

De ladrones y reglas. (Una visión del problema del «*Sensible Knave*» desde un utilitarismo de la regla atemperado)

About Thieves and Rules. (A View of the «*Sensible Knave*» problem from a temperate rule Utilitarianism)

JOSÉ L. TASSET*

Resumen: En el intento de defender una interpretación de la ética y la filosofía política de David Hume cercana al utilitarismo clásico, se interpone de un modo clave el llamado problema del «*Sensible Knave*» planteado por este autor al final de su obra más utilitarista, la *Investigación sobre los Principios de la Moral*.

Según la interpretación clásica de este fragmento, la racionalidad utilitarista en el ámbito político colisionaría con la moral volviendo a ésta última inútil. Por tanto, en el ámbito político la defensa de un utilitarismo moral sería una tarea autocontradictoria.

Con el fin de mostrar que, primero, Hume no dice nada parecido a esto y, segundo, que incluso señala la forma de superar esta aparente contradicción entre moralidad y racionalidad, se analizan brevemente los argumentos de los que procede básicamente esta interpretación «antiutilitarista» estándar y se defiende una interpretación de la discusión humeana sobre el problema del aparente conflicto entre moralidad y racionalidad, o de los incentivos racionales para el comportamiento inmoral, que permite explicar mejor la posición de Hume sobre este problema.

Finalmente, se propone una instancia de superación de la contradicción moralidad/racionalidad por medio de un utilitarismo de la regla *atemperado* centrado en la idea del «desarrollo progresivo de instituciones artificiales de refuerzo de la moralidad», que el propio Hume sugeriría

Abstract: In the attempt of defending an interpretation of David Hume's moral and political philosophy connected to classical utilitarianism, intervenes in a key way the so called problem of the «*Sensible Knave*» raised by this author at the end of his more utilitarian work, the *Enquiry Concerning the Principles of Morals*.

According to the classic interpretation of this fragment, the utilitarian rationality in politics would clash with morality turning useless the latter. Therefore, in the political area the defense of a moral utilitarianism would be an auto contradictory task.

In order to show that, first, Hume does not say anything similar to this and second, that even indicates the way of overcoming this apparent contradiction between morality and rationality, we analyze briefly the arguments from which there comes basically this «anti-utilitarian» standard interpretation, and we defend an interpretation of the humean discussion on the problem of the supposed conflict between morality and rationality, or of rational incentives for immoral behavior, which allows to explain better Hume's position on this problem.

Finally, we propose an instance of overcoming the contradiction morality/ rationality by a rule adjusted utilitarianism centered on the idea of the «progressive development of artificial institutions of reinforcement of morality», that Hume himself would suggest in other places in which he

Fecha de recepción: . Fecha de aceptación: .

* Universidade da Coruña, Facultad de Humanidades, Campus de Esteiro, 15403, Ferrol (A Coruña). Correo electrónico: jose.tasset@udc.es

en otros lugares en los que aborda el tema de la contradicción aparente entre «morality» y «knavery». Se proponen también posibles líneas de desarrollo futuro de esta idea, entre ellas su uso para aclarar la posible relación del pensamiento de David Hume con ciertas formas de liberalismo contemporáneo.

Palabras clave: David Hume, Utilitarismo de la regla, *Sensible Knave*, David Gauthier, Annete Baier, John Rawls, Instituciones, Árbitros Institucionales, Liberalismo, Socialismo.

approaches the topic of the apparent contradiction between «morality» and «knavery». We propose also possible lines of future development of this idea, between them its use to clarify the relation of David Hume's thought with certain forms of contemporary liberalism.

Key Words: David Hume, Rule Utilitarianism, *Sensible Knave*, David Gauthier, Annete Baier, John Rawls, Institutions, Institutional Referees, Liberalism, Socialism.

1. Introducción¹

David Hume, de cuyo nacimiento se celebrarán 300 años en 2011, es uno de los pensadores clásicos no sólo con mayor presencia teórica en las historias de la filosofía, sino con una presencia activa y real en las discusiones más actuales de la filosofía y de las ciencias sociales. Y ello es debido no sólo a la importancia de sus planteamientos teóricos, sino también a lo acertado de la elección y diseño de los ejemplos, imágenes y metáforas que pueblan sus obras.

En concreto, por ejemplo, toda la filosofía moral y política de David Hume puede explicarse como *una teoría del desarrollo de instituciones complejas de interacción social a gran escala a partir de esquemas de cooperación binaria*, y por tanto de raíz individualista, partiendo de dos textos famosos y geniales en los que el filósofo escocés idea un par de ejemplos o situaciones prácticas que sirven de verdaderas «bombas de intuición» para el desarrollo de la teoría. Estos ejemplos no sólo le servían para mostrar los elementos de su teoría sino para superar las dificultades de las teorías contrarias, en su caso, principalmente el contractualismo, tanto de raíz lockiana como de raíz hobbesiana.

En un famoso texto del *Tratado*, dice Hume:

«Vuestro grano madura hoy, el mío madurará mañana. Es provechoso para ambos que yo trabaje hoy contigo y que tú me ayudes mañana. No siento cariño ninguno por ti y sé que tú no lo sientes tampoco por mí. No debo, por consiguiente, preocuparme de tus cosas, sino que debo trabajar contigo por mi interés y esperando una acción recíproca; sé que puedo engañarme y que en vano esperaré tu gratitud. Así, pues, te dejo trabajar solo y me conduzco de la misma manera. La estación cambia y ambos perdemos nuestras cosechas por la falta de confianza mutua y seguridad.»²

1 Este trabajo se ha concebido y desarrollado dentro del proyecto de investigación MICINN referencia FFI2008-06414-C03-01. Se ha beneficiado de una estancia de investigación como *Honorary Research Fellow* en el Bentham Project, University College, Londres subvencionada parcialmente por el Vicerrectorado de Investigación de la Universidade da Coruña.

2 THN, 3.2.5.8. Todas las traducciones castellanas del *Tratado* proceden de David Hume, «Tratado De La Naturaleza Humana (Trad. De Vicente Viqueira, Revisada Y Actualizada),» en *David Hume: Obra Filosófica*, ed. José L. Tasset y Raquel Díaz Seijas (Madrid: Gredos, (en prensa)).

Y en otro texto aún más citado, que Hume repitió casi sin modificaciones tanto en el *Tra-tado* como en la *Investigación sobre los principios de la moral*, éste vuelve a insistir en que:

«Dos hombres que mueven los remos de un bote lo hacen por un acuerdo o conven-ción, aunque nunca se hayan prometido nada mutuamente. La regla que concierne a la estabilidad de la posesión no se deriva menos de las convenciones humanas, dado que surge gradualmente y adquiere fuerza en una lenta progresión por nuestra experiencia repetida de los inconvenientes de no cumplirla. Por el contrario, esta experiencia nos asegura aún más que el sentimiento de interés se ha hecho común entre nuestros conciudadanos y nos proporciona confianza en la regularidad futura de su conducta; solamente en la expectativa de ésta se basan nuestra moderación y abstinencia. Del mismo modo todas las lenguas se van estableciendo gradualmente por las convenciones humanas sin ninguna promesa. De igual forma el oro y la plata llegan a ser medidas corrientes de cambio y son considerados como pago suficiente de lo que vale cien veces más.»³

En ambos casos nos encontramos resumida en ejemplos en apariencia simples pero muy elaborados los elementos de la compleja teoría ético-política de Hume, que aquí no voy a desarrollar en sus detalles, pero que constituye un elemento crucial de las actuales teorías de la justicia y, en mi opinión, el sustrato básico del utilitarismo de la regla:⁴ debemos cooperar por motivos directos y básicos, naturales, en el dominio de la interacción primaria en el ámbito próximo, pero tenemos que construir reglas, basadas en un principio de utilidad a largo plazo, en el ámbito de la cooperación a gran escala o con desconocidos. En el ejemplo de los dos agricultores se muestra como la ausencia de reglas que regulen la cooperación y que contribuyan a la creación de «confianza» entre las partes, conduce a un desastre seguro, mientras que en la famoso «boat example» la demostración progresiva de la tendencia individual hacia un comportamiento cooperativo es la clave del éxito individual y colectivo.⁵

Pero, la racionalidad utilitarista en la que tanto ha influido el pensamiento de Hume no está exenta de paradojas y contradicciones; incluso su reconocimiento de éstas es una de sus principales virtudes.

Precisamente David Hume señaló una de estas paradojas, la aparente contradicción entre moralidad y racionalidad a la que parece conducir la justificación utilitarista de las reglas de justicia, en un célebre pasaje de la *Enquiry Concerning the Principles of Morals*, en el

3 THN, 3.2.2.10. Ibid.

4 Este trabajo parte de lo analizado previamente en algunos de mis trabajos, principalmente en: José L. Tasset, «La Ética, Las Pasiones Y La Razón. En Defensa Del Gobierno Ético De La Razón,» en *Decidir En Socie-dad. Homenaje a Julia Barragán*, ed. Fernando Aguiar, Francisco Lara, and Nelson Lara (Caracas: Ediciones Chyrime-k, C.A., 2009).

5 Recapitulo aquí de un modo mucho más esquemático lo dicho en José L. Tasset, *La Ética Y Las Pasiones (Un Estudio De La Filosofía Moral Y Política De David Hume)* (A Coruña: Universidade da Coruña, 1999); cfr. Cap. 5. Una ampliación de lo dicho en ese trabajo mucho más abierta al diálogo con las innovaciones derivadas de la teoría de la decisión racional y de la *Game Theory* en José L. Tasset, «La Ética Y Las Pasiones: Una Propuesta Neohumeana,» en *Los Sentimientos Morales [Cuaderno Gris, Época Iii (No. 7)]*, ed. Pura Sánchez Zamorano (Madrid: Universidad Autónoma de Madrid, 2003). Y sobre todo en Tasset, «La Ética, Las Pasiones Y La Razón. En Defensa Del Gobierno Ético De La Razón.»

que Hume imagina el punto de vista de un sujeto al que denomina «Sensible Knave»⁶ o «bribón inteligente».⁷ Por «Knave» creo entender que Hume quería decir «egoísta irrestricto y racional». Este fragmento normalmente ha sido usado para demostrar que Hume no puede ser utilitarista; yo voy a intentar demostrar aquí que el problema (no dilema) que plantea sólo puede ser entendido, paradójicamente, desde un cierto utilitarismo de la regla.⁸

2. Planteamiento del problema y análisis textual

Este tipo de utilitarismo de reglas o de la regla, o utilitarismo indirecto, es uno de los grandes hallazgos de la compleja historia del utilitarismo. Esta sofisticada modificación del utilitarismo directo, creo, es también una de las claves de la pervivencia actual del modelo utilitarista de racionalidad y de la ética utilitarista en su conjunto.

Pero, David Hume, quien *sin ser propiamente un utilitarista es una de las figuras clave en el nacimiento de esta teoría*, era amigo también de plantear, casi siempre en las páginas finales de sus obras, paradojas o contra-argumentos para sus propias ideas y teorías, en una suerte de *auto-test escéptico* que debía prevenir, de un modo muy literario, al lector contra la solidez final y definitiva de toda teoría, incluida la suya.

Así, al final de la *Enquiry Concerning the Principles of Morals*,⁹ sin duda una de las mejores obras de la historia de la Ética, David Hume se pregunta, casi a vuelapluma, si *cuando una regla está bien establecida y tiene una preeminencia amplia, casi universal, no habría un fuerte argumento racional (incentivo positivo diríamos hoy), de tipo individual al menos, para no cumplirla*.¹⁰

6 En adelante y para abreviar llamaremos a nuestro «Sensible Nave» simplemente «SK».

7 Hemos tomado en este trabajo en la mayoría de los casos la traducción que realiza Gerardo López Sastre de «Sensible Nave» como «bribón inteligente» porque el término «bribón» es clásico en la literatura sobre *Game Theory* con la que relacionamos en este trabajo la discusión llevada a cabo por Hume; por eso la hemos preferido; puede ser traducido también como «ladrón prudente», «inteligente» o «racional». En la edición española de las *Lecciones de filosofía política* de John Rawls se traduce como «truhán avisado» (John Rawls, *Lecciones Sobre La Historia De La Filosofía Política*, ed. Samuel Richard Freeman, trans. Albino Santos Mosquera (Barcelona [etc.]: Paidós, 2009): 247).

8 El concepto de «dilema» parece implicar un problema práctico como el planteado por Hume en el fragmento que nos ocupa, pero que no admite una única solución racional. Como intento mostrar no es el caso: la teoría de Hume está preparada para abordar el problema, obviamente en los términos de la propia teoría; así debe entenderse creo el comentario de Hume (v. más adelante el texto completo) acerca de la imposibilidad de dar razón de un modo externo a la teoría de la aparente contradicción entre «moralidad» y «racionalidad». Debo a Gerardo López Sastre la oportuna sugerencia sobre la falta de adecuación del término «dilema» en lo que se refiere al «problema» del SK.

9 La edición reciente más completa y rigurosa es: David Hume, *An Enquiry Concerning the Principles of Morals: A Critical Edition*, ed. L. Beauchamp Tom, *The Clarendon Edition of the Works of David Hume* (Oxford: Clarendon Press : Oxford University Press, 1998). Para el Tratado de la Naturaleza Humana empleamos la edición siguiente: David Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary J Norton, *Oxford Philosophical Texts* (Oxford, New York: Oxford University Press, 2000). En adelante se citarán la segunda investigación y el Tratado de acuerdo con la numeración marginal de estas nuevas ediciones. La versión castellana empleada de la segunda investigación es la de Gerardo López Sastre, cfr. David Hume, *Investigación Sobre Los Principios De La Moral*, ed. Gerardo López Sastre (Madrid: Espasa Calpe, 1991).

10 Aunque este famoso ejemplo aparece al final de la *Enquiry Concerning the Principles of Morals*, sin embargo, el tema de la cooperación entre agentes morales y el desarrollo progresivo de una estructura de armonización de sus comportamientos, así como del papel del egoísmo y el altruismo, a la vez como motivación y como impedi-

Ese es el sentido fundamental del conocido pasaje del «SK» contenido en la *Investigación sobre los Principios de la Moral* y, en principio, constituye una objeción para la propia teoría de la justicia defendida por Hume.

Uno de los primeros problemas que nos encontramos a la hora de evaluar la importancia del problema del SK es que David Hume suele ser citado de modo muy fragmentario, lo que de un modo general da lugar en muchos casos a malas comprensiones de algunos de sus planteamientos específicos y de muchas de sus intenciones generales.¹¹ Para evitar esto citaremos primero el texto relativo al «SK» de un modo específico y luego ampliaremos la cita añadiendo el resto de la discusión humeana de este tema, para que el problema que plantea Hume pueda ser contemplado en su contexto general, que no parece ser la insistencia en las *limitaciones racionales de la moral*, sino más bien el de afirmar la complejidad de cualquier punto de vista sensato de acercamiento a este fenómeno, que no es agotable sólo desde la racionalidad o moralidad individuales, que no valen para el fenómeno de la Justicia, sino que deben abrirse, por ejemplo, a la racionalidad estratégica o colectiva, característica de la interacción social entre seres humanos, todos ellos evidentemente seres morales, esto es, emocionalmente abiertos a otros, pero también *racionales y auto-interesados*, emocionalmente preocupados por sí mismos, especialmente en el contexto de la interacción con otros seres humanos iguales que ellos mismos.

El brillante texto de la EPM, al final, dice lo siguiente:

«Tratando al vicio con el mayor candor y haciéndole todas las concesiones posibles, tenemos que reconocer que no hay en ningún caso el pretexto más mínimo para, desde la perspectiva del interés propio, concederle preferencia sobre la virtud; excepto, quizá, **en el caso de la justicia, donde, considerando las cosas desde un cierto punto de vista, a menudo puede parecer que un hombre sale perdiendo por su integridad.** Y, aunque se admita que sin el respeto por la propiedad ninguna sociedad podría subsistir, sin embargo, de acuerdo con la forma imperfecta en que se dirigen los asuntos humanos, **un bribón inteligente puede pensar en casos particulares que un acto de iniquidad e infidelidad puede proporcionar una adición considerable a su fortuna sin causar una brecha considerable en la confederación o unión social.** *Que la honestidad constituye la mejor política* puede ser una buena regla general, pero es susceptible de muchas excepciones. Y quizás pueda pensarse que se conduce con más sabiduría quien observa la regla general y se aprovecha de todas las excepciones.»¹²

mento para el logro de ésta, ocupó a Hume a lo largo de toda su obra. He analizado estos temas anteriormente en el contexto de una reconstrucción de la teoría humeana de las instituciones morales y políticas en el capítulo 5 de Tasset, *La Ética Y Las Pasiones (Un Estudio De La Filosofía Moral Y Política De David Hume)*.

11 Una interpretación banal del fragmento precisamente por su lectura meramente fragmentaria, puede encontrarse por ejemplo en James Baillie, *Hume on Morality, Routledge Philosophy Guidebooks* (London ; New York: Routledge, 2000): 173.

12 «Treating vice with the greatest candour, and making it all possible concessions, we must acknowledge, that there is not, in any instance, the smallest pretext for giving it the preference above virtue, with a view to self-interest; except, perhaps, **in the case of justice, where a man, taking things in a certain light, may often seem to be a loser by his integrity.** And though it is allowed, that, without a regard to property, no society could subsist; yet, according to the imperfect way in which human affairs are conducted, **a Sensible Knave, in**

Es llamativo en una primera inspección de la literatura sobre este asunto no sólo que la mayoría de los intérpretes olviden que Hume habla de problemas similares o casi del mismo asunto en otros lugares de su obra,¹³ sino que incluso a mayores se olviden de colocar el llamado fragmento sobre el «SK» en el contexto tan sólo un poco más amplio de la sección de la EPM (9.2.) a la que pertenece, lo que ayudaría bastante no sólo a comprender su significado sino especialmente la solución que el propio Hume puede dar a este problema. En términos más precisos, el fragmento va precedido de un importante texto al inicio de la sección y de una conclusión al final que raramente se citan y que permiten en mi opinión ir más allá del problema del SK.

En el texto que antecede al fragmento Hume coloca el concepto de *interés real o auténtico de los individuos* en el centro de la discusión, lo que evidentemente implica que el SK, como otros muchos seres humanos, puede verse cegado por el atractivo de lo inmediatamente beneficioso (desertar, no cooperar, ser un egoísta irrestricto) olvidando su «true interest», que sería más distante y derivado del cumplimiento generalizado de las reglas. Dice Hume: «... ¿qué teoría de la moral puede servir a algún propósito útil si no puede mostrar mediante algún detalle concreto que todos los deberes que recomienda constituyen también el **verdadero interés** de cada individuo?»¹⁴

Esto es, resulta evidente que moralidad y racionalidad deben perseguir objetivos comunes,¹⁵ pero ello no significa que necesariamente una consideración miope y restringida del interés deba constituir el único modelo de racionalidad ni en consecuencia de moralidad.

particular incidents, may think, that an act of iniquity or infidelity will make a considerable addition to his fortune, without causing any considerable breach in the social union and confederacy. That honesty is the best policy, may be a good general rule; but is liable to many exceptions: And he, it may, perhaps, be thought, conducts himself with most wisdom, who observes the general rule, and takes advantage of all the exceptions.» Hume: EPM, 9.2.10. [Negritas mías.] Traducción castellana procedente de Hume, *Investigación Sobre Los Principios De La Moral.*: 155-156.

- 13 Contra la tendencia dominante, este hecho es tenido en cuenta de un modo muy adecuado en Daniel- M. Hausman, «Rationality and Knavery.» *Vienna Circle Institute Yearbook* (1997), quien cita oportunamente el comienzo del ensayo de Hume sobre «La independencia del Parlamento»: «Political writers have established it as a maxim, that, in contriving any system of government, and fixing the several checks and controuls of the constitution, every man ought to be supposed a knave, and to have no other end, in all his actions, than private interest. By this interest we must govern him, and, by means of it, make him, notwithstanding his insatiable avarice and ambition, co-operate to public good. (...) It is, therefore, a just political maxim, that every man must be supposed a knave: Though at the same time, it appears somewhat strange, that a maxim should be true in politics, which is false in fact.» Hume: EPM Pt. 1 E. 6 Para. 1/8. Hay trad. esp. en: David Hume, *Ensayos Políticos*, ed. Josep María Colomer, trans. César Armando Gómez (Madrid: Tecnos, 1987), David Hume, *Ensayos Políticos*, ed. Enrique Tierno Galván, *Civitas / Instituto De Estudios Políticos* (Madrid: Instituto de Estudios Políticos, 1955).
- 14 Hume: EPM, 9.2.16. [Negritas mías.] Traducción castellana procedente de Hume, *Investigación Sobre Los Principios De La Moral.*: 153.
- 15 El hablante ordinario a veces tiende a pensar que la moral, al ser una cuestión que tiene que ver con nuestros puntos de vista subjetivos, constituye una especie de puerta abierta a la arbitrariedad, a lo que Hume denominaba «confusión de la libertad con la indeterminación absoluta». No es así, al menos no de un modo general. Por ejemplo, es evidente que si un sujeto prefiere «a» a «b» (a<b) y «b» a «c» (b<c), deberá preferir «a» a «c» (a<c). La transitividad de las elecciones o preferencias desde un punto de vista inmediato es, de un modo evidente, un requerimiento de cualquier argumentación, elección o acción morales; no obstante, un cambio temporal en las circunstancias, o una comparación de los distintos bienes en distintos momentos entre sí, podría introducir excepciones en esta regla, que aun así debe cumplir nuestra moral de un modo general. En principio, pues, la moral en cualquiera de sus dimensiones está sometida a los requisitos de cualquier actividad racional.

En todo caso, Hume es consciente de que aunque en el ámbito de la justificación de la moral, esto sea o parezca ser suficiente, puede ocurrir que no podamos convencer y por tanto motivar a alguien cegado por esa consideración restrictiva y casi infantil del interés, en suma, si no compartimos unos ciertos supuestos morales básicos, el diálogo moral con alguien inmoral puede ser absurdo e inútil. Esto, que Hume ya había advertido en algún otro lugar de su obra,¹⁶ es precisamente lo que le lleva a la conclusión aparente de que quizás nuestros argumentos sean insuficientes para lograr un cambio en las motivaciones profundas del SK; más, esto no debe interpretarse como una derrota sino como un reconocimiento de que, por una parte, la razón no siempre puede modificar las pasiones y de que, en segundo lugar, no siempre las pasiones apacibles gobiernan sobre las violentas, en fin, de que aunque nuestra capacidad de simpatía o benevolencia sean el germen del desarrollo y de la riqueza expansiva de nuestra experiencia moral, a veces no pueden luchar contra el egocentrismo y la cortedad de miras. En esos casos, sólo nos cabe la admiración ante el sujeto moral, el desprecio ante el SK y, eso sí, el desarrollo de instituciones que le hagan ver y sufrir las desventajas de su erróneo punto de vista.

«He de reconocer que si un hombre piensa que este razonamiento está muy requerido de una respuesta, será un poco difícil encontrar alguna que le resulte satisfactoria y convincente. Si su corazón no se rebela contra tales máximas perniciosas, si no siente aversión por los pensamientos de maldad y bajeza, ha perdido, de hecho, un motivo considerable a favor de la virtud; y podemos esperar que su práctica estará de acuerdo con su especulación. Pero en todas las naturalezas candorosas la antipatía hacia la perfidia y la truhanería es demasiado fuerte para ser contrapesada por perspectivas de beneficios o ventajas pecuniarias. La paz interior del espíritu, la conciencia de inte-

Hume se dio cuenta de que esto no ocurría siempre al revés y que la moralidad no es necesariamente un requisito, digamos interno, de la racionalidad. Un simple ejemplo mostrará lo que quiero decir: Si prefiero «asesinar a 100 personas» (a) que «asesinar a 50 personas» (b), entonces será perfectamente racional realizar el razonamiento transitivo de que prefiero «asesinar a 50 personas» (b) que «asesinar a 25 personas» (c); el establecimiento de una premisa limitativa de este razonamiento según la cual «no debería preferir asesinar a ninguna persona», o al estilo del mandato de James Griffin (pero con un giro perverso), que debemos seguir la regla de «Limit the Damage» tratándose de vidas humanas, tendría que ser establecida de modo independiente y sería externa, desde un punto de vista formal, a la mera aplicación de la regla de transitividad.

Este creo que es el sentido básico del razonamiento humeano, muy famoso y mal entendido, relativo a la no irracionalidad de la preferencia por no sufrir ningún daño en mi dedo frente a la destrucción del mundo. Sólo se podría quizás objetar que mi dedo forma parte del mundo (es un bien «inclusivo») y por tanto quizás se olvida que la destrucción de éste incluiría la de mi dedo e incluso de la mano que lo soporta. Sería más claro haber escogido un ejemplo del tipo siguiente: no es irracional preferir no sufrir ningún daño en mi dedo a que se le corten las manos a todos los miembros de mi Universidad, menos a mí, ya que mi mano contiene el famoso dedo y lo necesito para seguir escribiendo este artículo.

- 16 Por ejemplo de un modo retóricamente muy parecido a la parte final de EPM, 9.2. lo hace también en el ensayo «El escéptico»: «...cuando uno nace con una constitución mental tan perversa, con una disposición tan cruel e insensible, como para no tener inclinación ninguna por la virtud y la humanidad, ninguna simpatía por sus prójimos, ningún deseo de estima o aplauso; tal individuo debe ser reconocido como enteramente incurable, y no existe remedio alguno en la filosofía. (...) Debo repetirlo: mi filosofía no ofrece ningún remedio en tal caso, y no podría sino lamentar la infeliz condición de tal persona.» En David Hume, *Disertación Sobre Las Pasiones Y Otros Ensayos Morales*, ed. José Luis Tasset Carmona, 2a (1ª ed.: 1990) ed., *Textos Y Documentos (Anthropos)* (Barcelona: Anthropos, 2004): 248-251.

gridad, un examen satisfactorio de nuestra propia conducta, éstas son circunstancias que resultan muy necesarias para la felicidad, y que serán apreciadas y cultivadas por todo hombre honesto que sienta su importancia.

Tal persona tiene, además, la frecuente satisfacción de contemplar a bribones que, a pesar de toda su astucia y habilidades, son traicionados por sus propias máximas, y que, aunque su propósito es estafar de una forma secreta y controlada, se presenta un acontecimiento tentador, la naturaleza es frágil, y caen en la trampa; de donde nunca logran salir sin la pérdida total de su reputación y de toda futura confianza y fe por parte de la humanidad.

Pero aunque consiguieran obrar en secreto y tener éxito, el hombre honrado, si posee algún barniz de filosofía, o incluso de la reflexión y la observación comunes, descubrirá que los bribones resultan al final las mayores víctimas, y que han sacrificado el disfrute inestimable de una reputación, en lo que se refiere al menos a ellos mismos, por la adquisición de juguetes y chucherías sin valor. ¿Qué pocas cosas son necesarias para satisfacer las *necesidades* de la naturaleza? Y, en lo que respecta al *placer*; ¿qué comparación entre la satisfacción no comprada de la conversación, las relaciones sociales, el estudio, incluso de la salud y de las bellezas usuales de la naturaleza; pero sobre todo la reflexión tranquila sobre la propia conducta; qué comparación, digo, entre estas satisfacciones y las diversiones vacías y febriles del lujo y el gasto? De hecho, esos placeres naturales realmente no tienen precio; tanto porque están por debajo de todo precio en su obtención, como porque se encuentran por encima de cualquiera en su disfrute.»¹⁷

3. Diferentes interpretaciones del problema del «*Sensible Knave*»

Las interpretaciones de este problema en Hume se han organizado clásicamente en *dos bandos opuestos*, el de aquellos que han interpretado que el fragmento del *SK* es el reconocimiento por parte de Hume de una derrota y el de los que sostienen que Hume plantea una aparente paradoja final, muy a su estilo, que sin embargo tiene una respuesta en el marco de su teoría.

Los críticos de Hume son liderados por el brillante filósofo político David Gauthier.¹⁸

Los defensores de la sostenibilidad de la solución humeana al problema del *SK*, entre quienes me sitúo aunque mis intenciones son más amplias, son liderados por la no menos brillante filósofa moral Annette Baier.¹⁹

17 Hume: EPM, 9.2.11-12. [Negritas mías.] Traducción castellana procedente de Hume, *Investigación Sobre Los Principios De La Moral*: 156-157.

18 La argumentación principal de Gauthier sobre este asunto se encuentra en David Gauthier, «Artificial Virtues and the Sensible Knave,» *Hume Studies* 18, no. 2 (1992). Este trabajo ha sido reeditado conservando la paginación original en Rachel Cohon, *Hume: Moral and Political Philosophy, The International Library of Critical Essays in the History of Philosophy* (Aldershot (England): Ashgate : Dartmouth, 2001). Cfr. También David Gauthier, «Three against Justice: The Foole, the Sensible Knave, and the Lydian Shepherd,» en *Midwest Studies in Philosophy VII* (Minneapolis: University of Minnesota Press, 1982).

19 Principalmente en Annette Baier, «Artificial Virtues and the Equally Sensible Non-Knaves: A Response to Gauthier,» *Hume Studies* 18 (1992). Baier, a su vez, dice basarse en algunos puntos importantes de su análisis en Gerald-J. Postema, «Hume's Reply to the Sensible Knave,» *History of Philosophy Quarterly* (1988).

David Gauthier realiza una polémica y refinada interpretación del problema del SK tal y como lo plantea Hume al final de la *Investigación sobre los principios de la moral*. Como mi intención primordial es construir una interpretación alternativa a la de Gauthier, voy a exponer en primer lugar algunos de sus argumentos y luego procederé a discutirlos y a formular mi propia visión del problema del SK, caracterizada por incardinarlo globalmente en la teoría de la Justicia de Hume.²⁰

En una acertada síntesis de Rachel Cohon,²¹ se señala que la intención de Gauthier es argumentar que el reconocimiento de cierta circularidad en la fundamentación de la justicia obligaría a Hume a reconocer que la justicia no es en el fondo una virtud. La única forma de escapar a esta conclusión sería afirmar que Hume creería lo que Marcia Baron ha llamado la «noble mentira», esto es, la suposición, fingida y consciente, de que la justicia posee realmente un motivo natural de aprobación (que no es tal), que sería el interés propio y el deseo de posesión.

Pero, además, Hume en el pasaje del «SK» reconocería que el interés personal, así como el deseo de posesión en el que éste al parecer se encarnaría de un modo principal, se satisfaría mejor al parecer por medio de una injusticia puntual que mediante una justicia global y sin excepciones, con lo que el interés personal no podría ser el motivo natural de la justicia y ésta al carecer, de algún modo, de una motivación consistente, no podría ser considerada una virtud. Ahora bien, como la tesis de la justicia como una virtud artificial es la clave de bóveda de toda la teoría de la justicia de Hume, ésta se quedaría seriamente dañada. Retengamos, antes de proseguir, que el propio Gauthier reconoce que lo que dice Hume es que la justicia es una virtud artificial, lo que dará lugar a que su motivación no sea directamente natural sino referida a los motivos naturales de comportamiento de un modo indirecto. Así pues, la tesis de Gauthier, quien sigue a Knud Haakonssen en este aspecto,²² acerca de la necesidad ineludible de un auto-engaño como condición de posibilidad parece no sostenerse si contemplamos la teoría de la justicia de Hume de un modo global.

Gauthier comienza señalando que la discusión del problema del SK debe colocarse en el contexto de la teoría humeana de las virtudes artificiales y surge en el desarrollo del análisis de la promesa; en todo caso, es un problema que «se resiste a una interpretación fácil».²³

El hecho de que se resista a una interpretación simple obedece, sobre todo, a que, como hemos explicado, Hume parece estar planteando una paradoja, un aparente problema sin solución, derivado del hecho de que toda virtud debe estar unida a una pasión natural y la justicia no parece estar unida a ninguna; en principio, parece estar unida al interés, pero precisamente el texto relativo al «SK» plantea la paradoja de por qué cumplir con la justicia si parece haber desaparecido el interés, esto es, el único motivo o pasión unida a la justicia; en ese caso, o podemos cumplir con la justicia por su valor en sí misma, por puro

20 Muy interesante y equilibrada es la aportación al debate desde la economía que se puede encontrar en Hausman, «Rationality and Knavery.» Muy interesante, aunque también muy breve, me ha resultado la pequeña aportación al análisis del problema contenida en: Matt Matravers, *Hume on Trust* (2002); available from http://www.open2.net/trust/society/soc_hume/soc_hume1.htm.

21 «Introducción» a Cohon, *Hume: Moral and Political Philosophy.*: xvii.

22 Knud Haakonssen, *The Science of a Legislator: The Natural Jurisprudence of David Hume and Adam Smith* (Cambridge ; New York: Cambridge University Press, 1981).

23 Gauthier, «Artificial Virtues and the Sensible Knave.»: 401

sentido del deber, pero ello no constituiría un motivo (en todo caso, el sentido de la justicia no puede constituir un motivo para la justicia misma);²⁴ o podemos buscar un motivo en el propio interés, pero contemplado desde un punto de vista diferente, desde una perspectiva no a corto plazo sino a largo plazo.

En contra de lo que cualquier lector o intérprete de la teoría de la justicia humeana podría esperar, y esa es en realidad la única paradoja que llama la atención realmente, en el fragmento sobre el SK, Hume parece defender un primado de la primera clase de argumentación, cuando tenía una buena oportunidad de defender claramente la segunda, lo que coincidiría muy bien con el tono utilitarista de la regla que en general domina toda o la mayor parte de la EPM.

La contradicción aparente a la que llega Hume en lo que se refiere al problema del «motivo de la Justicia», al no estar vinculada a una pasión natural, lleva a Gauthier a plantearse su análisis del siguiente modo:

«Propongo ahora seguir dos líneas de investigación. La primera ofrece una lectura diferente de Hume, distinguiendo la obligación natural para con la justicia, la fidelidad o el cumplimiento-de-las-promesas y la obediencia de la obligación moral, así como mostrando que la obligación natural no es ficticia y que la obligación moral puede estar relacionada con ella. Esta lectura distingue la teoría de Hume sobre las virtudes artificiales de su teoría de la moralidad y relega a la última a una posición subordinada. La segunda línea de investigación comienza presentando al SK de la *Enquiry Concerning the Principles of Morals*. El bribón supone una objeción no para esta lectura diferente de Hume en cuanto tal lectura, sino para el carácter adecuado de la teoría así leída. Consideraré cómo podría contestársele, si no por el Hume de la *Enquiry*, entonces por el del *Treatise*, y esto nos llevará de vuelta a una teoría del error que socava las virtudes artificiales. Concluiré considerando si Hume podría tener una respuesta alternativa que dejase intactas a estas virtudes.»²⁵

Como vemos, Gauthier hace algo parecido a lo que lleva a cabo el Segundo Rawls, una eliminación de la teoría de la moralidad en sentido fuerte a favor de una teoría política, carente por tanto de contenido moral en sentido estricto. Asimismo, parece que el dilema humeano del «SK» es el único freno para este deslizamiento.

Gauthier prosigue, tras su planteamiento, preguntándose cómo es posible que, por una parte, Hume considere que el motivo para el cumplimiento de las promesas sea el deber mismo, o para la equidad, un aspecto de la Justicia, sea ésta misma y el mérito procedente

24 O sí; podríamos estar aquí ante otro ejemplo del *razonamiento a término* que usan en algún momento todos los utilitaristas clásicos y en el que se acaba haciendo referencia de un modo que podríamos denominar «naturalista» a una pasión básica. La única candidata que se me ocurre para este papel es el denominado «auto-interés» o «self-love» en Hume, pero ello daría a toda la argumentación humeana sobre la Justicia un cierto tono «circular», lo que a veces ciertamente parece tener. Creo que la explicación de esto se halla en la no bien resuelta articulación de naturalismo y convencionalismo que se da en general en todo el utilitarismo clásico. Circularidades similares aparecen también en Mill por ejemplo cuando pretende *probar* el «principio de utilidad».

25 Gauthier, «Artificial Virtues and the Sensible Knave.»: 403-404.

de su observancia, y también afirme que sin embargo, dicho motivo existe, y es el interés propio.²⁶

Gauthier recoge de un modo muy certero, así, los textos de Hume en que se vincula Justicia e interés (lo que curiosamente parece abocar a Hume a una teoría utilitarista de la Justicia, en contra del propio interés de Gauthier), pero olvida comentar que Hume siempre añade una consideración relativa a que dicho interés es el motivo «...first...» de la Justicia, es decir, en un primer momento evolutivo del desarrollo de las instituciones de Justicia, cuando todavía no se ha superado el límite del grupo inmediato; más allá de este grupo, se hace necesaria la moral como estructura artificial de motivación para la Justicia. Sin esa distinción de niveles y sin una percepción del juego humeano entre lo natural y lo artificial (diferentes sólo en grado no en clase), no se logra comprender, creo, la teoría de la justicia de Hume, ni tampoco su teoría general de las virtudes. A su vez para ello es necesario comprender y valorar previamente de un modo diferente la teoría humeana de las pasiones.

No obstante, no se le pasa por alto a Gauthier que Hume intenta explicar cómo surge y se establece un «nuevo tipo de obligación» diferente del que nos vincula de modo natural, en el sentido de espontáneo, a nuestros prójimos más inmediatos.²⁷

Gauthier reconoce que la clave, al menos en lo que respecta a la obligación, está en el complicado uso que hace Hume de «natural», unas veces opuesto a «artificial», otras a «moral» e incluso a «civil».²⁸

El propio Gauthier, sin embargo, se da cuenta de que la única manera de fundar la motivación de justicia, aunque no lo suscribe como tesis general, es vincular el motivo para la Justicia con el desarrollo y cumplimiento de reglas:

«La tesis de Hume es, entonces, que no poseemos ningún motivo suficiente para la realización de los actos específicos requeridos por la justicia y la fidelidad, si consideramos esos actos abstraídos de las *prácticas* que ellos ejemplifican. Y esto es decir que no tenemos ningún motivo *natural* suficiente para la realización de estos actos.»²⁹

26 Cfr. Para el desarrollo de esta argumentación en Hume T 3.2.2. 24-28, T, 3.2.8.5 y finalmente T, 3.2.5.11-15). Gauthier glosa estos textos en *ibid.*: 404.

27 «Hume habla tanto de obligación natural como de obligación moral, particularmente en su discusión de las virtudes artificiales. Dice que la 'obligación natural para con la justicia reside en el interés', pero que la obligación moral consiste en el sentimiento de lo justo y lo injusto y éste deriva de la simpatía.» Introducción a Cohon, *Hume: Moral and Political Philosophy.*: xviii. Cfr. para este argumento de los dos tipos de obligación Tito Magri, «Natural Obligation and Normative Motivation in Hume's Treatise,» *Hume Studies* 22 (1996), quien argumenta que la obligación natural es redirigida por medio de la aceptación de las reglas generales con lo que adquiere el *status* de una norma. Por su parte, Stephen Darwall, «Motive and Obligation in Hume's Ethics,» *Nôus* 27 (1993). argumenta que habría una tercera clase de obligación implícita en la teoría de Hume, la generada por el estado en el que somos gobernados por reglas, y que éste es el que proporciona el motivo específico para la justicia y le permite salir a Hume del supuesto círculo en su argumentación.

28 Aunque Gauthier piensa que eso se puede superar bien en el ámbito de la obligación, no piensa que sea tan fácil hacerlo en el de la «motivación», en donde parece haber una contradicción no resuelta (la que al fin daría lugar recientemente a la polémica del «internalismo» humeano). Cfr. Sobre este problema el muy interesante trabajo de Charlotte Brown, «Is Hume an Internalist?,» *Journal of the History of Philosophy* 26 (1988)., así como Elizabeth Radcliffe, «How Does the Humean Sense of Duty Motivate?,» *Journal of the History of Philosophy* 34 (1996).

29 Gauthier, «Artificial Virtues and the Sensible Knave»: 405.

Esta dimensión clave de las reglas es, pues, el único elemento intermedio que vuelve lógica la restricción o renuncia al propio interés que puede suponer el cumplimiento de las demandas de la Justicia en un momento avanzado, es la única forma que una racionalidad a corto plazo quede compensada por una dimensión de racionalidad a medio-largo plazo:

«(...) Un acto único de justicia es frecuentemente contrario al *interés público*, y si existiese solo, sin ser seguido de otros actos, podría ser en sí mismo perjudicial a la sociedad. (...) Tampoco es cada acto de justicia, considerado de forma aislada, más favorable para el interés privado que para el público... Pero, aunque los actos particulares de la justicia puedan ser contrarios al interés público o al privado, es cierto que el plan o esquema total es altamente favorable, y de hecho absolutamente necesario, tanto para mantener la sociedad como para el bienestar de cada individuo. (...) Aunque en un caso el público sea la víctima, este daño momentáneo se halla ampliamente compensado por la continua validez de la regla y por la paz y orden que establece ésta en la sociedad. Incluso cada persona individual debe considerarse como beneficiada al hacer balance del asunto, pues sin la justicia la sociedad se disolvería inmediatamente y cada uno caería en la condición salvaje y solitaria, que es infinitamente peor que la situación peor que podemos imaginar en la sociedad.»³⁰

En términos simples, la Justicia introduce unas reglas de un nuevo juego, no natural, en el que incluso perdiendo en una situación concreta, se gana en el término general.

Es precisamente en ese momento, una vez establecida su apuesta por un utilitarismo de la regla en lo que respecta a la justificación de los actos individuales de justicia, que recordemos pueden ser completamente contrarios al interés puntual del sujeto afectado, cuando Hume se da cuenta de que esa justificación puede ser insuficiente desde el punto de vista de la motivación, que recordemos en este momento es «completamente artificial». Es decir, dado que ya no estamos en un ámbito en el que nos muevan nuestras pasiones naturales, saber que debemos hacer lo justo ¿sería suficiente para sentir que debemos hacer lo justo?

Hume, obviamente, tiene sus dudas (si no, jamás se plantearía este problema), pero recordemos que sólo podemos superar el reto del «SK» si, por una parte, distinguimos los momentos natural y artificial del desarrollo de las instituciones de justicia y, si por otra, recordamos que son fenómenos distintos *justificar* que *motivar*.

O en los términos de Gauthier:

«Cada persona debe esperar que cada elección que lleva a cabo entre actuar conforme a las reglas de justicia o violarlas, tiene un efecto sobre la conducta de otros, con consecuencias para su propio beneficio suficientes como para ofrecerle un motivo normalmente adecuado para la conformidad.»³¹

³⁰ T, 3.2.2.22.

³¹ Gauthier, «Artificial Virtues and the Sensible Knave»: 407.

En todo caso, el juego humeano de lo natural y lo artificial se acaba invirtiendo y en el ámbito de lo distante, de la cooperación multipersonal, lo artificial, la justicia y sus reglas, acaban siendo lo natural para el ser humano:

Así, en un resumen brillante del propio Gauthier:

«Esto complete mi intento de lectura de la investigación de Hume sobre las obligaciones de equidad, fidelidad y obediencia. En cada caso existe una convención o convenciones que son abrazadas por cada persona con el fin de incrementar sus posesiones y su seguridad. Aunque la adhesión a estas convenciones exige que cada persona restrinja su interés, esta restricción es ordenada por el propio interés, redirigido por la reflexión sobre la mayor satisfacción que éste alcanza al mantener las convenciones. La obligación natural (no-moral) está basada en este interés redirigido, y esta obligación es inicialmente eficaz para mantener el orden social. Pero, como las sociedades se vuelven más complejas, las personas pueden fallar de un modo importante al apreciar la fuerza de sus obligaciones interesadas, lo que exige, en la práctica si no en la teoría, que sean reforzadas por la moralidad (y, también, por supuesto, por los magistrados).»³²

O en palabras del propio Hume en el *Treatise*:

«En resumen, pues, tenemos que considerar que la distinción entre justicia e injusticia tiene dos fundamentos diferentes, a saber: el del *interés*, cuando los hombres observan que es imposible vivir en sociedad sin gobernarse por ciertas reglas, y el de la *moralidad*, cuando este interés es ya una vez observado y los hombres experimentan placer ante la vista de las acciones que tienden a la paz de la sociedad y dolor por las que son contrarias a ella. La convención voluntaria y el artificio de los hombres son los que hacen que el primer interés tenga lugar, y, por consiguiente, las leyes de la justicia deben considerarse de este modo *artificiales*. Después de que este interés se halla establecido y reconocido se sigue *naturalmente*, y de suyo, el sentimiento de la moralidad de la observancia de estas reglas...»³³

Aquí llegados, Gauthier recapitula para dar luego el salto a la paradoja que a su modo de ver supone el «SK»:

«Comencé citando tres fragmentos de Hume. Leí el primero como si sostuviese de un modo palmario que (1) una acción puede constituir un deber sólo si hay algún motivo para llevarla a cabo distinto del sentido del deber, y (2) que no existe ningún motivo para llevar a cabo un acto de cumplimiento de una promesa más que el sentido del deber, de lo que se seguiría que (3) un acto de cumplimiento de una promesa no puede ser un deber, y (4) no existe un motivo para llevar a cabo un acto de cumplimiento de una promesa que no implique la *errónea* aprehensión de que éste constituye un

32 Ibid.: 414-415.

33 T, 3.2.6.11.

deber. A la luz de mi discusión subsiguiente ahora debería desechar estas aparentes conclusiones. (2) se transforma en la tesis de que no existe ningún motivo *natural* (no-artificial) para llevar a cabo un acto de cumplimiento-de-una-promesa distinto del sentido del deber, de lo que se sigue que (1) un acto de cumplir-una-promesa no puede ser un deber *natural*. La última frase del primer fragmento, que ignoré previamente, en efecto lleva a esta conclusión e indica que el cumplimiento-de-una-promesa debe ser un deber artificial que descansa sobre las convenciones humanas.»³⁴

Después de la sofisticada, compleja y a veces hasta algo enrevesada argumentación de David Gauthier, la pregunta al llegar a este punto es obvia: ¿y qué? Hume no es un teórico del derecho natural y precisamente la principal y más retadora conclusión de toda su teoría de la Justicia, la que más problemas le generó, es la artificialidad del fundamento de ésta y en general, y por extensión, de todas las instituciones sociales.

Ahora bien, para sacar a Hume de la situación aporética que Gauthier le atribuye hay que entender, igual que en la teoría del conocimiento, que la alternativa al naturalismo teológico (que es el fundamento del iusnaturalismo) no es el escepticismo radical, sino alguna forma de planteamiento convencionalista moderado que redefina el naturalismo no en el sentido de referir directamente la naturaleza humana a un fundamento teológico sino a una teoría de las necesidades e intereses humanos individuales y colectivos. Este sería, a mi modo de ver, el camino para entender en toda su dimensión la tesis de Hume acerca de que *el artificio constituye una segunda naturaleza para el hombre*.

Nuestro objetivo, tras esta larga exposición de la interpretación del fragmento del «SK» llevada a cabo por David Gauthier, es proponer una teoría alternativa a la suya, en primer lugar, sugiriendo una interpretación diferente del texto de Hume, y en segundo lugar, viendo como precisamente de ese mismo fragmento surge una solución al problema del posible conflicto entre moralidad y racionalidad en el contexto de la teoría humeana de la justicia. Estas dos tareas las acometeré al hilo del análisis de algunos problemas adicionales de carácter interno que la posición de Gauthier presenta.

4. Una alternativa: el «Sensible Knave» y el utilitarismo de reglas

Si recordamos el problema que Hume planteaba con el fragmento del *SK*, éste vendría a decir que:

1. «El auto-interés ilustrado nos proporciona una obligación natural para con la justicia»
2. «...el hecho de que esto sea moralmente aprobado lo convierte en una obligación moral.»³⁵
3. Pero, el *SK* «...desafía el supuesto inicial de que cada uno, de hecho, gana más por medio de la estricta conformidad con las reglas de los artificios sociales humeanos, que por medio de ocasionales, calculadas e indetectadas rupturas de éstas.»³⁶

34 Gauthier, «Artificial Virtues and the Sensible Knave»: 415.

35 Baier, «Artificial Virtues and the Equally Sensible Non-Knaves: A Response to Gauthier»: 249.

36 Ibid.: 249.

Gauthier formula en su análisis del problema del SK de Hume dos argumentos y ambos me parece que presentan serias dificultades. En primer lugar, sugiere que Hume adoptaría una postura de un cierto cinismo al confesar la derrota de la moral al final de una obra sobre los principios de la moral, que de este modo era obvia y superflua por insincera; en segundo lugar, sostendría algo aún peor, puesto que Hume no sólo sería cínico sino torpe, puesto que su solución a la objeción del SK sería contradictoria con el resto de su sistema o específicamente con su teoría de la Justicia.

Lo que Gauthier estaría argumentando es que (a) Hume estaría reconociendo que algunas justificaciones de la conducta moral tendrían una naturaleza sofisticada, sobre todo debido a su circularidad, y (b) que Hume estaría construyendo una teoría implausible acerca del acto «fingido» de la voluntad que nos llevaría al acto de cumplir conscientemente una promesa, que estaría en el fondo actuando como soporte de todo el entramado de instituciones convencionales.

Con relación a esto, es cierto que Hume a veces es consciente de una cierta circularidad, al menos aparente, en todo lo que concierne a la Justicia. Pero, también es cierto, que aunque no es nuestro objetivo explicar todo lo relativo al detallado análisis que Hume hace de las promesas en el contexto del desarrollo del mecanismo gradual y evolutivo de la convención, sí es evidente que Gauthier no explica bien la tesis de Hume, quien estaría intentando todo el tiempo eliminar, en contra del contractualismo principalmente hobbesiano, la presencia real o virtual de cualquier clase de promesa o acto explícito de prometer.³⁷

En todo caso, David Gauthier, aunque posteriormente interpreta de un modo muy crítico los motivos de Hume en el fragmento del SK, sin embargo, proporciona una interpretación general interesante y correcta de la teoría humeana de la Justicia en el *Tratado*, una versión mucho más completa y quizás más compleja que la ofertada en la segunda investigación: «...la obligación moral para con la justicia sucede a la ‘obligación natural’, en relación con la cual Hume cree que la motivación natural ha de ser el auto-interés (o de familia) colectivamente redirigido.»³⁸

Por otra parte, Gauthier también está en lo cierto cuando insiste en el hecho de que Hume sostiene que, desde el punto de vista del SK, y sólo desde ese punto de vista, su respuesta es consciente de que quizás no es completamente satisfactoria.³⁹

37 De hecho, esta exclusión premeditada es la principal marca de diferenciación entre su teoría y el contractualismo, lo que volvería verdaderamente complicado el proyecto abordado por Gauthier de mostrar la compatibilidad entre la teoría humeana de la Justicia y alguna clase de contractualismo. Cfr. Por ejemplo: David Gauthier, «David Hume, Contractarian.» *Philosophical Review* 88 (1979). Un análisis más detallado de las razones que en mi opinión apoyan la imposibilidad de hablar de un contractualismo humeano necesitaría al menos otro trabajo de la misma extensión que el actual; basta por el momento con mostrar que una comprensión contractualista de Hume altera importantes elementos de su argumentación.

38 Baier, «Artificial Virtues and the Equally Sensible Non-Knaves: A Response to Gauthier»: 430. Cfr. también Annette Baier, *A Progress of Sentiments Reflections on Hume's Treatise* (Cambridge, Mass: Harvard University Press, 1991), cap. 10 para un análisis muy similar.

39 En relación con la explicación de los desacuerdos con Gauthier, Baier remite en Baier, *A Progress of Sentiments Reflections on Hume's Treatise*. y de un modo general, al trabajo ya mencionado de Postema, «Hume's Reply to the Sensible Knave.» En apoyo de la versión de Gauthier cita a Marcia Baron, «Hume's Noble Lie: An Account of His Artificial Virtues.» *Canadian Journal of Philosophy* 12, no. September (1982). Sobre la extrema originalidad de la teoría de Hume sobre el artificio social, recomienda su trabajo Anette C. Baier, «Hume's Account of Social Artifice --Its Origins and Originality.» *Ethics* 98, no. July (1998).

En todo caso, debemos recordar que Hume, en diversos lugares, usa el argumento de que no podemos discutir de moral con quien no acepta cierto tipo de premisas morales básicas. Y el *SK* sería uno de estos.

Así, quizás habría que entender ese cierto aire de apelación a la estética que aparece en el fragmento del *Tratado*, bien es cierto que acompañándolo de una muy utilitarista apelación a que el *SK* no tiene en cuenta el «true interest», esto es, que quizás en el estadio de la cooperación en el que se formula el problema ((a) seguir cooperando pero con menos beneficios inmediatos/ (b) defraudar con menos beneficios distantes derivados del perjuicio a la regla) el *Knave* sea menos «sensible» de lo que él cree.

En todo caso, y como señala acertadamente Annette Baier, Gauthier minusvalora la solidez de la respuesta de Hume al *SK* desde el punto de vista no del defraudador, sino del cumplidor que sigue-la-regla:

«Desde el punto de vista del virtuoso miembro, pagador-de-sus-deudas, del ‘partido de la humanidad’, la respuesta de Hume es perfectamente satisfactoria. Será completamente cierto, para una persona dotada de las otras virtudes humeanas (incluida la equidad, que es incluida en la lista como una virtud *natural*), que la justicia nos proporciona un mayor bien que la injusticia juiciosa.»⁴⁰

Hume parece llevar razón en los argumentos que formula desde el punto de vista no del *SK* sino de quien podríamos llamar, al menos en apariencia, «moral fool» o «moral idiot».

Pues, aunque pudiera ser erróneo que el *SK* ponga en peligro el sistema de reglas y normas de justicia porque, por ejemplo, su comportamiento no llegue a ser conocido (aunque esto nunca será completamente seguro y tendrá que ser considerado como dato posible en el cómputo o en el cálculo), sin embargo, sigue siendo verdad que el *SK* traiciona su integridad.

¿O no? Gauthier cuenta la historia del *SK* como si fuera el supuesto individuo moral quien en un arrebato incontrolado de fetichismo de la regla sacrifica su racionalidad, y con ello al parecer su integridad, por no vulnerar las reglas. Pero, este relato es miope, falso y además no coincide de forma general con lo que Hume parecía pensar sobre el asunto específico del *SK* o del problema general del desarrollo de la justicia.

Porque, hay reglas y reglas y no todas las situaciones son iguales. La posibilidad de la aparición del *SK* y el paralelo y conflictivo mantenimiento de nuestro respeto por la regla es un tipo de caso que, en mi opinión, sólo se puede solucionar adecuadamente adoptando un *punto de vista evolutivo respecto de las reglas e instituciones*. No hay, por tanto, en él una situación de auténtico fetichismo de la regla, sino una profunda convicción de que las reglas de justicia no admiten, sin peligro de desaparición, ninguna excepción.

Un auténtico fetichismo de la regla se daría si en el cumplimiento por ejemplo de una regla yo pusiese en peligro mi vida no estando comprometida más que la regla y ningún otro tipo de bien mayor.

Un ejemplo servirá para aclarar lo que estoy diciendo. Imaginemos que un guardia de tráfico me ordena pararme por necesidades del tráfico sobre una vía de tren y se olvida de mí

40 Baier, «Artificial Virtues and the Equally Sensible Non-Knaves: A Response to Gauthier»: 430.

debido a otras incidencias del tráfico. Un tren se acerca. ¿Alguien diría que las normas del tráfico y el bien público que las soporta se verían comprometidos si desplazo prudentemente mi coche fuera de la vía del tren incumpliendo así la regla de tráfico aplicada y aplicable en este caso, según la cual siempre tienen prioridad las indicaciones de un agente de tráfico? Yo diría, una vez más, que la moralidad racional no tiene por qué devenir en idiotez moral o que un sujeto moral no tiene necesariamente que transformarse en un «moral fool».

Quizás fuese oportuno, para poner el problema planteado por Hume en su sitio exacto, señalar que el *SK* no sería más que otro ejemplo de un problema del que ya se había ocupado Hume: el de las *agresiones al sistema*. Con la diferencia de que el *SK* acepta el sistema y lo incumple, lo vulnera, esto es, comete una *agresión intrasistemática*, mientras que la guerra sin reglas, el otro caso que Hume se plantea, nos conduciría a un contexto que podríamos llamar de *agresión extrasistemática*.⁴¹

Quedándonos en el ámbito de la amenaza o agresión intrasistemática, el punto de vista del *SK* se revela incoherente, porque de la misma forma que él/ella considera que está justificado aprovechar las excepciones a las reglas, sin que ello afecte a las propias reglas, se quedaría sin ningún argumento contra otro *SK* que aprovechara una nueva excepción pero en su contra, o lo que es peor, si el propio sistema admitiera excepciones, por ejemplo, a la norma de no ejercer violencia gratuita e ilegítima contra sus miembros, excepto si se dan las circunstancias de un seguro beneficio puntual y de una también segura ausencia de publicidad.

De este modo, una buena forma de autorrefutación del argumento racional en favor de la «knavery» sería una regla que permitiera la tortura en el caso de la detección de la existencia de un *SK* y cuando se diesen las condiciones arriba descritas. La defensa de la legitimidad de las excepciones implica, como sabe ver muy bien Hume, una mala comprensión de la fragilidad y dificultad del proceso de desarrollo de las normas e instituciones que, aunque artificiales, pueden ser denominadas como auténticamente morales (y en ese restringido sentido, «naturales»), por cuanto son necesarias para la ineludible vida social de los seres humanos.

Así que, en conjunto, llevaría razón Annette Baier,⁴² cuando niega que sea el individuo moral quien se traiciona a sí mismo. Ser moral no es necesariamente una forma de locura o de irracionalidad completa.

«...optar por las modestas ganancias materiales junto con la conciencia de que los medios usados para conseguirlas soportarían la inspección pública, de modo que uno no tendría ninguna necesidad de ocultamiento, con preferencia sobre la precaria y sometida a engaño satisfacción del próspero bribón, no parece una preferencia irracional.»⁴³

41 Aquí habría que decir que le falta a Hume en general una cierta conciencia de la necesidad de instituir límites también en este ámbito por medio de la creación de instituciones de derecho internacional, aunque es cierto que cuando habla de la agresión extrasistemática habla de una agresión sin reglas por lo que da a entender que podría haber alguna clase de agresión restringida por alguna clase de reglas. Cfr. El tratamiento de este tema en Hume: EPM, 3.1.11.

42 Baier, «Artificial Virtues and the Equally Sensible Non-Knaves: A Response to Gauthier»: 430.

43 Ibid.: 431.

Llegados a este punto, y acertadamente a mi juicio, Annete Baier considera que el problema debe contemplarse a la luz de toda la teoría de la Justicia de Hume y no sólo de una de sus partes.⁴⁴

Así pues, aclarar el problema del SK, exige «reconstruir» la teoría de la Justicia de Hume, teniendo en cuenta por ejemplo su teoría de la motivación, pero no quedándonos encerrados exclusivamente en ella, que es precisamente lo que David Gauthier pretendía conseguir con su análisis del problema humeano del SK; habiendo hallado un aparente déficit motivacional en el análisis humeano de la justicia, ello invalidaría toda la teoría, en especial su fundamentación en la utilidad.

Pero, el supuesto déficit motivacional es sólo aparente a pesar de alguna vacilación y hasta circularidad en Hume. La clave para demostrar esto se halla en la teoría de la redirección de los motivos naturales (válidos en el ámbito próximo), por medio de instituciones convencionales y/o artificiales, que constituyen para el ser humano una segunda naturaleza y que se justifican, como todas las reglas o normas, por su funcionamiento como sistemas que superan las tendencias de los seres humanos, en el nivel «natural» o más primario de interacción, a concentrarse en lo inmediato, en lo próximo, en el corto plazo.

La superación de esta tendencia debería llevarnos a ver la moral, en una especie de anticipo de las tesis cognitivistas de gran parte de la psicología moral del siglo XX, como la capacidad, destreza, o virtud al fin, que posibilitaría la cooperación a gran escala, siempre que vaya cambiando su estructura motivacional a medida que va aumentando el volumen y complejidad de esta cooperación.

El SK sería, así, el individuo sometido a una especie de *regresión infantil* a los motivos naturales de los primeros estadios de su comportamiento moral.

Esa es la razón de que Hume, en un determinado momento, lejos de apelar a la compleja argumentación utilitaria (aunque no completamente utilitarista) que da soporte a los motivos morales en el estadio complejo y avanzado de la interacción humana, la vida social, apele a un simple argumento naturalista o cuasi-intuicionista.

Si no comprendes las razones que apoyan el deber de ser moral, la clave de la articulación y compatibilidad entre racionalidad y moralidad, al menos deberías retrotraerte a lo que podríamos llamar la *programación moral natural* de los seres humanos en el entorno inmediato, esto es, si el SK no es capaz de sentir, por falta de desarrollo bajo la apariencia de exceso de racionalidad, el motivo artificial que crea una obligación con la Justicia, deberemos recordarle el motivo natural para la moral, que reside en nuestra sociabilidad, esto es, en la unión que existe en el ámbito social próximo entre interés propio y ajeno.

Ahí, nuestro comportamiento no necesita de refuerzos artificiales y el SK deja de ser «inteligente» y «racional» y aparece como lo que es, un «knave», un «sinvergüenza».

Así pues, la teoría de Hume acerca del desarrollo progresivo de la cooperación desde esquema binarios de cooperación digamos natural, en la que el motivo para dicha cooperación es la propia pasión natural, por ejemplo, de amor por los hijos, esto es, la tendencia al altruismo co-sanguíneo o de generosidad para quienes comparten nuestros mismos genes

44 En concreto, propone una muy interesante interpretación a la luz de una original teoría de la evolución de las instituciones como consecuencia de una redirección de los motivos naturales de comportamiento en cuatro niveles de complejidad progresiva.

(que por tanto puede ser explicada pero al tener un motivo natural no necesita ser fundamentada en ninguna otra razón), y de cooperación de grupo con una estructura artificial con los extraños o no emparentados genéticamente con nosotros, y que por tanto necesita ser conectada indirectamente con una pasión natural (o retrotraída a los esquemas de cooperación binaria o elemental de tipo individual) o bien generar lo que podríamos llamar también una «pasión artificial», y el continuo juego entre estos dos niveles de comportamiento (uno necesitado de explicación y el otro de justificación o fundamentación), constituye sin lugar a dudas uno de los grandes logros de la filosofía moral y política de Hume. Lo natural en Hume es la base de lo artificial y el artificio no se puede armar si no es partiendo de lo natural. Sólo desde ahí, también, puede interpretarse adecuadamente la paradoja profunda que propone el «SK passage» en torno al juego de las distintas estructuras de comportamiento de la «human nature». ⁴⁵

Desde este punto de vista, la figura del SK no sólo no constituye una objeción insuperable para Hume, sino que representa una ilustración de la riqueza de la compleja teoría de Hume sobre el paso de lo natural a lo artificial en el marco del desarrollo de las instituciones sociales y políticas, que recordemos, no obstante, están basadas en una moralidad natural basada en el principio del autointerés y de la generosidad parcialmente inclinada hacia los próximos a nosotros.

5. Consideraciones finales

En todo caso, este trabajo forma parte, en realidad, de un proyecto mucho más amplio del que no hemos desarrollado todos sus elementos o implicaciones. Es así una pequeña, pero creo que importante, batalla de un conflicto general de un gran interés teórico, y práctico, creo. Tiene que ver con nuestra manera de entender la historia de las ideas éticas y políticas, por lo que es interesante para los filósofos e historiadores de las ideas; pero, tiene también que ver con el desarrollo de soluciones al problema de la cooperación entre los seres humanos y, por tanto, es de interés para el simple agente moral. He intentado dar una respuesta al problema de por qué ser moral con un sujeto que no lo es puntualmente, pero que puede volver a serlo, o simplemente serlo, si se establecen los incentivos negativos adecuados. La idea humeana acerca del desarrollo de reglas del juego con una naturaleza institucional que vuelvan poco rentable no respetarlas evita la caída del sujeto moral ordinario en la figura poco atractiva, y nada práctica, del «idiota moral» o «moral fool» que coopera, coopera y coopera...mientras que los otros desertan y se aprovechan de él y de unas reglas de juego mal definidas. La institucionalización del juego y de sus reglas, así como la publicidad de éstas y la difícil creación, gradual, evolutiva y compleja de un «conocimiento compartido» entre quienes interactúan y quienes observan la interacción moral entre los agentes, aseguran, en un cierto sentido y en una cierta medida, que todos podamos seguir siendo morales sin traicionar nuestra racionalidad.

45 La «naturalidad» de lo artificial, tratándose del ser humano, y la importancia de esta desambiguación del término «natural» en Hume fue una de las primeras influencias de la lectura de Esperanza Guisán en mi trabajo y orientó decisivamente mi interpretación de la ética de David Hume. Ésta idea está desarrollada principalmente en Esperanza Guisán Seijas, *Como Ser Un Buen Empirista En Ética, Monografías Da Universidade De Santiago De Compostela* (Santiago de Compostela: Servicio de Publicaciones de la Universidad, 1985).

He partido, pues, de la idea de que David Hume en el fragmento sobre el SK plantea un problema crucial y le da una posible solución.

En todo caso, resulta ciertamente sorprendente la concentración de algunos autores contemporáneos en el fragmento del SK, olvidando, que Hume habla en otros lugares de la «knavery» y de los peligros que ésta supone para la Justicia, así como que, precisamente, este fragmento no lo es, sino que forma parte de una discusión mucho más amplia y de un conflicto mucho más general, que es del surgimiento y mantenimiento de la Justicia.

La razón que a mi modo de ver explica la sobre-discusión que se ha llevado a cabo de este problema reside en que diversas formas de contractualismo, principalmente de inspiración kantiana y hobbesiana, han visto este problema planteado por Hume como una amenaza directa para la línea de flotación de la posible teoría política surgida del utilitarismo clásico, que hundiría sus raíces, a pesar de las reticencias principalmente de John Stuart Mill,⁴⁶ en la teoría de la justicia humeana tal y como fue reformulada en la *Investigación sobre los principios de la moral*.

En esa disputa general de diversos contractualismos con lo que Brian Barry llamó la «teoría de la justicia como ventaja mutua»⁴⁷ se encierra, de un modo muy intenso y relevante, una lucha por el control interpretativo de las raíces teóricas de ambos modelos.

Esto nada tiene que ver en principio con el planteamiento o solución de problemas sino con lo que podríamos denominar «apropiación simbólica de las fuentes», esto es, con el uso simbólico de la historia del pensamiento para apoyar nuestras tesis. Es, pues, en principio y como he señalado más arriba, una disputa historiográfica.

En mi opinión, este tipo de ejercicio de apropiación simbólica es lo que rige las intenciones, por ejemplo, de John Rawls en su análisis, extenso y detallado, pero nada inocente ni neutral, de las ideas de David Hume tanto en sus *Lectures on the History of Moral Philosophy* como en sus *Lectures on the History of Political Philosophy*.⁴⁸ Asimismo, y de un modo todavía más claro, es la intención que a mi modo de ver anima los diversos trabajos en los que David Gauthier intenta mostrar que las ideas de Hume se acercan mucho a las suyas y al contractualismo hobbesiano.

Tanto uno como otro en realidad abordan y persiguen de formas diversas un mismo objetivo: *alejar a David Hume del utilitarismo*.

En este trabajo me he limitado a discutir y analizar el problema del «bribón inteligente» tal y como lo plantea Hume, he revisado los argumentos principalmente de David Gauthier sobre la teoría humeana acerca de tal problema, y he sostenido lo que paradójicamente creo que es la solución más simple a la posible contradicción del propio Hume consigo mismo y con toda su teoría de la Justicia: dicho problema se supera del modo más eficaz por medio de una visión utilitarista de la regla de la teoría ética y política humeana.

46 Cfr. Por ejemplo la negativa actitud de Mill frente a la figura de Hume, en contraste con la admiración que siempre mostró por Hume Jeremy Bentham, John Stuart Mill, «Bentham,» en *Essays on Ethics, Religion and Society (Collected Works of John Stuart Mill, X)*, ed. D. P. Dryer, F. E. L. Priestley, and John M. Robson (Toronto/London: University of Toronto Press; Routledge, 1969): 80.

47 Brian M. Barry, *Teorías De La Justicia* (Barcelona: Gedisa, 1995).

48 John Rawls, *Lectures on the History of Moral Philosophy*, ed. Barbara Herman (Cambridge, Mass: Harvard University Press, 2000), John Rawls, *Lectures on the History of Political Philosophy*, ed. Samuel Richard Freeman (Cambridge, Mass.: Belknap Press of Harvard University Press, 2007).

Añadiré, en mi contra, que esto no convierte automáticamente a Hume en un utilitarista sin más, pero eso es otro problema, del que me he ocupado en otros lugares y del que me ocuparé en otro momento. Es una guerra que ahora no he intentado ganar. Me he conformado con disputar esta pequeña batalla en torno al problema del SK.

Queda pendiente, sin duda, la tarea de abordar, por ejemplo, el análisis en profundidad de la presencia de las ideas de Hume en Rawls o el tratamiento general de hasta qué punto y en qué medida sostiene Hume alguna clase de posición utilitarista. Sentado esto, me he concentrado simplemente en el análisis del problema del SK tal y como lo plantea Hume. Los resultados de mi análisis creo que son contrarios a la interpretación estándar de este problema al llegar a la conclusión de que Hume piensa que dicho problema es soluble en los términos de su propia teoría.

Para acabar, señalaré lo que a mi modo de ver es un resultado no previsto de la interpretación utilitarista de la regla de la teoría de la Justicia de Hume que, a mi modo de ver permite solucionar el problema, que no dilema, del SK.

6. Una Conclusión paradójica. El «Sensible Knave» y el «liberalismo» de Hume

Junto a lo que espero sea alguna aportación teórica relevante, al menos desde el punto de vista de la reevaluación del papel de Hume en la Historia o Invención del Utilitarismo, habrá un aspecto, nada desdeñable para mí, que soy muy aficionado al humor filosófico, derivado de este trabajo que quizás no fuera la menos útil de sus conclusiones.

Si la posibilidad de contradicción entre el interés moral (global, normativo) y el puro interés individual (particular, relativo a un acto específico y concreto) sólo se puede superar mediante el desarrollo de instituciones digamos «arbitrales»⁴⁹ que permitan que ese conflicto no se resuelva sino que en realidad se disuelva por cuanto no vaya de ningún modo en interés del «bribón» ser eso, un «bribón», al menos si quiere seguir siendo «inteligente», en realidad, estaríamos concluyendo que no parece sostenibles de un modo global las interpretaciones «liberales» del pensamiento socio-político de David Hume, no al menos aquellas que consideran que ese liberalismo llevaría a una pura defensa de un combate abierto entre individuos en lucha por su propio interés y sin marcos institucionales estables y con posibilidad de intervención, digamos que «desde arriba» o meta-individualmente. Yendo más allá, esto parecería querer decir que *es imposible ser utilitarista de la regla y ser completamente liberal* en este sentido abierto y ampliamente desregulador y anti-institucional.

En realidad, la contradicción desaparece si se retrotrae uno a lo que el liberalismo clásico propuso hasta al menos Jeremy Bentham: *un primado de la defensa de los derechos civiles individuales en el marco de un fuerte contexto de instituciones posibilitantes y garantes de esos mismos derechos*, es decir, una conceptualización de la libertad que no escindía ni separaba radicalmente las dimensiones negativa y positiva de ésta.

Que el liberalismo posterior haya hecho desaparecer la importancia de los árbitros institucionales, o satanizado la propia existencia e importancia de ciertas instituciones

49 La influencia en mi posición de las teorías de Julia Barragán sobre el papel moral de las instituciones y el diseño institucional son muy relevantes. Tal y como se argumenta brillantemente por ejemplo en Julia Barragán, «El Mensaje De Las Instituciones,» *Telos: Revista Iberoamericana de Estudios Utilitaristas* 11, no. 1 (2002).

como necesariamente contrarias a la libertad, entendida únicamente como una ausencia de interferencias, no es responsabilidad de estos primeros liberales de la tradición utilitarista, quienes parecían creer posible volver compatible ser utilitarista y liberal, así como seguir aceptando la importancia de algunas «instituciones» y hasta de árbitros con una función de legítima interferencia cuando en el desarrollo del juego social, político o económico algún actor decide comportarse como un «ladrón inteligente», esto es, como alguien racional, autointeresado e informado, pero al fin y al cabo, un «ladrón».

La teoría utilitarista de la regla elaborada por David Hume en su teoría de la Justicia es la mejor explicación de las condiciones que dan lugar a este problema y creo, también, que una de las más elaboradas fundamentaciones filosóficas contemporáneas de los mecanismos para controlar y solventar este «ardid de la razón» por unos medios anclados en la propia naturaleza racional y pasional de los seres humanos. Sólo por eso merecería David Hume un lugar preeminente en la historia de la *invención del Utilitarismo*, a pesar de que ello parezca alejarlo irremediabilmente de algunas formas muy recientes de *interpretación del liberalismo*.

Bibliografía

- Baier, Anette C. «Hume's Account of Social Artifice --Its Origins and Originality.» *Ethics* 98, no. July (1998): 757-78.
- Baier, Anette. «Artificial Virtues and the Equally Sensible Non-Knaves: A Response to Gauthier.» *Hume Studies* 18 (1992): 429-39.
- Baier, Anette. *A Progress of Sentiments. Reflections on Hume's Treatise*. Cambridge, Mass: Harvard University Press, 1991.
- Baillie, James. *Hume on Morality, Routledge Philosophy Guidebooks*. London; New York: Routledge, 2000.
- Baron, Marcia. «Hume's Noble Lie: An Account of His Artificial Virtues.» *Canadian Journal of Philosophy* 12, no. September (1982): 539-55.
- Barragán, Julia. «El Mensaje De Las Instituciones.» *Telos: Revista Iberoamerica de Estudios Utilitaristas* 11, no. 1 (2002): 123-44.
- Barry, Brian M. *Teorías De La Justicia*. Barcelona: Gedisa, 1995.
- Brown, Charlotte. «Is Hume an Internalist?» *Journal of the History of Philosophy* 26 (1988): 69-87.
- Cohon, Rachel. *Hume: Moral and Political Philosophy, The International Library of Critical Essays in the History of Philosophy*. Aldershot (England): Ashgate : Dartmouth, 2001.
- Darwall, Stephen. «Motive and Obligation in Hume's Ethics.» *Nôus* 27 (1993): 415-48.
- Gauthier, David. «Artificial Virtues and the Sensible Knave.» *Hume Studies* 18, no. 2 (1992): 401-27.
- Gauthier, David. «David Hume, Contractarian.» *Philosophical Review* 88 (1979): 3-38.
- Gauthier, David. «Three against Justice: The Foole, the Sensible Knave, and the Lydian Shepherd.» En *Midwest Studies in Philosophy* vii, 11-29. Minneapolis: University of Minnesota Press, 1982.
- Guisán Seijas, Esperanza. *Como Ser Un Buen Empirista En Ética, Monografías Da Universidade De Santiago De Compostela*. Santiago de Compostela: Servicio de Publicaciones de la Universidad, 1985.

- Haakonssen, Knud. *The Science of a Legislator: The Natural Jurisprudence of David Hume and Adam Smith*. Cambridge ; New York: Cambridge University Press, 1981.
- Hausman, Daniel- M. «Rationality and Knavery.» *Vienna Circle Institute Yearbook* (1997): 67-79.
- Hume, David. *Disertación Sobre Las Pasiones Y Otros Ensayos Morales*. Editado por José Luis Tasset Carmona. 2a ed (1ª ed.: 1990). Barcelona: Anthropos, 2004.
- Hume, David. *An Enquiry Concerning the Principles of Morals: A Critical Edition*. Edited by L. Beauchamp Tom, *The Clarendon Edition of the Works of David Hume*. Oxford: Clarendon Press : Oxford University Press, 1998.
- Hume, David. *Ensayos Políticos*. Trad. de César Armando Gómez. Editado por Josep María Colomer. Madrid: Tecnos, 1987.
- Hume, David. *Ensayos Políticos*. Editado por Enrique Tierno Galván, *Civitas / Instituto De Estudios Políticos*. Madrid: Instituto de Estudios Políticos, 1955.
- Hume, David. *Investigación Sobre Los Principios De La Moral*. Editado por Gerardo López Sastre. Madrid: Espasa Calpe, 1991.
- Hume, David. «Tratado De La Naturaleza Humana (Trad. De Vicente Viqueira, Revisada Y Actualizada).» En *David Hume: Obra Filosófica*, Editado por José L. Tasset y Raquel Díaz Seijas. Madrid: Gredos, (en prensa).
- Hume, David. *A Treatise of Human Nature*. Edited by David Fate Norton and Mary J Norton, *Oxford Philosophical Texts*. Oxford, New York: Oxford University Press, 2000.
- Magri, Tito. «Natural Obligation and Normative Motivation in Hume's Treatise.» *Hume Studies* 22 (1996): 231-53.
- Matravers, Matt. 2002. Hume on Trust. En *BBC Reith Lectures: «The Philosophy of Trust»*, http://www.open2.net/trust/society/soc_hume/soc_hume1.htm
- Mill, John Stuart. «Bentham.» En *Essays on Ethics, Religion and Society (Collected Works of John Stuart Mill, X)*, edited by D. P. Dryer, F. E. L. Priestley and John M. Robson, CXXXIX, 577 p. Toronto/ London: University of Toronto Press; Routledge, 1969.
- Postema, Gerald- J. «Hume's Reply to the Sensible Knave.» *History of Philosophy Quarterly* (1988): 23-40.
- Radcliffe, Elizabeth. «How Does the Humean Sense of Duty Motivate?» *Journal of the History of Philosophy* 34 (1996): 383-407.
- Rawls, John. *Lecciones Sobre La Historia De La Filosofía Política*. Trad. De Albino Santos Mosquera. Editado por Samuel Richard Freeman. Barcelona [etc.]: Paidós, 2009.
- Rawls, John. *Lectures on the History of Moral Philosophy*. Edited by Barbara Herman. Cambridge, Mass: Harvard University Press, 2000.
- Rawls, John. *Lectures on the History of Political Philosophy*. Edited by Samuel Richard Freeman. Cambridge, Mass.: Belknap Press of Harvard University Press, 2007.
- Tasset, José L. *La Ética Y Las Pasiones (Un Estudio De La Filosofía Moral Y Política De David Hume)*. A Coruña:Tasset, José L. «La Ética Y Las Pasiones: Una Propuesta Neohumeana.» En *Los Sentimientos Morales [Cuaderno Grís, Época III (No. 7)]*, editado por Pura Sánchez Zamorano, 203-20. Madrid: Universidad Autónoma de Madrid, 2003.
- Tasset, José L. «La Ética, Las Pasiones Y La Razón. En Defensa Del Gobierno Ético De La Razón.» En *Decidir En Sociedad. Homenaje a Julia Barragán*, editado por Fernando Aguiar, Francisco Lara y Nelson Lara, 23-48. Caracas: Ediciones Chyrimé-k, C.A., 2009.

