

Consecuencialismo e imparcialidad

DAMIÁN SALCEDO
(Universidad de Granada)

Introducción

Hacia finales de los años 70 y principios de los 80 se presentaron las versiones mejor argumentadas del utilitarismo y, sin embargo, al mismo tiempo se desencadenó un proceso de revisión de los supuestos de dichas versiones dentro de la propia tradición. El motor de esa revisión radicaba en una noción de personalidad moral que exigía el reconocimiento de la complejidad del comportamiento humano¹ y que colisionaba radicalmente con la idea de agente moral que surgía de las propuestas utilitaristas. Estas, no obstante, defendieron su posición apelando a la coherencia de una propuesta global del comportamiento en la que exigencias de racionalidad y exigencias normativas básicas se entrelazaban férreamente. Desde su punto de vista, aquellos que querían ir más allá de esta propuesta aparecían como defensores del prejuicio y la intuición.

En los últimos años han aparecido propuestas de teorías éticas que, sin abandonar el componente consecuencialista básico del utilitarismo, tratan de incorporar la noción de agente moral que dio origen a la revisión. Lo que podemos llamar el neoconsecuencialismo es, de este modo, una familia de propuestas aún en un estado no muy desarrollado que tienen en común el haber abandonado algunos de los elementos del utilitarismo —los que llevan directamente a la noción criticada de agente moral—, pero que mantienen el valor de las consecuencias como criterio irrenunciable de la decisión moral.

En este artículo intentaré mostrar por qué el utilitarismo considera que tales vías son incorrectas y por qué califica de prejuicios lo que para muchos de nosotros son elementos indudables de la evaluación moral. Creo que en el fondo de la polémica se encuentra el problema tradicional de las razones para la acción, puesto que es la teoría que da respuesta

1. Como se verá, tengo en mente sobre todo las críticas al utilitarismo de J. Rawls (1971) y (1982), B. Williams (1973), D. A. J. Richards (1971) y (1988), St. Hampshire (1978) y Th. Nagel (1979).

a este problema la que le permite sustentar la racionalidad de la adopción de los principios utilitaristas frente a otros posibles. Por este motivo lo que el neoconsecuencialismo tiene que mostrar es la incorrección de la teoría de las razones para la acción del utilitarismo y la posibilidad de una explicación alternativa que le siga permitiendo la inclusión de las consecuencias entre los valores prioritarios a considerar. Por ello, también me propongo dejar indicada una vía para realizar este proyecto.

La racionalidad utilitarista

Para un utilitarista típico como J. C. Harsanyi, el principio de utilidad media es el único principio por el que puede guiarse la conducta ética racional, puesto que es el único principio capaz de satisfacer las características básicas de la explicación de cualquier otra conducta racional y así integrarse en una teoría general del comportamiento racional.

La noción de racionalidad que defiende el utilitarismo es equivalente a la noción de consistencia racional. En las posiciones utilitaristas que aquí analizamos dicha noción es tomada de la concepción básica que utiliza la ciencia económica para realizar sus explicaciones y predicciones. En el análisis económico el comportamiento racional de un individuo depende de la relación entre sus elecciones sobre fines alternativos y el conjunto de sus preferencias y oportunidades. Elegir racionalmente significa que se da una relación consistente entre su elección y el ordenamiento de sus preferencias. Dicha relación es consistente cuando la elección lo es de la alternativa que es preferida en primer lugar².

La noción de racionalidad como consistencia permite ir definiendo los diferentes principios de comportamiento racional para las diferentes situaciones de decisión desde un simple principio de maximización de la utilidad para situaciones de decisión individual bajo certidumbre hasta los diversos principios requeridos por las situaciones en que la decisión depende de la interacción entre diversos individuos propias de la teoría de juegos. En cada caso, lo único que se requiere es que se especifiquen las condiciones formales que han de satisfacer los ordenamientos de preferencias a efectos de su adecuación a las características de la situación decisional³.

2. Lo atractivo de este modelo es que supera la noción común de comportamiento racional que lo circunscribe a la consecución de los medios más adecuados para fines dados. En el modelo económico la racionalidad de la elección depende de la satisfacción de criterios prefijados en las preferencias y la racionalidad entendida en términos de medios-fines puede aparecer como un caso especial de esta concepción: el caso en el que la elección de un objeto satisface la condición de ser causalmente eficaz para producir un fin dado; cfr. HARSANYI, J. C.: (1976a).

3. Cuando el ordenamiento de preferencias de un individuo satisface ciertos axiomas de consistencia y continuidad, entonces las preferencias pueden ser expresadas numéricamente en términos de funciones de utilidad. Pero el criterio de racionalidad

En esta concepción, las decisiones éticas se caracterizan por producirse en situaciones en que la alternativa que ha de escogerse afecta a un colectivo de individuos que tienen intereses comunes, aunque divergentes⁴. Cuando un decisor se enfrenta a este tipo de situaciones ha de construir sus preferencias de forma que satisfagan un conjunto especial de postulados, de forma que las preferencias éticas sean aquellas que un individuo tendría «sólo en aquellos posiblemente raros momentos cuando se fuerze a sí mismo a adoptar una actitud imparcial e impersonal» (J. C. Harsanyi (1955), p. 14). Pero la decisión ética sólo podrá considerarse racional cuando existe una relación de consistencia entre las preferencias éticas del individuo y su elección.

Apelando a esta noción de racionalidad como consistencia, J. C. Harsanyi ha defendido que el criterio de decisión racional para situaciones éticas es el principio de utilidad media. O, de otro modo, que la elección racional en situaciones donde la decisión afecta a un colectivo es la elección de aquella alternativa a la que corresponda la función de utilidad social más alta. Del mismo modo que la racionalidad individual depende de la consistencia entre las preferencias personales y sus elecciones, la racionalidad ética depende de la consistencia entre las preferencias éticas de un individuo y su elección final.

En esta explicación del comportamiento ético racional el elemento esencial es esa noción de racionalidad como consistencia entre elección y preferencias. Pero esta noción es dependiente de una teoría especial de lo

sigue siendo el mismo: los individuos eligen según sus preferencias y eligen aquello que prefieren más. En esta noción la apelación a la utilidad no implica que sólo sea racional elegir egoístamente. La autonomía de las preferencias es respetada escrupulosamente. De forma que el individuo puede tener preferencias egoístas, altruistas o éticas. La utilidad no es más que una medida de la preferencia y, por consiguiente, no es la marca del egoísmo. Por otro lado, no es tampoco una norma o criterio con el que el individuo discrimine entre sus preferencias. Por el contrario, la utilidad depende de las preferencias y simplemente les asigna una relación dentro de una escala. Así, cuando se dice que el criterio de racionalidad de la conducta es la maximización de la utilidad simplemente se está diciendo que la decisión racional es aquella que se incline por la alternativa que está más alta en la escala de utilidad del individuo en cuestión: cfr. HARSANYI, J. C.: (1976a) y (1977); también puede consultarse la útil colección de artículos de GÄRDENFORS, P.; N. E. SAHLIN (eds.) (1988).

4. Esto lleva a hablar a Harsanyi de un «compromiso equitativo» inicial entre los individuos (HARSANYI, J. C.: 1955, p. 13) por el que deciden que las situaciones sociales se evalúen en términos de las preferencias de todos. Ciertos autores han interpretado dicha expresión como la exigencia de un procedimiento contractual y a hablar de la propuesta ética de Harsanyi como de un «contrato utilitarista» (MUELLER, D. C. (1979)). Esto es, sin embargo, erróneo puesto que la noción filosófica de contrato no solamente es ajena al utilitarismo, sino además un recurso del todo innecesario para su propuesta; cfr. DWORKIN, R. (1973). Como aquí queda expuesto, la mención a un compromiso inicial no es más que la descripción de las situaciones en las que se producen las decisiones éticas.

que se considera actuar por razones. Esta teoría señala a las preferencias como lo único que puede constituir una razón de esa clase. La noción de preferencia se utiliza como un signo para cubrir cualquier actividad volitiva; cualquier cosa que una persona quiera cuenta como una preferencia, no importa que sea simplemente un gusto o que sea un deseo o algo que se quiera porque se tenga un compromiso o un ideal o una obligación. Lo importante es que, cuando alguien quiere algo, la única actitud racional es intentar conseguirlo; o en nuestro caso, cuando alguien prefiriera algo, su única decisión racional es elegirlo.

Al llevar esta concepción al campo del comportamiento moral obtenemos la mencionada idea de que la única decisión moral que también es racional consiste en la elección de aquella alternativa que está ordenada más alta en nuestro ordenamiento ético de preferencias. ¿En qué consiste un ordenamiento ético de preferencias? La preferencia ética de un individuo es un tipo especial de preferencia que satisface un conjunto de especificaciones de construcción. Se exige de ella que tenga en cuenta las preferencias personales sobre el conjunto de alternativas de cada uno de los individuos afectados por la decisión. Por ello su forma es la de un agregado de preferencias personales. Se exige además que estas preferencias personales y las preferencias éticas cumplan una serie de requisitos de consistencia racional que permitan al decisor darles la forma matemática de funciones de utilidad. Se exige asimismo que esas funciones de utilidad sean construidas por éste bajo un principio de equiprobabilidad, de manera que sean comparables entre sí. Este principio de equiprobabilidad lleva aparejada una exigencia de imparcialidad que impide que las preferencias de un individuo reciban una ponderación mayor o menor que las de otro. Pues bien, cuando todas estas exigencias se ven satisfechas, la preferencia ética que resulta es una preferencia imparcial que tiene la forma de una función de bienestar social equivalente a la adición equiponderada de las funciones de utilidad personales de todos los individuos afectados⁵.

5. En la axiomatización conducente a justificar la adopción del utilitarismo se muestra claramente cómo actúan conjuntamente todas estas exigencias. De los tres axiomas necesarios, los dos primeros sólo exigen que las preferencias satisfagan los principios de la teoría bayesiana de la decisión a efectos de que puedan ser construidas como funciones de utilidad. Estas exigencias son importantes en cuanto que conllevan hipótesis controvertidas sobre la naturaleza humana, en particular sobre su actitud ante el riesgo (cfr. la agria discusión entre J. Rawls (1971) y J. C. Harsanyi (1975a)). Pero, con todo, es aún más significativo el tercer axioma, llamado «individualista» —un principio de la familia paretiana—, que impide que cualquier otra cosa que no sean las preferencias personales sean tenidas en cuenta como material del juicio ético. Harsanyi justifica su adopción en base a que admitir otra información es querer admitir valores que la sociedad puede no compartir y que harían de tal juicio algo arbitrario y subjetivo (HARSANYI, J. C. (1955)). Esta restricción de la información no sería tan negativa si no se combinara con la actitud que el modelo requiere que se tenga

Vemos aquí entrelazados de forma férrea supuestos de racionalidad y supuestos éticos que, finalmente, llevan a determinar como el único principio ético racional el principio de utilidad media. El utilitarismo defiende la adopción de los supuestos de racionalidad por ser principios aceptados en la teoría de la decisión racional bayesiana y defiende la adopción de los supuestos éticos por ser los más naturales, débiles e incontrovertibles compatibles con los primeros. Pero todos ellos juntos exigen más de lo que es natural aceptar. Exigen que el decisor considere las acciones alternativas valiosas únicamente por sus consecuencias; que estas consecuencias sólo sean registrables en términos de preferencias-funciones de utilidad; y que, por último, el agregado de estas preferencias sólo pueda ser una función lineal aditiva de las mismas. Si cada una de estas especificaciones es discutible aisladamente, cuando se considera que su conjunto implica o requiere que el decisor moral evalúe las acciones como si nadie —él mismo u otra persona— fuera el agente que origine lo que vaya a suceder y, por consiguiente, como si no tuviera responsabilidad sobre sus acciones, entonces la aceptación de las mismas se vuelve más controvertida. Como se suele reprochar, el utilitarismo sólo evalúa las acciones por los estados que resultarán de su acaecer en el mundo, no dando ningún significado moral a quién las ejecute o de qué manera han llegado a producirse. Para muchos de nosotros lo primero es importante para la comprensión de la naturaleza del juicio moral y lo segundo es peculiarmente importante en los juicios y decisiones referentes al ámbito de la justicia social ⁶.

Así las cosas, la teoría de las razones para la acción que sostiene que sólo las preferencias —entendidas del modo mencionado— pueden constituir tales razones, cuando es trasladada al terreno del comportamiento moral, hace desaparecer la idea más básica de agente moral. La idea que liga indisolublemente la acción a los objetivos, ideales o plan de vida de

ante estas preferencias. El postulado de equiprobabilidad requiere que el decisor construya las funciones de utilidad personales como si tuviera una igual probabilidad de estar en la piel de cada uno de los afectados en las distintas situaciones sociales que se seguirían de la adopción de cada una de las acciones alternativas; i. e., requiere que el decisor se comporte como un observador imparcial y simpatético (Harsanyi llama en su apoyo aquí a la influyente tradición que arranca de A. Smith en defensa de la teoría del observador ideal y a Kant a través de Hare como defensor de un principio de universalizabilidad; cfr. HARSANYI, J. C. (1977)). La imparcialidad ante las preferencias como definición del punto de vista moral es equivalente a la consideración de que las únicas razones que pueden recibir el calificativo de morales son las «no referidas al agente», como veremos más adelante.

6. En este artículo me centraré en la primera de estas cuestiones; sobre los juicios de justicia social ya he discutido un modelo en el que las exigencias utilitaristas son reducidas hasta el punto de hacer compatibles el consecuencialismo y un tratamiento moral adecuado a las necesidades que viene exigido por la noción de agente moral que aquí también utilizamos; cfr. SALCEDO, D. (1987a).

un agente. El decisor moral del utilitarismo no es más que un punto de vista neutral que recoge las preferencias de sí mismo y de los demás, manifestando finalmente una preferencia por un estado en lugar de por otro como si él mismo y los demás no tuvieran una relación moralmente significativa con dicho estado. De este modo, las razones morales para la acción son siempre única y exclusivamente neutrales con respecto al agente*.

Razones morales

Ahora bien, ¿sólo es posible reconocer como razones morales las razones neutrales? Si nos atenemos a los trabajos de Th. Nagel, esto no es así. El admite al menos cinco tipos de razones en lo que denomina el «territorio de la moralidad burguesa irreflexiva» ([1986], p. 166), diferenciados por sus fuentes, y no todos ellos son neutrales. Las que podrían llamarse razones de utilidad —i. e., «cómo afectan los actos de alguien a todo el mundo» (Th. Nagel [1979], p. 204)— y las razones que dan los valores perfeccionistas propios de aquellos logros y creaciones que tienen un valor intrínseco (*ibidem*), constituyen razones neutrales para que cualquiera actúe. Pero junto a ellas habría que reconocer algunos tipos de razones que sólo son razones para un individuo particular en una situación particular. Habría que reconocer, así, las razones que dan las obligaciones específicas que se tienen con personas e instituciones, obligaciones que «deben aceptarse ya sea por la aceptación deliberada o por alguna relación especial con la persona o institución en cuestión» (*ibidem*, p. 203). Hay que reconocer asimismo limitaciones a lo que uno puede hacer derivadas de los derechos generales que tiene todo el mundo, «ya sea para hacer ciertas cosas o para no ser tratado de cierta manera» (*ibidem*). Y, finalmente, hay que reconocer también las razones que surgen del compromiso con nuestros propios proyectos y actividades (*ibidem*, p. 205). La diferencia esencial entre estas últimas y las primeras es la apelación consustancial e ineliminable al agente de la acción de las razones referidas al agente. Cualquiera que suceda que sea yo, he de admitir que un acto que aumente la felicidad de los seres sintientes o que preserve la colección de cuadros del Prado es una razón poderosa para actuar. Pero si yo protejo a mis hijos de un daño, o cumplo con los deberes para con mis colegas, o me empeño en comprender problemas filosóficos abstrusos, éstas no son razones que yo reconozca como tales con independencia de que yo sea alguien que es padre, miembro de un colectivo profesional o una persona que ha decidido comprometerse con el estudio de la filoso-

* Traducimos los términos «*agent-relative reasons*» y «*agent-neutral reasons*» por expresiones como «razones referidas al agente» y «razones neutrales con respecto al agente» y variantes de las mismas que tratan de preservar el contenido técnico que en la literatura inglesa tienen.

fia. Son razones para mí en el sentido de que su formulación no puede desligarse de una referencia a quien soy yo y a mi posición particular (Cfr. Th. Nagel [1986], pp. 152-4).

¿Puede el utilitarismo reconocer estas razones que están fuertemente arraigadas en la moral común? La actitud general del utilitarismo —o, quizás mejor, de los utilitarismos— es reticente. Recuérdese que el utilitarismo surgió como una filosofía social radical que quiso borrar las convenciones morales y políticas de una sociedad que apelaba al mandato divino o a la tradición para hacerlas ineludibles. Sosteniendo que sólo aquellas prácticas e instituciones que pasasen el test de la utilidad social —un test racional e imparcial— serían aceptables, descalificaron un entramado de convicciones que aparecían como el producto del dogmatismo y el prejuicio. En razón de que nuestras sociedades occidentales son menos obscurantistas que las que les tocó vivir a los primeros utilitaristas, el moderno utilitarismo ha perdido su retórica virulenta. Pero, con todo sigue alentando esta misión ilustrada y así considera su deber denunciar cualquier nueva forma de prejuicio que se pueda esconder tras las propuestas filosóficas alternativas⁷. Esto es especialmente verdad de los utilitaristas del acto⁸. En la medida en que consideran que nuestras decisiones morales básicas se refieren a los actos individuales en situaciones morales específicas, únicamente aquellos actos que impliquen mayor utilidad social en comparación con otros alternativos pueden ser considerados correctos moralmente. Así les resulta imposible reconocer que alguien tenga una razón moral que no sea neutral con respecto al agente. Si moralmente me está permitido realizar un acto que favorezca la educación de mis hijos, será debido a que ese acto particular tiene mayor utilidad social que cualquier otro alternativo. Pero de ningún modo podrá deberse a que yo tengo una obligación especial como padre para con mis hijos.

Es ya antigua la crítica que trata de mostrar que el seguimiento por parte de una sociedad de una moral utilitarista del acto tendría efectos horribles para la misma⁹. En alguna de sus formas, éste ha sido el principal argumento para defender nuevas versiones del utilitarismo que reconocen algún valor moral a las razones no neutrales en la guía de nuestra conducta.

Lo que diferencia fuertemente al utilitarismo de la regla del utilitarismo del acto es su concepción de la utilidad social. El utilitarismo de la

7. Pueden verse a este respecto las invectivas de HARE, R. M. (1973), contra RAWLS, J.; o las de HARSANYI, J.:(1975b) contra DIAMOND, P. y SEN, A.

8. Su más conspicuo defensor actualmente es SMART, J. J. C. (1956), (1973) y (1978).

9. El origen de esta crítica fue HODGSON, D. H. (1967) y ha sido repetida muchas veces; pero al menos en la radicalidad en que este autor la expuso parece incierta: cfr. SINGER, P. (1972) y MACKIE, J. L. (1973).

regla orienta nuestras decisiones morales básicas a evaluar códigos morales generales: i. e., conjuntos integrados de reglas, prácticas, instituciones, etc., que gobiernan la parcela moral de la conducta de los miembros de una sociedad. Ahora bien, comparar códigos morales generales utilizando como criterio la utilidad social obliga a introducir variables en lo que ha de ser considerado como socialmente útil que no están presentes en la simple comparación de actos. Así señala J. C. Harsanyi que al valorar la utilidad de un código moral, un utilitarista de la regla tendrá que tener en cuenta distintos efectos del mismo sobre la conducta de los individuos —tanto en los directamente implicados en las acciones como en el resto de los individuos. Así, en primer lugar, habría que valorar los efectos de la propia obediencia al código (*compliance effects*), ya sean directos en términos de qué tipos de actos llevarían a cabo los individuos para cumplir las demandas del código, ya sean indirectos como los efectos que la aceptación del mismo tendría sobre los sentimientos de culpa o los sentimientos de aprobación y censura de las personas. En segundo lugar, habría de tenerse en cuenta qué condiciones determinarían esta aceptación y obediencia al código moral para coordinar las estrategias de los individuos en la sociedad (*coordination effects*). Puesto que la utilidad social de un código depende en gran medida de su capacidad para coordinar esfuerzos de una manera socialmente deseable. En tercer lugar, habría que considerar sus efectos sobre las expectativas de los miembros de la sociedad (*expectation effects*), particularmente sobre sus sentimientos de confianza y seguridad en relación a la fiabilidad de que otras personas se guíen por el código en sus actuaciones. Por último, es asimismo importante para determinar la utilidad social de un código moral sus efectos de incentivo (*incentive effects*) sobre los individuos para comprometerse en actividades socialmente beneficiosas¹⁰.

En consecuencia, el utilitarismo de la regla puede preguntarse sobre la utilidad social de, por ejemplo, las obligaciones especiales —y en esa medida reconocerlas como tales— atendiendo a los efectos que tendrían sobre una sociedad en que se siguieran prácticas que no defraudan las expectativas de aquellos que se encuentran en las situaciones en que dichos compromisos se dan. Esto naturalmente significa un paso adelante en la comprensión de las razones referidas al agente. Un utilitarista del acto, por centrarse exclusivamente en la utilidad social de la realización o no de un acto —p. e., el acto implicado por una de esas obligaciones especiales en una situación particular—, no tendría en cuenta los efectos

10. Hemos combinado aquí dos exposiciones diferenciadas de Harsanyi; la primera se encuentra en HARSANYI, J. C. (1978) y la segunda en HARSANYI, J. C. (1988), en la que desarrolla aspectos de BRANDT, R. B. (1970) y (1979), c. 15. Las implicaciones y desarrollos de esta forma de utilitarismo de la regla pueden hallarse en HARSANYI, J. C. (1980), (1982) y (1985).

sobre otras personas, ya fuera porque dichos efectos escapan a su consideración, o bien porque los efectos de un acto particular sobre ellas serían despreciablemente pequeños. Ello le llevaría a ignorar cualquier otra consideración que no fuera la estricta utilidad social de la realización del acto y así la propia existencia de tales obligaciones. Por el contrario, el utilitarista de la regla podría recomendar que las personas actúen por razones referidas al agente en la medida en que están incorporadas en el código moral socialmente más útil y de este modo admitir conductas morales guiadas por tales razones.

Bajo esta perspectiva las razones morales propias de la moral común reciben un papel subsidiario o instrumental. Mas, ¿es suficiente el reconocimiento que hace de las razones referidas al agente el utilitarismo de la regla? El utilitarista de la regla puede admitir que una razón que apele a su posición especial, sus compromisos consigo mismo o los derechos de otras personas es una razón moral si, al mismo tiempo, puede mostrar que es una consideración derivada del código moral general que es socialmente más útil. Por consiguiente, sólo se admiten razones referidas al agente cuando pueden derivarse de razones neutrales, siendo éstas la única fuente de moralidad de aquellas o cualesquiera otras¹¹.

Lo que diferencia la explicación neoconsecuencialista de la utilitarista sobre el valor de las razones referidas al agente es que no considera que tales razones sean morales sólo derivadamente y, por consiguiente, que la aceptabilidad de algo como una razón moral no depende de su neutralidad. El neoconsecuencialismo trata de integrar la pluralidad de los valores morales en una construcción consecuencial del juicio moral, eliminando la distinción entre una moralidad de sentido común en donde la pluralidad se admite y una moralidad «ilustrada» en la que sólo los valores neutrales cuentan. La novedad, pues, de esta teoría estriba en la propuesta de una reinterpretación de la naturaleza de los juicios de valor y de la noción de bondad moral. Y lo que tenemos que explicar, entonces,

11. La teoría de los «dos niveles» de R. M. Hare hace innecesaria la distinción tradicional entre utilitarismo del acto y de la regla; el «arcángel» que tiene por reino el «nivel crítico» del pensamiento moral es un utilitarista del acto muy especial: evalúa el cultivo y seguimiento de las disposiciones morales (reglas) por la utilidad social de la serie de actos de cultivarlas y seguirlas. Pero, para lo que a nosotros nos importa, al igual que en el utilitarismo de la regla, tales reglas especiales propias del «nivel intuitivo» que nos indican lo que debemos hacer son morales últimamente porque pueden ser justificadas en el nivel crítico por razones neutrales de utilidad social. Sin embargo, es de notar que esta teoría de los dos niveles da una explicación más sensible que la del utilitarismo de la regla del puesto especial que las razones referidas al agente tienen en el comportamiento humano al ligarlas a características empíricas de la naturaleza humana; cfr. HARE, R. M. (1976), (1981) y los comentarios de Hare a los artículos de Frankena, Harsanyi, Scanlon y Ursom en D. SEANOR: *N FOTION* (eds.), (1988).

es qué clase de juicio moral es el que puede hacer comprensible sin subordinarlas dicha pluralidad de razones morales¹².

Siguiendo a A. Sen (1982a) podemos distinguir entre «evaluación consecuencial» y «consecuencialismo». Ambos tipos de evaluación tienen por objeto estados de cosas resultantes —ambas son «*outcome moralities*». Lo que las distingue es la distinta descripción que admiten de dichos estados. Mientras que el consecuencialismo típico selecciona como relevantes para el estado de cosas los resultados de una sola fuente de información —ya sean los actos, los motivos, las prácticas o los códigos morales—, la evaluación consecuencial permite que en dicha descripción entre cualesquiera variables que hagan posible definir una situación como mejor en sus consecuencias que otra. Esta consideración no restrictiva de la información moralmente relevante tiene la ventaja de abrir la posibilidad de que cuenten con su peso propio en la determinación final de la bondad de un estado no sólo los efectos de realizar una acción, sino también el carácter moral de la acción misma y los criterios que el agente sostiene como relevantes en la realización de la acción¹³. La exclusión por parte del consecuencialismo típico de esta información parece ahora poco natural, puesto que siempre es parte del estado resultante de una decisión el que una acción sea ejecutada por alguien. Pues bien, con esta información sobre el agente y la acción, entran a formar parte del juicio de valor los valores referidos al agente en pie de igualdad con los valores neutrales de una manera que no le está permitida a otras explicaciones del juicio moral, ya sean consecuencialistas o no.

Ahora bien, admitir que se ha de dar un peso a los valores referidos al agente es imposible para el consecuencialismo tradicional, puesto que ello requiere violar lo que para él es el requisito esencial de un juicio moral: una persona ha de realizar sobre una situación dada el mismo juicio que cualquier otra persona —en igualdad de información—. Los valores referidos al agente son excluidos por esta exigencia que prohíbe que los juicios tengan una referencia a la perspectiva de quien los hace. Por ello el consecuencialismo sólo puede concebir valores neutrales. Pero éste no

12. En este trabajo me centraré en la propuesta neoconsecuencialista de A. Sen porque en ella se muestra mejor que en otras lo que requiere una teoría moral de este tipo. Pero los resultados que se obtendrán creo que son extensibles a las versiones más conocidas del mismo como las de SCANLON, T. (1975) y (1978); SCHEFFLER, S. (1982); SLOTE, M. (1985); PETIT, Ph. (1984) y PETIT, Ph.-BRENNAN, G. (1986).

13. En rigor, también el consecuencialismo —y no sólo la evaluación consecuencial— tiene en cuenta la acción y, de ahí, si bien limitadamente, algunas formas de valores referidos al agente; cfr. SEN, A.: (1982a). Sólo cuando el consecuencialismo es combinado con las exigencias de una moralidad «bienestarista» y aditivista toda consideración de la acción y del agente queda excluida al reducirse drásticamente toda información relevante a la suma de utilidades. Así si dos estados son iguales en su suma de utilidad, serán juzgados como igualmente buenos no importa cuán diferentes sean en sus otras características; cfr. SEN, A. (1979a).

es el caso de la evaluación consecuencial cuando es combinada con una exigencia que consiente la posicionalidad del evaluador. Los juicios de valor relativos posicionalmente son aquellos en los que la relación del evaluador con los estados de cosas evaluadas se ha constituido en una variable decisiva. Así cosas como que el evaluador sea él mismo el agente o no, el que él o los suyos sean beneficiarios del estado o no, el que realice o fruste valores y convicciones, etc., definen la implicación del evaluador con lo evaluado y ello, en suma, es lo que abre la descripción de los estados a una información que es vital para identificar la naturaleza de la implicación que el evaluador tiene en el asunto. Los juicios de valor posicionales no se dejan expresar con fórmulas de una sola variable —« α es bueno»—, sino que requieren la referencia el que emite el juicio —« α es bueno para i »—. Y, de este modo, la noción neutral de bondad moral deja paso a una noción posicional que contiene referencias ineliminables a la persona que realiza el juicio¹⁴.

Aunque hay que observar que los valores relativos al agente dan razones de modo diferente según sea su naturaleza¹⁵, todos ellos tienen en común la posibilidad que abren al evaluador de aducir su diferencia posicional para no verse obligado a valorar una situación como si él no fuera quien es y estuviera implicado en la situación del modo en que lo está. Así del hecho de que yo no tenga razones para impedir que alguien haga algo no se sigue que yo no tenga razones para hacerlo —o que yo deba de obviar estas razones—; o del hecho de que alguien no tenga razones para impedir que yo lo haga no se sigue que yo no las tenga para no hacerlo —o que yo deba obviarlas—. En relación a un conocido caso presentado por Th. Nagel ([1986], p. 176), yo puedo considerar malvado el acto de torcer el brazo al niño para conseguir que su abuela me ayude y así llevar al hospital a los heridos del accidente, tanto si yo mismo no fuera a impe-

14. Ello no impide que algo pueda ser considerado bueno desde cualquier posición; pero esta invarianza interposicional no equivaldría, por tanto, a una no posicionalidad del juicio moral; cfr. SEN, A. (1983), p. 115.

15. A. Sen (1982a) clasifica las distintas razones referidas al agente según impliquen y cuáles de las siguientes formas de relatividad:

1) Relatividad con respecto al que realiza la acción (*Doer Relativity*): No es el caso de que i pueda hacer este acto si, y sólo si, no tiene ninguna obligación de impedir que j lo haga.

2) Relatividad con respecto al que observa la acción (*Viewer Relativity*): No es el caso que i pueda hacer este acto si, y sólo si, j no tiene ninguna obligación de impedir que i lo haga.

3) Relatividad con respecto a la propia evaluación (*Self-evaluation Relativity*): No es el caso que i pueda hacer este acto si, y sólo si, j puede hacerlo.

Así, mientras las restricciones deontológicas requieren de estas tres formas de relatividad, la responsabilidad por la propia integridad sólo requiere DR y VR; cfr. SEN, A. (1982a), s. 5.

dir que otra persona lo hiciera como si esa persona tampoco considerara que debería impedírmelo. Mientras que la valoración del acto en cuanto a sus beneficios resultantes —llevar al hospital a los accidentados— tiene un valor positivo desde un punto de vista neutral, la evaluación del estado resultante que incluye el que yo tuerza el brazo al niño no admite esa neutralidad y requiere que se diferencie si el agente es el mismo que el evaluador. Tanto si finalmente yo decido obviar mi repugnancia moral a torcer el brazo al niño en razón del beneficio para los accidentados como si no, ineludiblemente el juicio hará referencia a las características de mi posición como evaluador marcada por esa repugnancia que yo siento¹⁶.

Al poder prestar atención como parte del estado de cosas resultantes a la realización o frustración de valores referidos al agente, la evaluación posicional está más cerca de nuestros juicios morales reales que cualquiera otra de las presentadas hasta ahora. En los ejemplos utilizados por A. Sen se puede apreciar ésto. Un caso en que es especialmente raro pedir que alguien evalúe una acción independientemente de que él mismo sea su ejecutor o no es el caso de Otelo. Uno no entiende la naturaleza compleja de los juicios de valor si pide que Otelo dé el mismo valor moral a la muerte de Desdémona que cualquier otra persona, puesto que para él tal muerte no puede dejar de ser el asesinato de su esposa que ha perpetrado él mismo. Uno puede estar a favor o en contra del tiranicidio por razones estrictamente neutrales, pero ello no habría de impedir comprender que es posible que un juicio así sea matizado por las obligaciones especiales que recaen sobre el agente. Así, el historiador Apiano manifiesta que el asesinato de César por Bruto y Casio fue un acto peor de lo que hubiera sido si lo hubieran realizado los enemigos de César, puesto que aquellos eran sus amigos y herederos. Sin embargo, hemos de hacerlo notar, para un utilitarista de viejo cuño tales juicios sólo podrían ser disparates, el reflejo de prejuicios o, en el mejor de los casos, producto de una intuición moral no sujeta a la debida inspección racional¹⁷.

16. El realizar una acción por razones referidas al agente que es peor en sus consecuencias que otra acción alternativa provoca la sospecha de que tal apelación puede ser sólo una forma encubierta de autoindulgencia; cfr. WILLIAMS, B. (1976a). En el neoconsecuencialismo se evita este riesgo, puesto que los valores referidos al agente no son absolutos; cfr. SEN, A. (1985), pp. 212 y ss.

17. Pero no sólo para el utilitarista. Muchos autores consideran que la comparación consecuencialista entre estados resultantes cuya descripción incluye el valor de la acción conlleva una comprensión incorrecta de la naturaleza de los valores relativos al agente, dado que lo característico de éstos, opinan, es imponer restricciones a lo que puede ser valorado en términos de buenas consecuencias. Esta objeción, sin embargo, sólo se sostiene si interpretamos el valor de las buenas consecuencias en términos de utilidad-bienestar; pero no cuando se hace en términos de una consideración informativamente plural. Objeciones de ese tipo se hallan en NOZICK, R. (1974) (cfr. la réplica de SEN, A. [1982a]) y FOOT, Ph. (1985) (cfr. la réplica de SCHEFFLER, S. [1985]).

Imparcialidad y razones para la acción

El utilitarismo no puede comprender los juicios que resultan de este consecuencialismo posicional. No puede comprenderlos como juicios morales en sí mismos. La explicación de esta imposibilidad ha de apelar, en mi opinión, a dos fuentes. Por una parte, la doctrina utilitarista sobre la naturaleza de lo que hace moral a un juicio de valor. Por otra, la doctrina que el utilitarismo acepta de lo que constituyen razones para actuar. Aunque cada una de ellas por separado puede llevar por sí misma a la incomprensión de las razones referidas al agente, cuando ambas se dan unidas se refuerzan de tal manera que sólo pueden admitir como morales y racionales los valores neutrales. Expondré cómo sucede tal cosa siguiendo como patrón la teoría de R. M. Hare bajo el supuesto de que los requisitos explicitados en ella son aceptados casi en su totalidad por cualquier utilitarista.

¿Qué hace moral a un juicio de valor? Esta pregunta es interpretada por Hare de la siguiente manera: ¿qué hace moral a una prescripción? Y su respuesta es inequívoca: la universalizabilidad. Ello es así porque una prescripción es la expresión de una preferencia por que algo suceda; pero una preferencia sólo puede ser moral si no es mi preferencia personal. Una preferencia ética es, como dijimos, una preferencia que amalgama todas las preferencias que los afectados por la acción podrían tener sobre ella. La preferencia para ser moral ha de ser impersonal y la impersonalidad requiere que se excluyan cualesquiera ponderaciones sobre las preferencias personales que procedan de características personales irrelevantes. Esto es lo que primariamente significa la universalizabilidad: se ha de valorar del mismo modo las cosas que son iguales una vez apartadas las diferencias que son moralmente irrelevantes.

J. L. Mackie distinguió tres interpretaciones de la universalizabilidad en relación al grado en que excluyen diferencias irrelevantes¹⁸. En su forma más débil la universalizabilidad requiere que no se tengan en cuenta las meras diferencias numéricas; i. e., «las diferencias entre un individuo y otros simplemente como tales» (J. L. Mackie [1977], p. 83). Pero parece natural que tampoco las diferencias cualitativas marquen diferencias en lo que es relevante para la validez del juicio moral. Así, en una segunda interpretación, la universalizabilidad exige que las prescripciones se apliquen a todas las personas, «no importa cómo los individuos varíen en sus cualidades mentales y físicas, sus recursos y *status* social» (*ibidem*, p. 90). No obstante, ninguna de estas dos interpretaciones parece bastar al

Hemos tratado una cuestión similar a ésta en relación a la valoración de las políticas públicas en SALCEDO, D. (1988).

18. Seguimos en este trabajo el análisis de MACKIE, J. L. (1977) en tres «*stages*» de la noción de universalizabilidad, pero hay análisis más complejos; p. e., cfr. las 16 tesis de la universalizabilidad de NARVERSON, J. (1985).

utilitarismo. Según éste, la universalizabilidad impide que alguna de las concepciones del bien —sean del evaluador o de cualquiera de los afectados— tenga un peso especial en nuestro juicio. Sólo las preferencias de los afectados han de ser tenidas en cuenta y cualquier otra información ha de ser excluida.

Bajo tal interpretación, cuando el utilitarismo señala que las prescripciones para ser morales han de ser universalizables, ha de entenderse que nos está conminando a «mirar las cosas desde el propio punto de vista y desde el de la otra persona, y descubrir los principios que guíen la acción (...) que uno pueda aceptar desde ambos puntos de vista» (J. L. Mackie [1977], p. 93). No se trata simplemente de poder imaginarse lo que el otro preferiría, sino de ponerse realmente en su piel adoptando sus deseos, gustos, ideales, valores, sus cualidades, sus capacidades y hasta su posición social y económica¹⁹. Así es como una simple exigencia de impersonalidad se refuerza hasta ser una fuerte exigencia moral de imparcialidad. Cuando esta interpretación se acepta, una prescripción moral consiste en una expresión de preferencia porque algo se haga siempre, dado que todos los afectados por la acción tengan las mismas preferencias.

Naturalmente la imparcialidad que requiere esta interpretación de la universalizabilidad excluye del ámbito de la moralidad los valores relativos a la posición del evaluador que acepta el neoconsecuencialismo. Al manifestar que, si bien yo no tengo razones para impedir que tu hagas X, sí las tengo para no hacerlo yo, el utilitarista me advierte de que estoy introduciendo una diferencia moralmente irrelevante entre mis preferencias y las tuyas. Además, según él, estoy cometiendo una incoherencia al prescribir una acción que no considero que deba hacerse siempre. Este privilegio que otorgo a mis ideales, a la forma en que veo las cosas en relación a los de otra persona imposibilita el reconocimiento como moral de mi prescripción puesto que de ningún modo es imparcial.

Hemos visto que la exigencia de universalizabilidad proviene de la necesidad que tienen los juicios morales de ser impersonales. La impersonalidad parece un requisito natural, puesto que sería raro que alguien tratara de pasar como moral algún juicio que sólo fuera la expresión de

19. Este requisito suele ser más explícito en el utilitarismo económico en donde viene impuesto por la necesidad de las comparaciones interpersonales; cfr. HARSANYI, J. C. (1977); SEN, A. (1979c), y MIRLEES, J. A. (1982). En HARE, R. M. (1981) se presenta como un supuesto que da materialidad a la mera exigencia formal de la imparcialidad; GIBBARD, A. (1988) lo ha llamado «*Conditional Reflection Principle*», denominación que Hare ha aceptado. Por otro lado, es ese principio el que hace de intermediario para pasar desde la imparcialidad al utilitarismo y de su no plena comprensión dependió que en *Freedom and Reason* Hare no diera finalmente ese paso; cfr. GORR, M. (1985) y NIELSEN, K. (1985) que argumentan la imposibilidad de derivar de la universalizabilidad un principio sustantivo de imparcialidad como el requerido por el utilitarismo.

su punto de vista más irreductiblemente subjetivo. No obstante, ¿por qué la impersonalidad ha de ser interpretada como imparcialidad? Mi opinión es que el utilitarismo se ve abocado a esa interpretación tan fuerte en razón de lo que ha tomado como hechos moralmente relevantes. Recordemos que es característico del utilitarismo desde su fundación por J. Bentham señalar que la capacidad de goce y sufrimiento es el dato último de la moral. Es de esta capacidad de donde surgen los intereses de los seres sintientes y las preferencias que manifiestan las personas. En mi opinión, son las nociones de «interés» y «preferencia» las que obligan y al mismo tiempo permiten despojar a los juicios morales de referencias posicionales.

Por un lado, los intereses y preferencias son peculiarmente asuntos muy personales; son productos de la especificidad más irreductiblemente particular de las personas y su situación. Ello impide que puedan constituir inmediatamente valores morales. Una preferencia sólo puede ser ética, si no es la preferencia de nadie en particular. Por otro lado, los intereses y preferencias no conllevan ninguna demanda moral especial sobre su consideración. Todas las preferencias son igualmente preferencias en la medida en que todos los individuos tienen igualmente una capacidad de sufrir y gozar. Así un principio que tenga por fuente de información los intereses o las preferencias de todos los individuos afectados será moral sólo si las trata bajo un supuesto de igual consideración. Podríamos decir que el hecho de que intereses y preferencias siempre sean intereses y preferencias para alguien obliga a la preferencia ética a suprimir la referencia a la persona de quienes provienen; pero ello a su vez es permitido por el hecho de que todas las personas sean igualmente portadores de intereses y preferencias. Por ello sólo la fuerte exigencia de imparcialidad es capaz de lograr que los juicios morales basados en preferencias sean impersonales²⁰.

Ahora bien, ¿por qué los intereses y preferencias? ¿Por qué sólo son ellos la información moralmente relevante? R. M. Hare ([1981], s. 5.1) sostiene que tales hechos son seleccionados por los principios morales que es racionalmente adoptar. Según él, como asunto de prioridad lógica, no sabremos qué información es la que requiere la moral hasta que no sepamos cuáles son los principios morales. Pero yo sospecho que es verdad también de la teoría de Hare algo que lo es del resto de los utilitaristas. Ellos han adoptado la interpretación más fuerte de la universalizabilidad

20. El utilitarismo no sólo excluye cualquier otra información moralmente relevante que no sea la que aportan las preferencias, sus intensidades y las comparaciones interpersonales: además trata la información sobre cualesquiera otras cosas — ideales, obligaciones, derechos, etc.— como si fuera información de preferencias. Así, decir que algo es un valor importante, siempre significa para un utilitarista que alguien lo prefiere. De ahí que no haya modo de poner límites a las exigencias de la imparcialidad: cfr. SINGER, P. (1979) y (1988).

porque han tomado una decisión previa sobre lo que constituyen los datos morales. No al contrario. Sólo porque los intereses y preferencias surgen de una capacidad para gozar y sufrir que todos los seres sintientes tienen en cuanto tales, la imparcialidad es la actitud adecuada ante ellas. Hacer de un interés o una preferencia un punto privilegiado sería admitir un privilegio personal que la moral rechazaría.

Si este es el caso, si no son los principios los que determinan cuál es la información moralmente relevante en el utilitarismo²¹, entonces ¿por qué los intereses y preferencias son considerados los únicos hechos relevantes? Creo que la explicación de ello se encuentra en la teoría de las razones para la acción que acepta el utilitarismo.

La teoría que el utilitarismo sostiene, como es conocido, es que cualquier razón para actuar tiene que surgir de los deseos, preferencias, intereses, etc., del agente. Hay maneras diversas de defender esta tesis. Una de ellas es la que vimos incorporada en la teoría general del comportamiento racional de Harsanyi. Actuar racionalmente es actuar consistentemente en relación a un conjunto de preferencias y al coste de oportunidad de las alternativas. Por consiguiente, actuar en base a lo único que puede dar una razón para actuar: las preferencias de los individuos. Al identificar el ámbito de la moral con aquellas situaciones en las que hay que tomar una decisión que afecta a una sociedad o a un conjunto de ella, aparece especificado el problema de la racionalidad ética: determinar en qué consistirá la racionalidad de esa decisión, una racionalidad que ya es colectiva. Para resolverlo el utilitarismo aplica el mismo esquema que ha sido eficaz en las decisiones individuales. La racionalidad de una decisión social depende de la consistencia entre las preferencias sociales y la elección social. Los principios utilitaristas tienen que velar por esa consistencia. Pero no sólo por ella. Han de velar también por una relación aceptable entre preferencias individuales y preferencias sociales. Aquí es donde la imparcialidad juega su papel, permitiendo una relación no viciada entre las razones de los individuos y las razones sociales; o, de otro modo, entre lo que quiere cada uno y lo que todos quieren. Así la fuerza de las razones personales es transferida a las razones morales.

Sin embargo, una vez entendido este proceso, la cuestión fundamental sigue pendiente: ¿por qué se considera que los deseos, los intereses o las

21. En cualquier teoría pueden analizarse los principios a partir de la información que hacen moralmente relevante o viceversa; cfr. SEN, A. (1979b). Lo que aquí sostenemos es que mientras los principios utilitaristas determinan la forma precisa en que han de ser tratadas las preferencias —ya sea como relaciones ordinales, como funciones cardinales de utilidad, o de cualquier otro modo lógicamente posible—, el que la información se limite a las preferencias no es algo que en la lógica interna del utilitarismo esté determinado por los principios, sino que proviene de la teoría de las razones por las que las personas actúan y, últimamente, de lo que es importante para el hombre.

preferencias son las únicas razones para actuar? La teoría de las razones para la acción se configura en torno al problema de la motivación y lo característico del utilitarismo es entender ese problema bajo dos supuestos²²:

1. Una explicación de la razón práctica debe ser tal que haya una conexión necesaria entre las razones y la acción.
2. Cualquier hecho que constituya una razón última para actuar debe ser algo que pueda ser conocido sin necesidad de una prueba o una evidencia ulterior (R. Norman [1971], p. 77).

La apelación a los deseos —palabra que utilizaremos en un sentido amplio que permita abarcar intereses, gustos, preferencias, etc.— es lo único que satisface estas dos exigencias. Por un lado, tener un deseo de algo establece una conexión necesaria entre la acción y la razón por la que se realiza. El deseo puede hacer tal cosa por su carácter dual que implica tanto la apreciación del objeto como la descripción del estado mental del sujeto. Alguien que manifiesta que desea algo tiene que intentar realizarlo; en caso contrario, hay que suponer que no lo desea realmente. Por otro lado, ese mismo carácter dual presta una vivacidad a la apelación de deseos que les hace aparecer como si ellos constituyesen el fundamento último inteligible de la acción. Cuando alguien dice que ha realizado algo porque lo quiere, ya no hay nada más que preguntar. Ninguna otra cosa puede explicar la conducta intencional como los deseos. Y, de este modo, actuar por razones sólo puede ser actuar por deseos.

Vemos así cómo se anudan las distintas tesis a efectos de conducirnos a la conclusión de que sólo pueden ser reconocidos como morales los valores neutrales de carácter utilitarista. Por un lado, sólo los deseos son razones para actuar. Por otro, sólo aquel valor que atribuimos a la capacidad de sufrimiento y de goce marca el límite de la actuación moral. Así los valores o las prescripciones son expresiones de preferencias y, por tanto, dan razones por sí mismas para actuar. Pero una prescripción para ser moral tiene que ser impersonal. Por consiguiente, la única preferencia/prescripción que es moral es la que es imparcial. Por último, la única

22. Sigo la explicación de NORMAN, R. (1971) que resume bien la fuerza intuitiva que tiene la doctrina utilitarista de las razones para la acción. Sin embargo, hay que notar que esta tesis puede ser defendida de otras maneras. Una de ellas, como hemos visto, es la que proviene de las teorías formales de la decisión racional. Otra tiene su fuente en Hume y en su tesis del carácter motivacionalmente inerte de la razón. Un tercer tipo es la que puede encontrarse en HARMAN, G. (1976), quien sostiene que la razón práctica consiste sólo en hacer coherentes nuestras intenciones y deseos. Finalmente, citaré la corriente muy popular en las ciencias sociales y económicas que ejemplifica DAVIDSON, D. (1963), quien sostiene que, puesto que la única explicación de la acción intencional es la que proporcionan los deseos y creencias del agente, sólo ellos pueden ser razones de su actuación. Puede verse una valoración crítica de todos estos argumentos en DARWALL, S. L. (1983).

preferencia imparcial es la preferencia utilitarista. El círculo de supuestos racionales y morales es férreo. Si intentamos criticar alguno de estos supuestos, nos encontraremos enfrentados a los otros. La no admisión de uno de ellos significa la violación de los otros. Y la cuestión a la que siempre remiten es la misma: pero ¿qué otra cosa que los deseos da razones para actuar? Y si es así, ¿qué otra cosa pueden ser los valores morales, sino imparciales?

Impersonalidad y consideraciones racionales

El neoconsecuencialismo ha de defender una teoría alternativa de las razones para la acción que no le comprometa con la interpretación más fuerte de la universalizabilidad en orden a evitar verse abocado a alguna forma de utilitarismo. Para ello ha de mostrar dos cosas. En primer lugar, la posibilidad de razones con fuerza motivacional que no surjan de deseos. En segundo lugar, la posibilidad de que los juicios morales sean impersonales sin que también tengan que ser imparciales.

Comenzaremos por lo último. Hay que admitir que hacer un juicio moral sobre lo que se considera bueno o sobre lo que se debe hacer compromete al hablante a mantener que siempre valorará o prescribirá del mismo modo en las mismas circunstancias. Este es el significado general de la exigencia de universalizabilidad sobre los juicios morales. El neoconsecuencialismo ha de mostrar, entonces, que la explicación que da del juicio moral al hacerlo depender de la posición del evaluador no mina ese requisito de universalizabilidad.

El neoconsecuencialismo, como vimos, requiere una noción posicional de la bondad moral. Un juicio de valor sobre los estados que resultarían de adoptar un curso de acción —incluyendo los valores referidos al agente que definen la relación del agente y el valor de la acción— no puede obviar la forma en que está implicado el evaluador en ese estado resultante. Esta posicionalidad impide que la exigencia de imparcialidad se aplique al juicio, puesto que con el cambio de posición del evaluador el juicio puede variar. Así Otelo puede juzgar con más crueldad la muerte de Desdémona de lo que lo hago yo, puesto que él es ineludiblemente su esposo y asesino. Si analíticamente podemos distinguir al Otelo evaluador del Otelo esposo y asesino, hemos de admitir que la posición del Otelo evaluador está marcada por el papel de esposo y asesino del Otelo agente de la acción, de forma que su juicio inevitablemente incluirá parcialmente una ponderación particular del elemento del estado de cosas resultante que consiste en que Otelo ha asesinado a su esposa.

Sin embargo, el juicio de valor que admite esta varianza posicional no tiene por qué negar que cualquier otra persona deba valorar del mismo modo los hechos. Los hechos son relevantemente similares de acuerdo con las restricciones que impone la universalizabilidad. El estado que se

juzga es el mismo para Otelo y para mí: Desdémona ha sido asesinada por su esposo. En su juicio y en el mío no puede haber diferencias por el hecho de que Otelo y yo seamos «cuantitativa» y «cualitativamente» diferentes —en el sentido en que J. L. Mackie daba a estos términos. Yo juzgaría tan cruelmente como lo hace Otelo el asesinato de Desdémona, si yo estuviera en su posición²³. Lo que diferencia nuestros juicios es la manera en que Otelo y yo estamos implicados en la acción. Yo juzgo la muerte de Desdémona por su esposo, pero es decisivo que yo no sea su esposo y asesino. Parecería poco natural exigir que los dos atribuyésemos un mismo valor (neutral) a esa acción.

Así cumplir con los requisitos de las dos primeras interpretaciones de la universalizabilidad puede ser suficiente para que el juicio de valor sea impersonal y, por consiguiente, moral. Lo que no requerimos es que sea imparcial; i. e., que no tome en cuenta cómo los compromisos, obligaciones, objetivos, etc., del agente marcan la evaluación de un modo que no haya de tener sentido decir que podría haber evaluaciones distintas correspondientemente a cómo el evaluador esté implicado en la acción. Con ello las exigencias primarias a la universalizabilidad quedan satisfechas, puesto que el evaluador queda comprometido en la prescripción universal del juicio, si bien ahora de una forma similar a ésta: «Quienquiera que tenga mi misma posición ha de juzgar del mismo modo los hechos».

La noción posicional de bondad moral, por consiguiente, lo que hace desaparecer es la idea cara al utilitarismo de que se pueden juzgar las consecuencias de una acción por una función de bienestar igual para todas las personas; o, en palabras de D. H. Regan, la posibilidad de definir «un bien universal que todos los agentes tienen que promover» ([1983], p. 93). La acción correcta sigue siendo determinada por sus buenas consecuencias. Pero lo que son buenas consecuencias ya no es un valor único y neutral alcanzable imparcialmente. Por el contrario, se trata de una noción informativamente plural que varía con la posición del evaluador²⁴.

Como se sigue de nuestra argumentación, la posibilidad de una evaluación consecuencial que tiene en cuenta valores referidos al agente depende de que sea posible valorar los estados del mundo por algo distinto a la actitud mental hacia esos estados de los afectados —ya sean prefe-

23. La distinción fundamental aquí es entre la posición del evaluador y la del autor. Mientras que el juicio de valor puede variar con el cambio de posición del evaluador, la universalizabilidad exige que el juicio de valor no varíe con la identidad del autor de la acción. Así, si yo hubiera sido el autor de la muerte de Desdémona y su esposo, he de juzgar lo mismo de cruelmente que lo hace Otelo; cfr. SEN, A. (1985), pp. 183-4.

24. Gracias a esta estructura plural se mantiene la relación entre las nociones de lo bueno y lo correcto propia de una ética consecuencialista: cfr. REGAN, D. (1983) y la respuesta de SEN, A. (1983) y (1985).

rencias, deseos, intereses, etc. El utilitarismo está comprometido con esta concepción por aceptar una doctrina de las razones para la acción que identifica a los deseos como las únicas razones válidamente justificativas de la conducta. Pero, si sólo se tienen en cuenta los deseos o preferencias para valorar estados, ni la acción ni el agente que la realiza pueden entrar en esa evaluación. Así la evaluación está condenada a ser neutral y sólo se pronunciará sobre lo que habrá de acontecer en el mundo para que se vea satisfecho tal valor neutral.

Th. Nagel ha criticado esta posición en la doctrina de Hare, apelando a una distinción conceptual que nos parece pertinente para la comprensión del neoconsecuencialismo:

(...) su idea (de Hare) de que los juicios morales son universalmente prescriptivos significa que dependen de lo que *se quiera que suceda*, considerando las cuestiones desde todos los puntos de vista —más bien que de aquello que se piensa que los individuos *tienen razón de hacer*, considerando la cuestión de este modo. (Th. Nagel [1986], p. 163; subrayados del autor de la cita)

Lo que Nagel señala puede parafrasearse en términos de la distinción entre «preferencias que dan razones» y «razones para preferir». La comprensión de las exigencias de estas dos concepciones de las razones para la acción es fundamental. Cuando se concibe que sólo del deseo puede surgir una razón para actuar, uno se ve abocado a sostener que una prescripción moral es aquella que manifiesta un deseo de que todos hagan algo. Puesto que un deseo sólo puede ser moral si es el deseo imparcial de todos por la acción. Sin embargo, tener una razón para preferir no implica un deseo de que cualquier otra persona haga lo mismo que prescribo. Simplemente expresa una demanda sobre cómo cada uno debe actuar «tal que si alguien sinceramente asienta a ella, ha de reconocer una razón para actuar de ese modo» (Th. Nagel [1988], p. 106)²⁵. Reconocer una prioridad de las razones sobre las preferencias nos permite dejar abierto el concepto de razones siendo posible que sean tanto neutrales como referidas al agente. Serán neutrales cuando los intereses de todas las partes afectadas reciban una ponderación igual. Pero permite también que las razones que el agente acepta reciban un peso especial en la evaluación fi-

25. R. M. Hare señala que Nagel ha confundido dos cosas al hacer esta interpretación de lo que es prescribir. Una verdadera: hacer un juicio moral es pretender que cualquiera que lo acepte habrá adquirido una razón; y una falsa: hacer un juicio moral es pretender que todos —lo acepten o no— ya tienen una razón para actuar de acuerdo con él; cfr. los comentarios de Hare al artículo citado de Nagel en D. Seanor-N. Fotion (eds.) (1988). Sin embargo, esto no me parece una crítica que se pueda hacer a Nagel o a la teoría que estamos exponiendo. Hacer un juicio moral es pretender que cualquiera tendría una razón, si considerase las cosas desde la posición del que lo hace.

nal. Nagel considera que una buena parte de estas razones establecen restricciones absolutas sobre el tipo de acción que tenemos razón en llevar a cabo. Pero tal interpretación no es posible ni necesaria en el caso de una evaluación consecuencial. Juzgar los estados por su bondad moral — contrariamente a juzgarlos por las preferencias de los afectados— equivale a ponderar diversas variables que conforman ese estado y, naturalmente, entre ellas lo que el agente considera que está obligado a hacer en base a los valores que acepta en su actuación como agente. La realización o violación de estos últimos son, por consiguiente, datos que entran a formar parte de las consecuencias que se tratan de determinar.

El concepto central en esta explicación es, entonces, el de razones como «consideraciones racionales». La consideración racional de las consecuencias de una acción requiere reflexionar sobre el estado de cosas que resulta de ella. Si dicha reflexión es capaz de mover al evaluador hacia la acción o la recomendación de la acción, entonces inmediatamente esa reflexión es una razón para el evaluador. De este modo, el criterio para determinar si algo es una razón sigue permaneciendo fiel a la exigencia de que sea capaz de motivar a la acción²⁶. Pero, a diferencia de la doctrina utilitarista, se sostiene que los deseos no satisfacen el requisito de que sólo puede ser una razón aquello que parece más evidente en la explicación de la conducta. Argumentar que los estados mentales referidos no tienen la inteligibilidad propia que le atribuye el utilitarismo no es difícil (cfr. R. Norman [1971]). Y al contravenir este presupuesto, se puede evitar identificar razones y deseos. Esta doctrina excluye, por consiguiente, cualquier tentación de dar prioridad a los motivos sobre las razones.

Lo que más nos importa de esta noción de razones como consideraciones racionales es, naturalmente, que nos permite evitar la exigencia de que los juicios de valor sean imparciales. A menos que podamos suponer que todos los evaluadores serán motivados por la consideración de las mismas consecuencias a preferir lo mismo, no habrá ninguna base para esperar que existan juicios de valor imparciales. Lo más que parece razonable esperar es que en algunos casos suceda que desde toda posición se coincida en una misma evaluación; pero desde luego no en todos. En aquellos en que se registren variaciones posicionales significativas es seguro que los juicios serán diferentes. No obstante, ello no implica que estos juicios no satisfagan un requisito de universalizabilidad que asegure su moralidad. Tal cosa es posible en la medida en que la posicionalidad no implica arbitrariedad. Cualquiera que se sitúe en una posición evalua-

26. Así la noción de consideración racional permanece dentro de una visión «internalista» de las razones para la acción que, en mi opinión, podría desarrollarse siguiendo las pautas sugeridas por Hume en relación a la posibilidad de reglas generales: cfr. SALCEDO, D. (1987b). La noción de consideración racional está sugerida por A. Sen en su crítica a los supuestos conductuales de la teoría económica; cfr. SEN, A. (1973), (1974), (1976a y b).

tiva determinada hará el mismo juicio. Esta exigencia de impersonalidad es suficiente para garantizar que los juicios de valor resultantes de una evaluación consecuencial posicional son morales. Un requisito más fuerte de imparcialidad es entonces una opción de una teoría ética específica que requiere justificación.

Por último, esta noción de razones como consideraciones racionales nos permite contemplar una noción de persona moral diferente a la que se extrae del utilitarismo y que, como señalamos, ha sido objeto tradicional de crítica y motor de la revisión neoconsecuencialista. Ciertamente, la defensa de valores neutrales basados en preferencias obliga a pensar en las personas como simples ocasiones en las que ciertos estados mentales ocurren y, de ahí, considerar la dimensión moral como una esfera a la que le incumben exclusivamente estados de cosas en el mundo y cuánto éstos maximicen la aparición de estados mentales deseables. Se pierde así algo que es esencial a la noción de persona como agente moral, a saber: que cuando deseamos que algo suceda, no sólo es importante que tal cosa suceda; sino que ese deseo es inseparable del deseo de que se produzca como un resultado de nuestro desear que suceda. De ahí la importancia que para la persona moral tiene su propia capacidad de obrar en el mundo y de ver realizados sus planes y objetivos. Este cuadro más complejo —que el que el utilitarismo consiente— de las razones para la acción es el que permite comprender el papel de la integridad, el respeto por las decisiones autónomas, las obligaciones especiales, los compromisos, etc. Puesto que sólo esta noción de agente moral da sentido a las acciones que se llevan a cabo en relación con una concepción —más o menos elaborada— que cada agente moral tiene de lo que es una vida valiosa. Cuando se entienden así las cosas, uno comprende la legitimidad de la crítica antiutilitarista que señala ya sea la incompreensión de esta teoría ética hacia los distintos «modos de vida valiosos» (Hampshire), ya sea la violación que lleva a cabo del principio fundamental de la incomparabilidad (*separateness*) de las personas (Rawls, Richards), o bien la alienación que exige del sentido de integridad y responsabilidad de las personas (Williams)²⁷.

Lo que nos interesa señalar es que bajo la concepción de las razones para la acción como consideraciones racionales la imagen de la persona

27. Comprender la legitimidad de estas críticas no es, naturalmente, aceptar como correctas sus fuentes. Mi opinión es que todas aquellas teorías éticas en que la imparcialidad es un requisito de la moralidad del juicio de valor producen una imagen errónea del agente moral. Esto es verdad del utilitarismo como argumentan sus críticos; pero también de una buena parte de las teorías de éstos y, en particular, de las de filiación kantiana —como las de Rawls y Richards—. En este sentido creo acertada la crítica de WILLIAMS, B. (1976b) que se dirige tanto al primero como a las segundas, fundada en el hecho de que ambos pueden demandar de las personas que renuncien a los proyectos que son el núcleo esencial de su identidad personal.

como agente moral que se sigue es radicalmente diferente de la utilitarista. Fundada en la noción de «razones para preferir», permite entender la conducta racional no como una actuación consistente con un ordenamiento de preferencias, sino como el ejercicio de una capacidad de considerar y revisar las preferencias en orden a buscar razones que las apoyen o las rechacen. Así la racionalidad no consiste primariamente en elegir adecuadamente —en una maximización—, sino en determinar nuestras preferencias de acuerdo con razones que consideremos poderosas. En la perspectiva neoconsecuencialista estas razones no implican una apelación independiente a normas de lo correcto. Por el contrario se trata de apelación a una noción de bondad moral capaz de unificar en una sola perspectiva el deseo del agente de que suceda lo mejor como su deseo de que lo que suceda se origine de modo que él lo reconozca sin repugnancia como suyo. Así es posible dar un peso importante a la diversidad de concepciones del bien que pueden tener las personas y a su valor de agentes únicos e inconmensurables. Pero ello sin perder de vista que la consideración racional de las consecuencias no es una forma encubierta de establecer límites absolutos a lo que se puede considerar bueno²⁸. Los juicios de bondad moral son el resultado de una ponderación peculiar de valores referidos al agente y valores neutrales que realiza un sujeto moral capaz de comprender la importancia tanto de realizar lo correcto como de realizar lo mejor.

Conclusiones

Mientras que nuestro pensamiento moral común parece fundamentalmente estructurado en torno a razones morales referidas al agente, el utilitarismo ha sostenido que la validez moral de cualquiera de estas razones sólo podría establecerse en relación a razones neutrales con respecto al agente. La tradición utilitarista y sus críticos comparten la idea de que es imposible un consecuencialismo referido al agente. Nuestro objetivo primario, por consiguiente, ha sido mostrar la posibilidad de una teoría ética de esas características. Tal y como hemos comprendido esta cuestión, dicha posibilidad depende de que seamos capaces de realizar dos cosas.

En primer lugar, hemos de mostrar que la neutralidad viene impuesta

28. En la discusión sobre las razones para la acción se suele presentar la cuestión de la universalizabilidad como si sólo hubiera dos caminos: el que lleva hacia una moral heterónoma que funda las razones en deseos y el que lleva hacia una moral autónoma fundada en la capacidad autónoma de la voluntad; cfr. DARWALL, S. L. (1983) y O'NEILL, O. (1985). La propuesta neoconsecuencialista apunta la posibilidad de definir una teoría que no apela como último criterio de aceptabilidad moral a lo que se quiere sin que por ello tenga que renunciar a mantener la prioridad de lo bueno sobre lo correcto.

por una interpretación muy fuerte de la universalizabilidad. Esta interpretación parte de una idea simple y fácil de aceptar: los juicios morales que emite una persona han de estar libres de ciertas clases de arbitrariedad que tienen su origen en que el juicio sea de *esa* persona. Hemos sostenido que si desde esta exigencia natural de impersonalidad el utilitarismo deriva una exigencia muy restrictiva de imparcialidad, ello se debe a la naturaleza que atribuye a los juicios morales. Dado que en esta teoría los juicios morales no son más que expresiones de preferencias, deseos, actitudes, etc., la única forma de conseguir que sean completamente impersonales es que sean imparciales. Gran parte de la fortaleza de la posición utilitarista proviene de la demostración de que no se puede aceptar la universalizabilidad y al mismo tiempo no entenderla como imparcialidad. Pero, creo que hemos conseguido hacer ver que se puede asegurar la universalizabilidad de juicios morales no imparciales. Naturalmente, para ello ha sido necesario abandonar la idea utilitarista de la bondad moral.

En la teoría neoconsecuencialista en que nos apoyamos los juicios morales son siempre posicionales. Así, la bondad moral no es la satisfacción de la preferencia (universalizable) de alguien por que algo suceda en lugar de otra cosa; sino, de forma más compleja, la preferencia de alguien por que las cosas que sucedan se realicen de un determinado modo (o no se realicen de un determinado modo). Esta relevancia de la «realización» de la acción en el juicio moral es la que hace inexcusable la referencia al «alguien» que emite el juicio moral. En esta concepción de la naturaleza de la bondad moral queda recogida la idea de que la preferencia por que sucedan buenas cosas no se ha de hacer al margen de la visión del evaluador sobre lo que está moralmente permitido hacer o no. A diferencia de los utilitarismos indirectos, esta teoría ética no subordina las razones referidas al agente a las razones neutrales, sino que es en la propia naturaleza plural de la bondad moral donde el evaluador ha de encontrar el balance adecuado a su propia posición de unas y otras. De este balance resultará un juicio *moral*, si el juicio se ajusta a las restricciones propias de la impersonalidad; y, en este sentido, satisfará las exigencias de la noción de universalizabilidad —en las dos primeras interpretaciones de Mackie. Pero, una vez concebida de esta manera plural la bondad moral, tratar de suprimir las características de la posicionalidad evaluativa imponiendo también la imparcialidad es ya algo difícil de entender.

La segunda clave de una teoría ética neoconsecuencialista radica en la revisión de la noción de razones para la acción del utilitarismo. Naturalmente, se podría insistir en que la noción de bondad moral es más compleja de lo que el utilitarismo supone. Pero si esa afirmación no ha de quedarse en una mera observación gratuita, ha de encontrar algún fundamento. Los pilares profundos del utilitarismo se asientan sobre la idea de que sólo pueden constituir razones para la acción de los deseos. Así, puesto que los juicios morales son guías de acción o prescripciones;

y puesto que las razones para la acción son de naturaleza no cognitiva, el utilitarismo sostiene que la única manera de entender una guía de acción como razón para actuar es entendiéndola como la expresión de una preferencia. Desafiar estas ideas es desafiar teorías fuertemente asentadas en el campo especulativo y empírico de las ciencias sociales y, últimamente, ideas centrales sobre lo que es el comportamiento racional. Pero el neoconsecuencialista tiene que llevar a cabo ese desafío para defender la naturaleza posicional de los juicios morales y, finalmente, una noción pluralista de bondad moral.

En este trabajo no hemos pretendido proporcionar la teoría completa de las razones para la acción que el neoconsecuencialismo requiere —una tarea que de todos modos será ineludible—. Más modestamente hemos tratado de mostrar que desafiar a la teoría utilitarista no conduce necesariamente al fracaso, sugiriendo algunas líneas por las que podría discurrir una teoría alternativa. La idea de «consideración racional» pretende señalar el lugar de una distinción fundamental para nuestro caso: la distinción entre la noción de «preferencias que dan razones» y la noción de «razones para preferir». Mientras que en la primera se produce una identificación entre razones y motivos, en la segunda la razón aparece como dotada de una inteligibilidad anterior que no pertenece por derecho propio al motivo. Dicha razón habrá de ser una combinación de consideraciones sobre lo que son las mejores consecuencias en un caso dado y entre dichas consideraciones pueden encontrarse elementos cognitivos evaluativos —pues ésta parece la naturaleza de los valores deontológicos—. Cuando dicha razón además genere un motivo para actuar —una preferencia por la acción y sus resultados—, entonces estaremos en presencia de una razón capaz de guiar nuestra conducta.

Con este esbozo es suficiente para permitirnos señalar que las preferencias por sí mismas no agotan el significado de los juicios morales. Un juicio moral siempre alude a una combinación compleja de consideraciones entre las que no son las menos importantes las que afectan a la situación del agente y a la posición del propio evaluador con respecto a éste y su acción. La noción posicional de bondad moral refleja esta realidad compleja, al fundar su posibilidad en la idea de que la razón práctica misma no se deja agotar en su fuerza motivacional.

BIBLIOGRAFIA

- ARTHUR, J.-SHAW, W. H. (eds.) (1978): *Justice and Economic Distribution*, Englewood-Cliffs, N. J.: Prentice Hall.
- BRANDT, R. B. (1970): «Some Merits of One Form of Rule Utilitarianism», en K. Pabel-M. Schiller (eds.), 1970.
- BRANDT, R. B. (1979): *A Theory of the Good and the Right*, Oxford: Clarendon P.
- DANIELS, N. (ed.) (1975): *Reading Rawls*, Oxford: Blackwell.

- DARWALL, S. L. (1983): *Impartial Reason*, Ithaca, N. Y.: Cornell U. P.
- DAVIDSON, D. (1963): «Actions, Reasons and Causes», *Journal of Philosophy*, 60.
- DWORKIN, R. (1973): «The Original Position», N. Daniels (ed.), 1975.
- FOOT, Ph. (ed.) (1967): *Teorías sobre la ética*, trd. M. Arbolí, México: F.C.E., 1974.
- FOOT, Ph. (1985): «Utilitarianism and the Virtues», S. Scheffler (ed.), 1988.
- GARDENFORS, P.-SAHLIN, N.-E. (eds.) (1988): *Decision, Probability and Utility*, Cambridge: Cambridge U. P.
- GIBBARD, A. (1988): «Hare's Analysis of "Ought" and its Implications», D. Seanor-N. Fotion (eds.), 1988.
- GORR, M. (1985): «Reason, Impartiality and Utilitarianism», N. T. Potter-M. Timmons (eds.), 1985.
- GOTTINGER, H. W.-LEINFELLNER, W. (eds.) (1978): *Decision Theory and Social Ethics*, Dordrecht: D. Reidel Pu. Co.
- HAMPSHIRE, St. (1978a): «La moral y el pesimismo», St. Hampshire (ed.), 1978b.
- HAMPSHIRE, St. (ed.) (1978b): *Moral pública y privada*, trad. M. Córdoba, México: F.C.E., 1983.
- HARE, R. M. (1973): «Rawls' Theory of Justice», N. Daniels (ed.), 1975.
- HARE, R. M. (1976): «Ethical Theory and Utilitarianism», A. Sen-B. Williams (eds.), 1982.
- HARE, R. M. (1981): *Moral Thinking*, Oxford: Clarendon P.
- HARMAN, G. (1976): «Practical Reasoning», *Review of Metaphysics*, 29.
- HARRISON, R. (ed.) (1979): *Rational Action*, Cambridge: Cambridge U. P.
- HARSANYI, J. C. (1955): «Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility», J. C. Harsanyi, 1976b.
- HARSANYI, J. C. (1975a): «Can the Maximin Principle Serve as a Basis for Morality? A critique of J. Rawls' Theory», J. C. Harsanyi, 1976b.
- HARSANYI, J. C. (1975b): «Non-Linear Social Welfare Functions: Do Welfare Economists Have a Special Exemption from Bayesian Rationality?», J. C. Harsanyi, 1976b.
- HARSANYI, J. C. (1976a): «Advances in Understanding Rational Behavior», J. C. Harsanyi, 1976b.
- HARSANYI, J. C. (1976b): *Essays on Ethics, Social Behavior and Scientific Explanation*, Dordrecht: D. Reidel Pu. Co.
- HARSANYI, J. C. (1977): «Morality and the Theory of Rational Behavior», A. Sen-B. Williams (eds.), 1982.
- HARSANYI, J. C. (1978): «Rule Utilitarianism and Decision Theory», H. W. Gottinger-W. Leinfellner (eds.), 1978.
- HARSANYI, J. C. (1980): «Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior», *Theory and Decision*, 12.
- HARSANYI, J. C. (1982): «Some Epistemological Advantage of the Rule Utilitarianism Position in Ethics», *Midwest Studies in Philosophy*, 7.
- HARSANYI, J. C. (1985): «Rule Utilitarianism, Equality and Justice», E. F. Paul, et al. (eds.), 1985.
- HARSANYI, J. C. (1988): «Problems with Act Utilitarianism and with Malevolent Preferences», D. Seanor-N. Fotion (eds.), 1988.
- HODGSON, D. H. (1967): *Consequences of Utilitarianism*, Oxford U. P.
- KORNER, S. (ed.) (1974): *Practical Reason*, Oxford: Blackwell.
- MACKIE, J. L. (1973): «The Disutility of Act Utilitarianism», *Philosophical Quarterly*, 23.
- MACKIE, J. L. (1977): *Ethics: Inventing Right and Wrong*, Harmondsworth, Mid.: Penguin.
- MIRLEES, J. A. (1982): «The Economic Uses of Utilitarianism», A. Sen-B. Williams (eds.), 1982.

- MUELLER, D. C. (1979): *Elección pública*, trad. J. C. Zapatero, Madrid: Alianza, 1984.
- NAGEL, Th. (1979): *La muerte en cuestión*, trad. C. Valdés, México: F.C.E., 1981.
- NAGEL, Th. (1986): *The View from Nowhere*, N. Y.: Oxford U. P.
- NAGEL, Th. (1988): «*The Foundations of Impartiality*», D. Seanor-N. Fotion (eds.), 1988.
- NARVERSON, J. (1985): «*The How and Why of Universalizability*», N. T. Potter-N. Timmons (eds.), 1985.
- NIELSEN, K. (1985): «*Universalizability and the Commitment to Impartiality*», N. T. Potter-M. Timmons (eds.), 1985.
- NORMAN, R. (1971): *Reasons for Actions*, Oxford: Blackwell.
- NOZICK, R. (1974): *Anarchy, State and Utopia*, Oxford: Blackwell.
- O'NEILL, O. (1985): «*Consistency in Action*», N. T. Potter-M. Timmons (eds.), 1985.
- PAHEL, K.-SCHILLER, M. (eds.) (1970): *Readings in Contemporary Ethical Theory*, Englewood-Cliffs, N. J.: Prentice Hall.
- PAUL, E. F., et alii (eds.) (1985): *Ethics and Economics*, Oxford: Blackwell.
- PETTIT, Ph. (1984): «*Satisficing consequentialism*», Proceedings of the Aristotelian Society, Supp. Vol. 58.
- PETTIT, Ph.-BRENNAN, G. (1986): «*Restrictive Consequentialism*», Australasian Journal of Philosophy, 64.
- POTTER, N. T.-TIMMONS, M. (eds.) (1985): *Morality and Universality*, Dordrecht: D. Reidel Pu. Co.
- RAWLS, J. (1971): *A Theory of Justice*, Oxford: Oxford U. P.
- RAWLS, J. (1982): «*Social Unity and Primary Goods*», A. Sen-B. Williams (eds.), 1982.
- RICHARDS, D. A. J. (1971): *A Theory of Reasons for Action*, Oxford: Oxford U. P.
- RICHARDS, D. A. J. (1988): «*Prescriptivism, Constructivism and Rights*», D. Seanor-N. Fotion (eds.), 1988.
- REGAN, D. H. (1983): «*Against Evaluator Relativity*», Philosophy and Public Affairs, 12.
- SALCEDO, D. (1987a): «*La fundamentación normativa de un principio igualitarista*», Anuario de Filosofía del Derecho, 4.
- SALCEDO, D. (1987b): «*Intersubjetividad y reglas generales en la teoría del valor estético de Hume*», Agora, 6.
- SALCEDO, D. (1988): «*Derechos y bienestar*», Anales de la Cátedra Francisco Suárez, 28.
- SCANLON, T. M. (1975): «*Preference and Urgency*», Journal of Philosophy, 72.
- SCANLON, T. M. (1978): «*Los derechos, las metas y la justicia*», St. Hampshire (ed.), 1978.
- SCHIFFLER, S. (1982): *The Rejection of Consequentialism*, Oxford: Clarendon P.
- SCHIFFLER, S. (1985): «*Agent-Centred Restrictions, Rationality and the Virtues*», S. Schiffler (ed.), 1988.
- SCHIFFLER, S. (ed.) (1988): *Consequentialism and its Critics*, Oxford: Oxford U. P.
- SEANOR, D.-FOTION, N. (eds.) (1988): *Hare and Critics*, Oxford: Clarendon P.
- SEN, A. (1973): «*Behaviour and the Concept of Preference*», A. Sen, 1982b.
- SEN, A. (1974): «*Choice, Orderings and Morality*», S. Körner (ed.), 1974.
- SEN, A. (1976a): «*Liberty, Unanimity and Rights*», A. Sen 1982b.
- SEN, A. (1976b): «*Rational Fools. A Critique of the Behavioural Foundations of Economic Theory*», A. Sen, 1982b.
- SEN, A. (1979a): «*Utilitarianism and Welfarism*», The Journal of Philosophy, 76.
- SEN, A. (1979b): «*Informational Analysis of Moral principles*», R. Harrison (ed.), 1979.
- SEN, A. (1979c): «*Interpersonal Comparison of Welfare*», A. Sen, 1982b.

- SEN, A. (1982a): «*Rights and Agency*». *Philosophy and Public Affairs*, 11.
- SEN, A. (1982b): *Choice, Welfare and Measurement*, Oxford: Blackwell.
- SEN, A. (1983): «*Evaluator Relativity and Consequential Evaluation*», *Philosophy and Public Affairs*, 12.
- SEN, A. (1985): «*Well-being, Agency and Freedom*», *The Journal of Philosophy*, 82.
- SEN, A.-WILLIAMS, B. (eds.) (1982): *Utilitarianism and beyond*, Cambridge: Cambridge U. P.
- SINGER, P. (1972): «*Is Act-Utilitarianism Self-Defeating?*», *Philosophical Review*, 81.
- SINGER, P. (1979): *Ética práctica*, trad. M. I. Guastavino, Barcelona: Ariel, 1984.
- SINGER, P. (1988): «*Reasoning towards Utilitarianism*», D. Seanor-N. Fotion (eds.), 1988.
- SLOTE, M. (1985): *Common-sense Morality and Consequentialism*, London: Routledge&K. Paul.
- SMART, J. J. C. (1956): «*Utilitarismo extremo y restringido*», Ph. Foot (ed.), 1967.
- SMART, J. J. C. (1973): «*Bosquejo de un sistema de ética utilitarista*», J. J. C. Smart-B. Williams, 1973.
- SMART, J. J. C. (1978): «*Distributive Justice and Utilitarianism*», J. Arthur-W. H. Shaw (eds.), 1978.
- SMART, J. J. C.-WILLIAMS, B. (1973): *Utilitarismo: pro y contra*, trad. J. Rodríguez, Madrid: Tecnos, 1981.
- WILLIAMS, B. (1973): «*Una crítica del utilitarismo*», J. J. C. Smart-B. Williams, 1973.
- WILLIAMS, B. (1976a): «*Persons, Character and Morality*», B. Williams, 1981.
- WILLIAMS, B. (1976b): «*Utilitarianism and Moral Self-indulgence*», B. Williams, 1981.
- WILLIAMS, B. (1981): *Moral Luck*, Cambridge: Cambridge U. P.