo que nos interesa es ver el modo en que el lenguaje permite al yo individual ser considerado como un agente al que se le puede atribuir responsabilidad moral.

Siguiendo la argumentación dennettiana, lo primero que viene a mostrarnos éste es que "un yo es, sobre todo, un locus de autocontrol" (p. 81).⁴ Pero, por otra parte, también considera Dennett que la responsabilidad implica una gradación, ya que depende de la capacidad evaluativa del individuo el que nuestras actitudes protocolares sean revisadas, criticadas y modificadas según el modo en que cada uno quiere ser, una revisión que,

³ Dennett, D. C., 1995, p. 624.

⁴ Dennett, D. C., 1984, p. 81.

obviamente, presupone una atención especial al entorno socio-político en el que está inmiscuido el agente.

Alguien podría criticar la postura dennettiana afirmando que tan sólo es cuestión de suerte el que un individuo esté más capacitado que otros para autorreflexionar sobre sus propias evaluaciones, por lo que depende de la fortuna de nacer con mayor capacidad de racionalidad el llegar a ser considerado como una agencialidad responsable. Pero Dennett no es partícipe de esta idea, antes bien para este autor todos los seres humanos estarían igualmente capacitados para controlar sus propias acciones, por lo que aunque alguien esté mayor dotado que otros (en cierto sentido), dado el largo proceso que se requiere para la construcción del propio yo, todos ellos tienen (en principio) las mismas posibilidades de llegar a adquirir la responsabilidad (moral) sobre sus acciones.

Del desarrollo de esta concepción de la acción moral, Dennett extrae su concepto del yo y es que en un momento determinado, la teoría dennettiana se ve en la necesidad de dotar de cierta forma a toda la enorme cantidad de deliberaciones y decisiones que sobre sus propias acciones realiza un determinado organismo. Lejos de suponer un Teatro Cartesiano en el que se representarían dichas deliberaciones y sin admitir un único punto en el que se coordinarían dichas metarreflexiones, Dennett postula la existencia de lo que denomina 'centro de gravedad narrativo' donde se daría cita todo aquello que no puede ser atribuido en tanto agente de acción.

Sin embargo, una primera crítica surgiría a raíz de esta idea y es que si el yo se considera como una creación, como una ficción, parecería adolecer de cierta debilidad metafísica y tendría serios problemas para soportar el peso de la responsabilidad moral. Pero la respuesta de Dennett es que precisamente dicha debilidad es su virtud, ya que aunque el yo sea una ficción, cualquiera estaría orgulloso de haber sido su creador, ya que éste englobaría todo aquello que se ha hecho, se ha creado, se ha inventado e, incluso, se ha fingido, por lo que en cualquier momento podríamos crear

uno nuevo, sustituyendo así al anterior y adecuándonos así a las nuevas circunstancias.

Pese a todo, la debilidad ontológica seguiría presente al afirmar el mismo Dennett que, en el proceso de constitución del yo, el sujeto no dispone de ningún tipo de entidad directriz, que no existe un Jefe de la Mente que a modo de capataz organice las diferentes narraciones que van surgiendo sino que, muy al contrario, se da una explosión (inconsciente) de identidades y personalidades (entendidas como narraciones pasadas) que se van desarrollando, configurando y abandonando conforme una de ellas cobra cierto protagonismo y se va imponiendo sobre las demás en tanto es elegida (o simplemente aparece como más atractiva).

A primera vista, Dennett parece que ha sido capaz de incluir dentro de la cadena determinista un cierto tipo de libertad, porque parece que su noción de yo capacita para, en última instancia, elegir nuestra propia personalidad en el sentido en que parece permitir que seamos nosotros los que nos demos nuestras propias razones para ser como somos, en tanto que son nuestras autoevaluaciones las que permiten incorporar cierto tipo de metarreflexiones a nuestra configuración personal. Sin embargo, desde el punto de vista del determinismo radical cabría plantear una duda y es que si hemos considerado que el lenguaje es producto de una serie causal determinada por la ley de la evolución, entonces parece que el propio lenguaje (que es el requisito básico para llegar a dotarnos de libertad, según la teoría dennettiana) es un mecanismo automático creado por las leyes físicas (en tanto que no puede ser improvisado por ser el vínculo histórico entre una determinada sociedad y su propio pasado) para construirnos a nosotros y no un instrumento que nosotros construimos para que nos permita autoconfigurarnos como agentes responsables.

Una respuesta de tipo denettiano podría afirmar que el lenguaje, en tanto que tradición recibida y heredada pero sometida a revisión y crítica, posibilita una nueva configuración de conceptos y categorías, por lo que en realidad permite tomar distancias de

las originales razones y creencias que por medio de él pudieran transmitirse, admitiendo así una nueva cosmovisión que diera lugar a la nueva generación de narraciones no-determinadas. Pero también existe una crítica más débil, de circularidad, que afirma que si tenemos por cierta la tesis de la constitución narrativa del yo, entonces deberíamos presuponer una cierta entidad previa que narre todo el proceso de construcción y que, por tanto, estaría fuera de nuestro control o bien tendríamos que suponer que el yo emerge de una encarnizada lucha a muerte entre yoes que van surgiendo según acontece la historia cognitiva del sujeto, sin necesidad de una entidad previa, pero dejando al lenguaje como un elemento que se autogestiona. La solución dennettiana, sin embargo, es clara: lo que ocurre en el momento previo a la configuración del yo es lo segundo, pero el lenguaje no es un elemento autogestionado, sino que es un elemento democrático que cualquier narración o sujeto tiene a mano para poder autoconstituirse como tal.

Dada esta explicación, más o menos satisfactoria, del carácter ontológico del yo, todavía queda pendiente la acusación de debilidad. ¿Puede el yo, tal y como lo ha tejido Dennett, soportar el peso de la responsabilidad moral? Si, por una parte, la ficcionalidad del yo parece debilitar la base sobre la que podemos referir la atribución de responsabilidad moral y si, por otra, la misma ficcionalidad permite una mayor y más radical autoevaluación que un yo sustancial y limitado, entonces parece que la problemática de la ontología del yo y de la posibilidad de su autoconfiguración dentro de un mundo determinista se mueve entre el establecimiento de un yo menos ficcionalizado, pero capaz de soportar el peso de la responsabilidad moral o la concesión de un yo ficcionalizado que a cambio de un mayor autocontrol ceda determinado grado de responsabilidad.

La posición compatibilista de Dennett es capaz de superar la segunda opción, pero debe enfrentarse con la primera, ya que un centro de gravedad narrativo no garantiza el suficiente nivel de autogestión del yo necesario para soportar la carga de la res-

ponsabilidad moral. Parece que el complemento que requiere la configuración de un agente capaz de originar sus acciones y decisiones a partir de sus propios deseos y creencias es que dicho principio de autonomía descansa sobre un lecho de posibles alternativas de acción todavía no realizadas, pero que podrían serlo en el momento de la acción (o elección) del agente, el denominado *Principio de Posibilidades Alternativas* (PPA). Para que una acción pueda ser catalogada como libre y voluntaria el agente, parece, debe ser capaz de poder explicarla a partir de razones que lo movieron a actuar del modo en que lo hizo, razones que actuarían como guías de la acción humana, incluidas dentro del mundo, con poder causal.

Si atendemos a su desarrollo ontogenético, como aquí hemos hecho, la cuestión principal acerca de la responsabilidad moral de un agente es optimista. Porque a pesar de que los deseos y creencias que nos mueven a actuar, según Dennett, son producto de una acumulación e interiorización de normas y convenciones sociales, esto no nos debe hacer desesperar en nuestro intento de actuar libremente dentro de un mundo concebido como determinista. Pero, ¿cómo es posible que en un mundo dominado por la explicación física de los hechos tenga cabida la acción llevada a cabo desde la vida mental de las máquinas (humanas)? Esto es, si concebimos que todo hecho del mundo tiene una explicación física, ¿cómo podemos atribuir poder causal a lo mental? Dennett concibe al ser humano como intencional; esto es, explica las actitudes comportamentales de un agente desde sus propios propósitos y fines. Aunque dicha intencionalidad es concebida como derivada (Dennett, 1987): lo que los agentes suponen como sus intereses en realidad son fines debidos a la acumulación aleatoria y arbitraria de modificaciones ambientales indicadas por algoritmos que permitieron la configuración de un cierto tipo de orden, una rutina, capaz de crear, desarrollar y hacer evolucionar cierto tipo de organismos que consiguieron elevarse hasta la posibilidad de controlar su propio desarrollo (vista desde el punto de vista filogenético).

Es decir, todo agente es un artefacto, incluidos los seres humanos. La única diferencia existente entre la inteligencia humana y algún otro tipo de inteligencia (artificial) es que esta última ha sido creada por la primera, mientras que la inteligencia humana es producto de nuestros genes egoístas (Dennett, 1995). Pero todos responden al mismo principio: su intencionalidad responde a los intereses de quien los diseñó.

Dennett afirma que la cultura humana es producto de la evolución memética. En la evolución del ser humano como organismo biológico, apareció en un determinado momento una serie de mecanismos capaces de adaptarse al medio según los elementos externos (culturales) presentes, mecanismos necesarios para la supervivencia de la especie. Dichos mecanismos fueron incorporados a la estructura fenotípica de la especie, como elementos que permiten la acción. Así, lo que en un principio parecían elecciones libres de los individuos, en realidad son respuestas de los mecanismos implementadores de la acción, mecanismos que a pesar de ser flexibles y adaptables a los contextos tienen una explicación física.

Sin embargo, la dificultad añadida surge cuando se pretende la negación del mencionado PPA. Tendemos a pensar que somos libres y que en un momento dado podríamos haber elegido actuar de modo diferente a como realmente actuamos, por lo que cuando no podemos hacer otra cosa consideramos que se ha dado un cierto tipo de determinismo puntual que impide la posibilidad de actuar de otro modo. Dennett niega el PPA porque, por un lado, no existen evidencias que indiquen que una persona podría, o puede, haber hecho otra cosa que lo que realmente hace y porque, por otra, la responsabilidad del agente respecto de la acción no remite aunque éste no pudiera haber actuado de otro modo. Parece que la diferencia entre poder haber hecho otra cosa y no poder haberla hecho es el determinismo, pero no es así para Dennett, la responsabilidad es exactamente la misma si estamos determinados como si no lo estamos, porque en un momento dado nosotros hicimos lo que hicimos por nuestro diseño.

Actuamos como actuamos porque tenemos los contenidos mentales que tenemos, pero ello no supone una reducción de la responsabilidad de nuestra acción porque responde a nuestra propia historia cognitiva, la acción nos pertenece por estar causada por los contenidos cognitivos del agente. Es una ilusión pensar que si estamos determinados no somos responsables de lo que hacemos porque es nuestra estructura la que causa las acciones que realizamos y nuestra estructura depende de nuestra historia cognitiva y del modo en que consideremos las enseñanzas. Para Dennett, el PPA es una ilusión porque no podemos nunca hacer otra cosa que la que realmente hicimos, pero no por ello dejamos de ser responsables de lo que hacemos. Parece que la solución para la toma de decisiones es estar diseñados para poder actuar del modo en que debemos, como hemos mostrado anteriormente.

Pero aquí Dennett parece suponer la existencia de una causa completa para toda acción. Es decir, supone que dada una determinada acción existe un, y sólo un, modo factible de explicación. Pero, si todo hecho físico, en este caso una acción humana, puede ser explicado a partir de una causa física, entonces cualquier pretensión de causalidad de lo mental quedaría anulada. Cuando negamos la posibilidad de actuar de otro modo distinto a como en realidad actuamos, damos por supuesto que no tenemos otra salida que realizar la acción que hacemos; esto es, que estamos determinados.

Cuando hablamos de acciones correctas no importa en absoluto que no pudiéramos hacer otra cosa, porque damos por supuesto que lo que cabe hacer en dicha situación es lo que haría cualquiera de nosotros en el mismo caso. Pero no ocurre así con acciones que consideramos incorrectas. Cuando por lo que nos preguntamos es por la censura de un cierto acto, sí tiene importancia el poder llegar a determinar que un cierto agente no tenía más opciones que realizar el acto que definitivamente hizo. Porque si llegáramos a la conclusión de que el agente tenía otras posibilidades de acción, entonces cabe la posibilidad de que el agente se equivocara al hacer lo que hizo y, por lo tanto, ser un acto censurable. Si, por el contrario, demostramos que no es posible

que el agente actuara de modo diferente, entonces demostramos que el agente estaba obligado a actuar como lo hizo y desaparece (o, por lo menos, disminuye) la responsabilidad de sus actos.

Vemos ahora que lo que realmente le ocurre a Dennett es que, en su intento por desligarse de la antigua noción de yo como agente-mónada, se queda todavía a un paso, ya que aunque es consciente de que no existe una finalidad desde el punto de vista filogenético, se queda corto al no conseguir romper definitivamente con la tradición metafísica. Requerimos de una ruptura todavía más radical que la dennettiana para lograr dar con un concepto humano de responsabilidad moral. Al dar demasiada importancia al marco naturalista y evolutivo y descuidar el marco social y cultural, Dennett no tiene en cuenta el modo en que el sujeto se entrena moralmente para formar parte del mundo al que pertenece; esto es, no tiene en cuenta la capacitación para llegar a ser considerados como seres responsables.

Lo que se trataría es de hacer referencia al marco socio-cultural como un proveedor de mecanismos reflexivos y de mecanismos no-reflexivos (al modo en que lo defienden Fischer & Ravizza) en vistas a conseguir lo que la sociedad realmente conlleva: la consecución de la mayor cantidad posible de agentes responsables moralmente. Lo que sería deseable es establecer una concepción del yo que tuviera en cuenta ambos marcos (el naturalista y el sociocultural) en tanto que permitiría que el agente tuviera responsabilidad moral de sus actos, aunque no fuera *causa sui*, dado que los condicionantes sociales e históricos (los estados mentales que escaparían a su control) serían entendidos no como propiedades de un agente esencialmente determinado, sino como condiciones de configuración del individuo y posibilidades abiertas a la deliberación y actuación de un sujeto sociohistórico (lo que dotaría del suficiente autocontrol al sujeto) que posibilita un mayor o menor grado de desarrollo moral. Es decir, sería conveniente instaurar una nueva imagen del sujeto en la que no aparezca tan separado de los factores sociales, históricos y políticos que lo capacitan y constituyen como agente moral y que le permiten ser lo que es.

REFERENCIAS

- Dennett, D. C. (1984a). *Elbow Room: the Varieties of Free Will Worth Wanting*. Cambridge (Ma.): MIT Press.
- Dennett, D. C. (1984b). I could not have done otherwise – so what? *The Journal of Philosophy*, 81, 553-567.
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge (Ma.): MIT Press.
- Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little Brown.
- Dennett, D. C. (1995). *Darwin's Dangerous Idea: Evolution and the Meaning of Life*. Boston: Simon and Schuster.
- Dennett, D. C. (1996). *Kinds of Minds: toward an Understanding of Consciousness*. Londres: Basic Books.
- Fischer, J. M. & Ravizza, M. (1998). *Responsibility and Control*. Cambridge: Cambridge U. P.
- Frankfurt, H. G. (1982). *Freedom of the Will and the Concept of a Person*. En G. Watson (ed.), *Free Will*. (pp. 81-95). New York: Oxford U. P.
- Honderich, T. (2002). Determinism as True, both Compatibilism and Incompatibilism as False, and the Real Problem. En R. Kane (ed.), *Oxford Handbook of Free Will*. (pp. 461-476). Oxford: Oxford U. P.
- Hume, D. (1988). *Tratado II*. España: Tecnos.
- Kane, R. (1996). *The Significance of Free Will*. New York/Oxford: Oxford U.P.
- Strawson, G. (1986). *Freedom and Belief*. Oxford: Clarendon Press.
- Strawson, G. (1994). The Impossibility of Moral Responsibility. *Philosophical Studies*, 75, 5-24.
- Watson, G. (1982). Free Agency. En G. Watson (ed.), *Free Will*. (pp. 96-110). New York: Oxford U. P.
- Watson, G. (1987). Responsibility and the Limits of Evil. En F. Schoeman (ed.), *Responsibility, Character and the Emotions*. (pp. 256-286). New York: Cambridge U. P.
- Williams, B. (1993). *La fortuna moral*. México D. F.: UNAM.
- Wolf, S. (1987). Sanity and the Metaphysics of Responsibility. En F. Schoeman (ed.), *Responsibility, Character and the Emotions*. (pp. 46-62). New York: Cambridge U. P.
- Wolf, S. (1990). *Freedom within Reason*. Oxford: Oxford U. P.

jjcolomina@hotmail.com

Vicente.Raga@uv.es