

*Reseñas*

**Elliot Sober & David Sloan Wilson** *Unto Others. The Evolution and Psychology of Unselfish Behavior.* Cambridge, Ma./London: Harvard University Press, 1998. 394 p. ISBN 0-674-93047-9.<sup>1</sup>

El debate en torno a la realidad psicológica del altruismo se ha convertido en este libro en un punto de encuentro entre la filosofía, la psicología y la biología. El filósofo de la ciencia y de la biología E. Sober y el biólogo D.S. Wilson han unido esfuerzos para presentarnos una defensa de dos "especies en peligro de extinción" en las ciencias sociales y biológicas: el concepto evolucionista del altruismo, y el concepto psicológico y filosófico de la motivación desinteresada. Se trata de dos conceptos lógicamente separables. Un comportamiento es biológicamente altruista cuando sus *efectos* mejoran la aptitud biológica de un individuo distinto al actor. En este concepto no cuentan los motivos; el comportamiento puede ser biológicamente altruista, incluso cuando se trata de un organismo que no goza del privilegio de la intencionalidad. En cambio, en el concepto psicológico el motivo es fundamental: sólo se es altruista en sentido psicológico, si el bienestar de otro es representado por el organismo como

fin último, y no como un medio para su propio bienestar. Ambos conceptos han tenido una historia desafortunada, que sin embargo está en curiosa desproporción con las evidencias y argumentos que se han presentado en su contra. Los autores se proponen revocar el "derecho de fuerza" que de facto, pero ilegítimamente, ha ostentado el concepto de egoísmo, tanto en las ciencias sociales como en la biología.

La argumentación para cada caso se presenta en partes separadas, pues la diferencia lógica entre ambos conceptos no permite inferir del uno al otro. Se puede ser biológicamente altruista y psicológicamente egoísta, y viceversa (203-204). Además, las vicisitudes de ambos conceptos no son del todo paralelas. Respecto del altruismo biológico, su suerte está ligada a la teoría de la selección de grupos. Si se puede probar que la selección de grupos ha sido un factor importante en la evolución, entonces queda demostrada también la existencia del altruismo biológico. Los autores sostienen que la selección de grupos está suficientemente documentada en diversas especies, y que la resistencia de muchos biólogos a aceptar la evidencia es un hecho curioso e irracional de la historia reciente de la biología. Probablemente la idea de que los conceptos de egoísmo biológico y psicológico se apoyan mutuamente, junto con la suposición de que el egoísmo psicológico es una verdad evidente, hayan jugado algún

<sup>1</sup> Agradezco aquí a la *Fundación Alexander von Humboldt*, que financia el proyecto de investigación en cuyo marco se escribieron estas reseñas.

papel en este curioso fenómeno. En todo caso, dado que biólogos evolucionistas de la talla de G. C. Williams y W. D Hamilton, inicialmente fuertes críticos de la selección de grupos, han reconsiderado su actitud hacia ella y aceptado su papel en la evolución (49-59; 75-77), los autores del libro consideran que el caso en favor de la existencia del altruismo biológico está cerrado, a pesar de que muchos biólogos siguen presos de un prejuicio en su contra. Por tanto, presentan la primera parte de su obra como el reporte y la sustentación de una transición en plena marcha en el pensamiento evolucionista. Cuentan allí parte de una historia que debería interesar a los historiadores de la ciencia, en calidad de ejemplo de los problemas asociados al cambio científico. Sería en todo caso un error ver a los autores, por su defensa de la selección de grupos y del altruismo biológico, como “dos herejes solitarios clamando en el desierto” (7).

El caso del altruismo psicológico es distinto. Los autores sostienen que el debate con argumentos psicológicos y filosóficos en contra y a favor ha sido no concluyente. No se ha podido demostrar ni una ni otra posición. El derecho de fuerza del que ha gozado el concepto de egoísmo es también ilegítimo, pero no puede decirse que las evidencias estén, como en el caso anterior, en favor del altruismo. El debate está aún abierto y su definición exige luces nuevas, distintas a las proporcionadas hasta ahora por la filosofía y la psicología. Los autores tienen en este caso la ambición de mostrar que una

perspectiva evolucionista sobre el problema aporta evidencias importantes e inéditas en favor del altruismo psicológico.

Los primeros cinco capítulos del libro están, pues, dedicados a defender la idea de la selección de grupos, y a ponerla en relación con la evolución humana. Se trata de una idea con tradición, tanto en el pensamiento evolucionista, como en el de las ciencias sociales, especialmente ligada a la idea del grupo como superorganismo, como una unidad de partes funcionalmente adaptadas unas a otras y a su entorno. Bajo esta forma se abusó ampliamente de ella, al menos hasta que G. C. Williams le propinara un golpe mortal en 1966 con su libro ya clásico: *Adaptation and Natural Selection*. Williams hizo estudios postdoctorales en Chicago en los años 50s, cuando estaban allí los defensores más prominentes de la idea de la selección de grupos y del ecosistema como un gran superorganismo. “Si esto es biología evolucionista —pensaba entonces Williams—, prefiero dedicarme a vender seguros” (36). Muchos biólogos prominentes vieron en la obra de Williams la refutación definitiva, y durante muchos años la selección de grupos se consideró una teoría desechada, como la del flogisto en física. Los autores no quieren revivir los abusos, y están lejos de querer impulsar el misticismo de la superarmonía del universo o del ecosistema. Es cierto que la idea de la selección de grupos está asociada a la evolución de grupos como unidades en armonía funcional y adaptativa al entorno, exigien-

do en su límite ideal la eliminación del conflicto interno y el altruismo perfecto entre sus miembros. Pero esto no elimina la idea de la competencia y del conflicto, propia de la biología evolucionista. La selección de grupos puede darse lejos de su límite ideal en un proceso paralelo a la selección individual, y el altruismo biológico promovido por ella no excluye un conflicto entre grupos, que puede ser tan descarnado como el de la selección individual (9).

Sin embargo, no deja de ser cierto que, al menos inicialmente, el altruismo biológico fue visto, incluso por el mismo Darwin, como un enigma y un problema —si no un contraejemplo— para la teoría de la evolución por selección natural y su base en la lucha individual por la supervivencia del más apto. Si un individuo tiene un diseño genético que en un entorno determinado se expresa en un comportamiento altruista (el comportamiento beneficia a otros, sin importar sus motivos, puede incluso carecer totalmente de ellos), tendría que ser víctima tarde o temprano de la selección natural. Esto es inevitable, si ese individuo tiene que competir con otros que no son altruistas. Los egoístas se beneficiarán del altruista y tendrán un mayor éxito reproductivo que éste, pues el altruismo exige por definición un costo. Aunque inicialmente hubiese una mayoría altruista, esta disminuiría paulatinamente hasta desaparecer. Sin embargo, hay ejemplos claros de altruismo en todas las especies: el altruismo biológico existe. ¿Cómo puede ser esto posible, si la selección natural es un hecho?

Darwin sugirió, poniendo como ejemplos las castas estériles en los insectos sociales y la moralidad en los humanos, que la explicación estaba en que los rasgos altruistas, si bien disminuyen la aptitud del altruista, benefician al grupo al cual pertenece. Aquí está el origen de la idea de la selección de grupos. Pues explicar la existencia de rasgos altruistas por el beneficio que confieren a los grupos, supone postular un proceso de selección natural actuando sobre grupos y no sobre individuos. Si un grupo incluye una proporción grande de altruistas, su éxito reproductivo como grupo supera a otros que no incluyen altruistas. Dentro del grupo, los altruistas tienen menos éxito reproductivo que los egoístas. Pero ellos contribuyen a que su grupo crezca más rápidamente que los grupos sin altruistas. Si el grupo que contiene a los altruistas termina por imponerse en la lucha por la supervivencia, la proporción de altruistas en la población total aumenta. De este modo, si las condiciones evolutivas son tales que la competencia entre grupos es lo suficientemente acentuada como para neutralizar y superar la competencia al interior del grupo, entonces es posible que el beneficio que el altruismo confiere a los grupos explique su existencia en diversos organismos actuales. Por otro lado, si el proceso de selección natural se centra en la competencia entre individuos, los altruistas deben desaparecer (19-27).

La idea es pues simple y clara, y así la explican los autores, quienes la expresan también por medio de un

modelo matemático, consistente con los que se usan en la teoría de la genética de poblaciones (19-21). Dada la consistencia del modelo, sólo quedan dos estrategias a los críticos del mismo. La primera es que la consistencia nada dice acerca de su realidad empírica. Para que pueda haber evolución de rasgos altruistas, la competencia entre grupos debe tener una fuerza mayor que la de la competencia al interior de los grupos, donde los altruistas siempre pierden. Las condiciones para que esto sea posible son raras y no se dan en la naturaleza. En cierto modo, ese fue el núcleo de la crítica de Williams. No quiso mostrar que la selección de grupos fuese evolutivamente imposible. Más bien la consideró una posibilidad remota, ligada a condiciones empíricamente improbables. Diseñó además un test que debía servir para confirmar o desconfirmar la existencia de procesos de selección de grupos, basándose en las predicciones divergentes que la selección de grupos y de individuos proyecta sobre las proporciones de los sexos en las especies sexuadas, y comparando estas predicciones con los datos empíricos disponibles (36-40).

En la década de los 30s, los fundadores de la genética de poblaciones repararon en que la capacidad de engendrar una proporción mayor de hembras tiene efectos positivos sobre el crecimiento de la población, o la aptitud del grupo. Ello se debe que el crecimiento de la población está limitado por el volumen de óvulos, no por el volumen de esperma, supuesto que una

proporción menor de varones puede fertilizar a todas las hembras. Por lo tanto, si un individuo engendra una mayor proporción de hembras entre sus descendientes (se supone que su capacidad para ello está controlada por un gen o grupo de genes y es por lo tanto heredable, aunque esto no excluye que el gen requiera de un entorno específico), está contribuyendo a un crecimiento acelerado del grupo del que forma parte. Pero el efecto de esa capacidad sobre su propio éxito reproductivo individual es negativo. Para medirlo, supongamos una población en donde los individuos no difieren en su éxito reproductivo por ningún concepto que no sea derivable de la mera capacidad de engendrar una proporción distinta de sexos. Tomemos a dos individuos de esta población, A y E. Cada uno engendra 10 descendientes, que suponemos ser para esta población el número máximo para las hembras, así como para los varones cuando la proporción de los sexos es 1:1 y ningún varón se queda sin aparear. El individuo de tipo A engendra 9 hembras y 1 varón, mientras que el de tipo E engendra 5 hembras y 5 varones. Suponiendo que sus descendientes se aparean entre sí, y que ninguno se queda sin aparear (recuérdese que no difieren en su éxito reproductivo por ningún concepto que no sea derivable de la mera proporción de los sexos), las 14 hembras de la segunda generación no podrán superar su número máximo, 10 cada una, mientras que los seis varones fertilizarán a las 14 hembras y tendrán en promedio  $140/5 = 28$  descendientes. Si evalua-

mos ahora el éxito reproductivo de los individuos originales A y E por el número de sus nietos, vemos que A tiene  $9(10) + 1(23) = 113$  nietos, mientras que E tiene  $5(10) + 5(23) = 165$  nietos. Estas cifras permiten afirmar que engendrar una mayor proporción de hembras es un rasgo altruista: incrementa el éxito reproductivo de otros miembros del grupo a costa del éxito reproductivo del individuo con ese rasgo. Fisher había mostrado que, por esta razón, y suponiendo que la selección natural opera sobre individuos, la proporción de los sexos tiende siempre a la paridad. Los individuos que engendran mayor proporción de hembras son eliminados por la selección individual. Williams utilizó esta idea para elaborar un argumento contra la selección de grupos. Señaló que, si la selección de grupos fuese un hecho, entonces debían encontrarse especies con proporción mayoritaria de hembras, o especies en las que se hubiese observado la capacidad de variar la proporción de los sexos, aumentando por ejemplo el número de hembras cuando al grupo le conviene aumentar su productividad. De existir procesos de selección de grupos, la proporción de los sexos no debe mantenerse constante cerca de la paridad, como lo predice la selección de individuos, sino que debe variar según si el grupo está por encima o por debajo de su número óptimo. Con esta predicción divergente entre la selección individual y la de grupos, Williams hizo una evaluación de los datos científicos disponibles sobre las proporciones de los sexos y concluyó que ellos refuta-

ban la hipótesis de la selección de grupos.

Los datos de Williams no habían sido completos. Apenas un año después (1967) W. D. Hamilton, que ya era famoso por su teoría de la selección de familias (*kin selection*), publicó un artículo en donde ofrecía una explicación para una serie de ejemplos de especies que presentaban proporciones de sexos con mayoría de hembras. Parte del curioso destino científico de la selección de grupos, es que prácticamente nadie dedujo que, según el test diseñado por Williams, la mera existencia de esos datos era ya una prueba en favor de la selección de grupos. El propio Hamilton no dejó de mencionar, aunque sólo en una nota, que la proporción de sexos en los ejemplos presentados por él era un resultante de dos fuerzas distintas y opuestas en la evolución: la selección de individuos al interior de grupos, y la selección entre grupos que exhiben distintas proporciones (42). Pero nadie parece haber reparado en esta observación.

Sucede que Hamilton ya había publicado tres años antes (1964) su teoría de selección de familias, utilizando conceptos y estrategias que desde entonces han sido entendidas como alternativas teóricas a la selección de grupos para explicar casos de "aparente" altruismo biológico. El mismo Hamilton había creído estar ofreciendo una explicación alternativa y distinta. Pero la verdad de las cosas, según Sober y Wilson, es que sólo ofreció una perspectiva distinta para mirar el mismo proceso al que alude la teoría de la se-

lección de grupos (57). La explicación de Hamilton del altruismo implicaba, por un lado, asumir la perspectiva de los genes. Si el comportamiento altruista beneficia a parientes cercanos, hermanos por ejemplo, hay una probabilidad de un 50% (el grado de parentesco) de beneficiar a un individuo que tiene el mismo gen que determina el comportamiento altruista (la determinación genética es válida en un rango específico de entornos, es decir, no excluye que el entorno sea también en factor causal importante). Por otro lado, Hamilton introdujo el concepto de aptitud inclusiva, que incluye no sólo el número promedio de los descendientes de un individuo, sino también el de sus parientes cercanos valorado por su grado de parentesco. Esto permite formular el proceso de selección como enfocado sobre la aptitud inclusiva individual, y transmite la impresión de que apelar a la selección de grupos es superfluo. Tanto la perspectiva del gen (que después dio origen a la celebre formula del "gen egoísta" de Dawkins), como el concepto de aptitud inclusiva, permiten reinterpretar lo que desde la selección de grupos parece un rasgo altruista como un rasgo biológicamente egoísta, que beneficia al fin de cuentas al individuo que lo posee (65-67).

Sober y Wilson explican que, en la medida en que las teorías de la selección de familias y del "gen egoísta" creen ofrecer explicaciones alternativas y distintas de la selección de grupos, son víctimas de lo que ellos llaman la "falacia promedial" (*averaging fallacy*).

cy). Ésta consiste en desconocer la presencia de grupos en el proceso selectivo, y calcular la aptitud tanto de los altruistas como de los egoístas promediando las aptitudes de esos rasgos a través de las fronteras de los grupos. En el modelo de Hamilton, la aptitud inclusiva que permite explicar la evolución de comportamientos altruistas, es en realidad la aptitud del altruista en un grupo de alta concentración de altruistas, el grupo representado por los críos de una misma camada. El propio Williams había publicado en 1957 un modelo de selección de grupos, utilizando camadas de críos que actúan entre sí durante una parte de su ciclo vital (62-64, 67).

La falacia promedial desconoce que la elevada aptitud relativa del altruista en una población sólo puede ser producto de la estructura de grupos: la aptitud de los altruistas es siempre menor que la de los egoístas *que están en su mismo grupo*, pero la aptitud de los altruistas en grupos de alta concentración de altruistas es mucho mayor que la de los egoístas en grupos de alta concentración de egoístas. La diferencia puede ser tal que, si se promedian las aptitudes que estos rasgos presentan en grupos de distinta composición, la de los altruistas puede resultar mayor que la de los egoístas. Para los biólogos que sucumben a esta falacia, esto significa que el altruismo no es sino aparente: según la definición del altruismo biológico, el rasgo debe disminuir la aptitud del que lo exhibe en relación con los que no lo exhiben. Pero los números no mienten, y ellos mues-

tran que la aptitud del pretendido altruista es mayor. Luego no hay tal altruismo biológico. La crítica de Sober y Wilson a este argumento es convincente: basarse en un promedio es simplemente desconocer el proceso por el cual el rasgo altruista tiene una aptitud neta mayor en la población. Para explicar este resultado neto, hay que atender a la estructura de grupos. Hay que notar que, dentro de los grupos, la aptitud del altruista siempre es menor que la del egoísta. Sólo cuando los grupos exhiben grandes diferencias en sus proporciones de altruistas y egoístas, especialmente cuando hay extremos de alta concentración de egoístas y de alta concentración de altruistas en grupos distintos, puede darse que el promedio de la aptitud de los altruistas supere a la de los egoístas. Pero esto es un fenómeno pasajero, pues con el paso del tiempo los altruistas desaparecen en todos los grupos. Para que sobrevivan, es necesario que cada cierto tiempo se vuelvan a reconstituir grupos con alta concentración de altruistas (31-34). La teoría de Hamilton, conocida como *kin selection*, que apela a un proceso de selección de familias, es en realidad un modelo de selección de grupos, en donde la estructura de grupos no se hace aparente. Los comportamientos que se seleccionan por beneficiar a los miembros de una misma camada de críos, suponen una estructura de grupos, pues las camadas conforman grupos de individuos que actúan entre sí y están aislados durante un tiempo de las otras camadas en una población.

Los autores sostienen que la fala-

cia promedial está también detrás de las explicaciones del altruismo por medio de la teoría de juegos (79-86). Ninguna de estas explicaciones es verdaderamente alternativa a la selección de grupos. Todas apelan de manera no aparente a procesos de selección de grupos y los miran bajo otra perspectiva. La teoría de juegos presenta sus modelos sin el concepto de grupo. Asumiendo, no obstante, una concepción minimalista de lo que es un grupo (un conjunto de individuos que afectan mutuamente su aptitud con respecto a un rasgo), el estudio de interacciones entre individuos propio de la teoría de juegos puede verse como implicando grupos, conformados por los individuos que interactúan en cada caso. El hecho de que los grupos sólo duren lo que dura la interacción es irrelevante. Los autores dedican también una sección a exponer y defender su concepción minimalista de lo que es un grupo (92-98).

Los dos capítulos finales de la primera parte argumentan que la selección de grupos ha jugado un papel importante en la evolución humana. Las sociedades humanas muestran un nivel elevado de socialidad, comparable a la ultrasocialidad en los insectos sociales. Algunos biólogos evolucionistas, como Alexander, rechazan esta comparación. Su argumento es que el alto grado de parentesco genético propio de los insectos sociales no se da en las sociedades humanas. En ello presupone que sólo un alto grado de parentesco genético es capaz de explicar el "aparente" altruismo de la socialidad. Sober y



Wilson arguyen que el parentesco puede ser substituido por otros mecanismos en los humanos. Ya Hamilton había señalado en 1975, respaldando explícitamente la idea de selección de grupos –lo cual sorprendentemente ha pasado desapercibido en la historia de la controversia–, que “no importa si los altruistas se agrupan con altruistas porque están genéticamente relacionados, o porque se reconocen entre sí, o porque el gen altruista determina por un efecto pleiotrópico una preferencia hacia un hábitat determinado” (77). (Los autores relatan el episódico intercambio intelectual entre Hamilton y la figura romántica y trágica de George Price, que contribuyó al cambio de actitud de aquél hacia la selección de grupos, y a su aceptación de la selección natural actuando simultáneamente al nivel del individuo y al nivel del grupo. Price, sin ninguna formación biológica, ideó una ecuación matemática para modelar el proceso de selección a múltiples niveles. Abandonaría poco después sus estudios evolucionistas para dedicarse a la interpretación del Nuevo Testamento y a ayudar a los *clochards* en Londres. Finalmente se suicidó en un edificio abandonado, donde había vivido algún tiempo como *clochard*, después de haberse despojado de todas sus pertenencias).

Después de describir inicialmente un mecanismo conocido como interacción selectiva, el cual supone que los organismos son capaces de reconocer cognitivamente el altruismo de sus congéneres y de asociarse selectivamente, con independencia del parentesco, los

autores se concentran sobre las normas sociales y las recompensas y castigos que las respaldan. Una norma social asociada a recompensas y castigos es capaz de reducir los costos naturales de comportamientos altruistas necesarios para que el grupo funcione como unidad adaptativa (por ejemplo, las actividades de la caza y la defensa, propias de los grupos de cazadores y recolectores, presuntamente la forma de organización que sirvió como escenario a la evolución de la humanidad). Si no fuera por el costo que las normas sancionadas imponen a los comportamientos egoístas dentro del grupo, éstos terminarían por desplazar evolutivamente a los altruistas y la socialidad difícilmente se extendería más allá de la familia. Los comportamientos altruistas con las normas que los sancionan conforman así un paquete que, siendo altruista y en favor del grupo, tiene un costo relativo menor que el de los rasgos altruistas sin sanción. Además aumentan el costo del egoísmo dentro del grupo, de manera que la selección individual de los egoístas dentro del grupo no pueda contrarrestar fácilmente la selección intergrupala. La selección de grupos puede no ser suficientemente fuerte para explicar el comportamiento altruista sin sanción en sociedades numerosas, pero puede ser suficiente para explicar la evolución conjunta del comportamiento altruista y de la norma sancionada que lo prescribe (142-152).

Los autores consideran que el mecanismo de la norma sancionada para beneficio del grupo constituye un mo-



delo posible para explicar cómo grupos grandes de individuos sin parentesco pueden evolucionar como unidades. Estudiando además una muestra al azar de 25 culturas, a partir de las etnografías contenidas en la base de datos antropológicos conocida como Archivo del Área de las Relaciones Humanas, concluyen que todas las culturas se muestran no sólo fuertemente reguladas por normas, sino que éstas tienen las características predecibles si su función ha de consistir en beneficiar al grupo como unidad. "La mayoría de las sociedades tradicionales parece diseñadas para suprimir procesos intragrupal que son disfuncionales para el grupo como unidad, y como resultado, la selección natural ha operado al nivel del grupo y ha producido adaptaciones a ese nivel" (192). El modelo, sin embargo, está abierto en cuanto a qué mecanismos psicológicos y culturales son los que promueven el comportamiento altruista (156). Los datos antropológicos sugieren una diversidad de mecanismos culturales, de los cuales los autores mencionan sólo dos: la educación para la internalización de las normas, y la vigilancia y escarnio público (156, 182). Los autores parecen asociar a estos mecanismos culturales, respectivamente, el motivo psicológico de un concernimiento directo con el bienestar de otros (altruismo psicológico) y el motivo egoísta de evitar la sanción externa. El altruismo biológico atañe tan sólo al comportamiento conforme a la norma, no al mecanismo psicológico responsable de su producción, que puede ser altruista

o no. Con esto queda introducida la segunda parte del libro: el debate en torno a la realidad del altruismo psicológico.

Los autores se proponen exponer aquí momentos destacados del debate en torno al altruismo, tanto en la psicología como en la filosofía, para sustentar su tesis de que los argumentos de uno y otro lado no han podido decidir la cuestión, antes de proponer su solución apoyada en la teoría de la evolución. Su exposición es valiosa por su riqueza y claridad, comenzando por precisar exactamente el sentido en que se enfrentan la tesis del altruismo y la del egoísmo. Quien defienda la realidad del egoísmo psicológico, está comprometido con la negación total de la motivación altruista: en ningún caso ningún ser humano desea el bienestar de otro como fin último. Si lo desea, es sólo de manera instrumental, como parte de o medio para su propio bienestar (225-226). El altruismo, en cambio, no está comprometido con negar de plano la motivación egoísta. Afirma que, al menos a veces, los seres humanos deseamos el bienestar de otros como un fin último. La tesis altruista es pues compatible con un pluralismo motivacional. Por otra parte, la tesis tampoco afirma que la motivación altruista, para valer como tal, debe actuar sola, sin concurso de una motivación egoísta. No excluye entonces que ayudemos porque el bienestar de otro nos importa en sí mismo y *al mismo tiempo porque nos place* ayudar (246). Tampoco implica nada respecto de la fuerza de la motivación altruista-

ta. El pluralismo defendido puede en principio incluir la tesis de que, en caso de un conflicto entre el deseo altruista y el egoísta, el egoísta es el más fuerte (245-247).

La cuidadosa aclaración de la tesis altruista lleva a los autores a puntualizar también sus relaciones con la moralidad. A diferencia de los tratamientos usuales del tema, a los autores les parece importante señalar las diferencias entre el altruismo y la moralidad. Es posible adherirse a concepciones de la moral en donde el principio supremo no es, ni implica, el deseo altruista, el bienestar de otro por sí mismo. Puede ser el respeto a la libertad del individuo, por encima de su bienestar, o puede ser una moral de respeto a los mandatos divinos, independientes de si conllevan o no al bienestar de la humanidad. Por otro lado, acciones motivadas por el altruismo pueden ser inmorales, por ejemplo, cuando actuamos en favor de alguien desconociendo los derechos de otras personas. Es obvio que habría que cualificar el tipo de moral para la cual el altruismo es clave, y también quizás redefinir la tesis altruista para evitar conflictos del segundo tipo. Sin embargo, pienso que lo dicho por los autores aquí no disminuye la importancia de la tesis del altruismo para la moral de tipo naturalista, y quizás tampoco para las concepciones morales de tipo kantiano, si no es implausible sostener que el respeto a la libertad de la persona se inserta en una concepción integral y jerarquizada de lo que implica su bienestar.

Otra aclaración importante consiste en la relación del altruismo con la emoción conocida como empatía. Se trata de una emoción que desempeña un papel central en las reflexiones contemporáneas sobre la moralidad desde la psicología. No faltan quienes piensan incluso que su mera existencia habla en favor del realismo moral. Pero una aclaración cuidadosa de estas suposiciones es más bien inusual. Los autores señalan tres elementos indispensables propios de la empatía. La empatía supone, en primer lugar, que se comparte la emoción de otra persona, se coincide emocionalmente con ella (233-234). En segundo lugar, que la coincidencia emocional tiene un componente cognitivo: exige que haya una representación de la otra persona en un determinado estado emocional (236). En tercer lugar, la coincidencia emocional exige la presencia de un elemento que la distingue de la aflicción personal. En la aflicción, yo coincido con el estado emocional de otra persona, pero mi atención no se enfoca sobre la otra persona, sino sobre mí mismo, faltando entonces ese elemento de la empatía, que es sentirse bien (o mal) *por razón de (feeling for)* otra persona (234-5). A pesar de resaltar correctamente este elemento, los autores sostienen que, si bien la empatía generalmente causa el deseo de ayudar a otros, no está claro que este deseo sea verdaderamente altruista (232). Se podría tratar de un deseo instrumental, por ejemplo, ayudar a otros como medio para eliminar el sentimiento de aversión que la empatía implica cuando

empatizamos con una persona a quien le sucede algo malo. En este caso, el deseo no tendría el bienestar de otro como fin último. Pero si al empatía produjese un deseo de este tipo, parece natural dudar que se trate de verdadera empatía, como la definen los autores, y no de simple aflicción personal. Creo que no es injusto notar que, en este punto, el análisis de los autores no es acertado. Si lo que causa el deseo de ayudar a otro es empatía, como ellos la definen, entonces implica constitutivamente un deseo altruista. Los autores se acercan a este análisis cuando tratan de explicar qué es sentirse mal por alguien que nos importa. Señalan que se trata de un sentimiento de aversión, y que como tal nos gustaría librarnos de él; pero dan a entender que si se trata de verdadera empatía, entonces, necesariamente, sólo aceptaremos aliviar el sentimiento de aversión por un cambio real en la situación de la persona y su bienestar (235). Esto implica que en la empatía hay, al menos *prima facie*, un deseo del bienestar ajeno por sí mismo.

¿Significa esto entonces que la mera existencia de la empatía prueba la existencia del altruismo? Tal y como la entienden y analizan los autores, creo que la respuesta debe ser positiva, en contra de lo que ellos explícitamente afirman (232). Sin embargo, esto no resuelve la controversia, sino que más bien la desplaza a la pregunta: ¿sentimos los humanos verdaderamente empatía así entendida, o sentimos algo veladamente egoísta, que la introspección nos presenta de manera deceptiva

como lo que usualmente entendemos por empatía, y por tanto como altruismo? La verdadera duda es, pues, sobre el carácter de los deseos *constitutivos* de lo que llamamos empatía. Esto es importante para la manera como en el libro se presentan y evalúan los experimentos psicológicos con los que Batson (1991) ha querido demostrar empíricamente la realidad del altruismo. Los autores piensan que estos experimentos no han podido definir el debate. Según ellos, aunque los experimentos pueden mostrar que la empatía hace una diferencia en el comportamiento, no pueden mostrar que la empatía genere deseos verdaderamente altruistas (266, 272). Pero esta manera de plantear la posición escéptica parece inadecuada. La pregunta es más bien si los seres humanos sentimos realmente empatía, entendida como constituida por deseos altruistas. Los experimentos de Batson se pueden entender fácilmente de este modo. Están diseñados para tratar de dar pruebas en contra de diversas hipótesis egoístas. Podemos constatar, si miramos de cerca a estas hipótesis, que todas dan una lectura alternativa, no-altruista, del sentimiento empático y del comportamiento suscitado por él, y que, de ser ciertas, el elemento de la empatía que los autores destacan con la fórmula: 'sentirse bien o mal *por razón de* otra persona', no encontraría aplicación. Veamos estas distintas hipótesis.

1. Hipótesis egoísta de la reducción del sentimiento de aversión: sostiene que el deseo de ayudar se explica porque este sentimiento, producido por la

coincidencia empática con la necesidad ajena, nos incomoda y queremos eliminarlo (261). Ayudar al otro es un medio para restablecer nuestro confort psicológico. El deseo está dirigido a uno mismo y no al otro. A la vista de la necesidad ajena nos sentimos mal, pero no por razón del otro, sino por nosotros mismos.

2. Hipótesis egoísta de evitar la autocensura: el deseo de ayudar se explica porque queremos evitar la autocensura (264-5) (por lo tanto, se trata de un sentimiento desagradable respecto de uno mismo). Otra vez se observa que el deseo está enfocado sobre el propio bienestar psicológico. El elemento clave de la empatía, que es sentirse mal por razón del otro, está aquí también ausente.

3. Hipótesis egoísta del sentimiento de recompensa: sostiene que el deseo de ayudar se explica por la recompensa emocional de sentirnos bien con nosotros mismos (266). De nuevo el sentimiento que se pone como base del comportamiento altruista está dirigido al propio bienestar y no al del otro.

4. Hipótesis egoísta del sentimiento compartido: sostiene que el sentimiento de recompensa no tiene que venir de ser uno mismo el que ayuda, sino de compartir representativamente el bienestar de otro (267). No nos interesa el bien del otro por sí mismo, sino el sentimiento agradable que derivamos de saber que está bien.

5. Las hipótesis egoístas se pueden volver, como se observa, cada vez más sutiles y complejas. La estrategia de Batson y sus colegas fue diseñar situa-

ciones experimentales en donde las hipótesis rivales del egoísmo y del altruismo predijesen comportamientos diferentes, y luego realizar los experimentos para obtener datos con los cuales contrastar las hipótesis. Los diseños experimentales son en sí mismos interesantes y están explicados con claridad. Sin embargo, los experimentos sólo pueden refutar las hipótesis egoístas una por una, y siempre parece haber una hipótesis egoísta ulterior y más refinada para reemplazar a las refutadas (271-272). La conclusión negativa de los autores parece razonable, pero en todo caso me interesa destacar que las diversas hipótesis egoístas que Batson y sus colaboradores han querido refutar, no se proponen otra cosa que cuestionar la naturaleza presuntamente altruista del sentimiento empático, de manera que se elimine de él el elemento altruista que los autores, al parecer sin tener una conciencia clara de ello, incluyen en su definición (el sentirse mal o bien *por razón* del otro).

De la exposición del debate psicológico en pro y en contra del altruismo, puede observarse que la hipótesis egoísta plantea dificultades al altruismo, precisamente cuando postula entre sus fines últimos estados internos placenteros (o de ausencia de dolor) más o menos sofisticados. Al definir el hedonismo, los autores señalaron que el placer y la ausencia de dolor es postulado por él como el *único* fin último, y que ello incluye a los placeres sofisticados, por ejemplo, el placer de saber que a alguien le va bien, o mal (224). Quizás esto explica que los au-

tores digan luego que el hedonismo es el competidor principal al que se enfrenta el altruismo (297). Esta observación es imprecisa. Más adelante se verá que la refutación del hedonismo no es suficiente para establecer la verdad del altruismo. Lo que es cierto, y quizás ha sido la intención de los autores, es que, lo que causa dificultad a la hipótesis altruista, es un tipo específico de deseos hedonistas, dirigidos a placeres sofisticados de carácter psicológico. Pero los autores ya definieron el egoísmo como incluyendo entre sus fines últimos este tipo de placeres sofisticados (226). Son ellos, precisamente, los que permiten explicar el deseo de ayudar, o el deseo de que otros estén bien, como medio adecuado de alcanzar un estado interno confortable, ofreciendo así una alternativa egoísta a la explicación altruista.

Hasta ahora, la evidencia experimental producida por los psicólogos no ha sido capaz de inclinar la balanza en favor de una de las dos hipótesis rivales. Es tentador afirmar que la cuestión no es decidible por evidencia comportamental, pero los autores no quieren llegar tan lejos (272); quizás algún día logren los psicólogos encontrar evidencia decisiva. Volcándose luego sobre los experimentos mentales y demás argumentos de los filósofos, intentan mostrar que tampoco ellos logran cerrar el debate en uno u otro sentido. Entre los filósofos han predominado los argumentos que intentan mostrar que el hedonismo o el egoísmo psicológico son falsos. Los autores exponen y evalúan una serie de argumentos, que

incluyen el argumento de Butler contra el hedonismo, el experimento mental de Nozick contra la misma doctrina, el argumento de Nagel que intenta mostrar la realidad del altruismo, y diversos argumentos basados en el peso de la prueba, en la infalsabilidad de las hipótesis y en la parsimonia o simplicidad de las mismas. La exposición y evaluación de todos estos argumentos es cuidadosa, sofisticada y presta atención a muchas de las complejidades del tema; merece la atención de todos los interesados en el mismo. Me limito aquí a una observación de carácter general y a comentar dos argumentos, el de Butler con cierto detalle, y el de Nagel con el propósito específico de remover un malentendido interpretativo.

La observación de carácter general atañe a los argumentos antihedonistas de Butler y Nozick. Los autores consideran que estos argumentos no logran su objetivo de refutar el hedonismo. Pero omiten señalar que los argumentos no pueden apoyar la tesis altruista. Pues están diseñados para refutar el hedonismo y establecer la tesis general contraria, de que entre nuestros fines últimos debemos incluir un interés irreductible por estados de cosas externos, distintos del placer mismo, que no es sino un estado interno de la conciencia. De ser exitosos, no lograrían demostrar la tesis más específica de que, entre los estados externos a la conciencia que constituyen fines últimos, debemos incluir el bienestar objetivo de otras personas. Podría tratarse en principio sólo de estados exter-

nos de cosas relacionados con nuestro propio bienestar. Como ya se mencionó dos párrafos más arriba, la refutación del hedonismo no implica el establecimiento de la tesis altruista. Si acaso los argumentos contra el hedonismo son en realidad más fuertes de lo que los autores los hacen aparecer, no se habría todavía avanzado nada en dirección a decidir el debate entre altruismo y egoísmo.

Butler quiso refutar la idea hedonista de que los deseos están dirigidos directamente al placer producido por su satisfacción. Sostuvo que el placer supone una adecuación del objeto deseado al apetito, es decir, que es el objeto deseado el verdadero fin del deseo y no el placer (276). Butler parece haber querido defender que esta es la única relación existente entre el placer y el deseo, y en este caso es fácil señalar contraejemplos. Hay placeres que, como las sensaciones placenteras que afectan nuestros sentidos, nos pueden venir de manera inesperada, y no suponen un deseo previo dirigido a un objeto. Por otro lado, a veces deseamos el placer por sí mismo y sólo indirectamente un objeto externo como medio. Pero esta objeción no es decisiva. Su argumento tendría igualmente éxito, con sólo que mostrase que hay otros fines últimos además del placer. Esto es lo que parece querer decir, cuando sostiene que el sentimiento de placer supone que se desea algo distinto del placer mismo. El argumento ha sido influyente, y los autores citan pasajes de C. D. Broad y de Joel Feinberg que se inspiran en él. Pero no que-

dan convencidos, y sostienen que el argumento de Butler confunde el *sentimiento* de placer que sigue a la satisfacción de un deseo, con el *deseo* de placer que en cualquier caso puede precederlo. Es cierto que el placer que sucede a la satisfacción del deseo de comer, supone que se desea algo distinto al placer, a saber, la comida. Pero esto, recalcan los autores, nada dice en contra de la hipótesis hedonista de que el deseo de placer precede al deseo de comida, y lo explica por derivación (278). El hedonismo no es la tesis monomaniaca que dice que lo único que se desea es el placer. Lo que dice es que ese deseo es el único fin último que explica a los demás deseos; y esto es compatible con tener una multiplicidad de deseos instrumentales, entre ellos, por ejemplo, el deseo de comer (280). Obviamente los autores no pretenden con esto defender el hedonismo. Tan sólo afirman que nada de lo que dice Butler, o quienes han utilizado argumentos Butlerianos, logra refutarlo.

El pasaje que los autores citan de Broad parece anticipar y bloquear la estrategia de los autores contra Butler, que consiste en sostener, por ejemplo, que deseo el poder porque me da placer y que mi deseo último de placer explica mi deseo de poder. Broad dice respecto de esto: "este tipo de placer presupone un deseo de algo distinto al placer, y por tanto este deseo no puede derivarse del deseo de este tipo de placer" (277). Los autores no comentan esto en detalle. El argumento parece correcto, pero sólo porque Broad pre-

supone aquí que la derivación hedonista debe partir de un tipo específico de placer (el placer que da el poder, por ejemplo), que no puede existir sin que previamente se desee el poder. Aunque sea correcto afirmar que el deseo de poder no se puede derivar de manera hedonista del placer específico que da el poder, una derivación hedonista no tiene que partir del deseo dirigido a placeres específicos, sino que puede hacerlo del deseo del placer en general.

Hay, sin embargo, una línea de argumentación que se apoya en Butler y que fue explotada por Hume, en un pasaje que los autores mencionan pero no presentan (358, n.15). Se trata de la tesis de que los deseos dirigidos al placer son de segundo orden o reflexivos, y que nuestros deseos de primer orden están dirigidos a objetos externos. La idea es que el sujeto adquiere deseos dirigidos al placer en un proceso que supone observación reflexiva y deliberación. El sujeto tiene primero que experimentar sus deseos y apetitos dirigidos a cosas externas, y el placer que produce su satisfacción. Sólo por medio de esta experiencia reflexiva puede adquirir deseos dirigidos al placer y las creencias pertinentes sobre los modos de asegurárselo. No hay duda de que esta explicación del origen del deseo del placer en un proceso reflexivo que se apoya en los deseos primarios, tiene un atractivo que no parece tener la teoría hedonista contraria, según la cual el deseo de placer y de evitar el dolor es primario, y los demás deseos se adquirirían en un proceso reflexivo

y deliberativo acerca de las cosas que nos podrían producir placer. Quizás lo que quiso decir Broad, en el pasaje mencionado arriba, es que es implausible que los deseos que todos observamos en nosotros se originen en un proceso reflexivo como el que supone el hedonismo. La subjetividad tendría que partir de un único deseo de placer, y adquirir, acaso en un proceso de observación reflexiva de actividades ciegamente perseguidas, creencias sobre cuáles le producen placer y cuáles no. Como explicación de la forma como adquirimos nuestros deseos, esto no tiene ninguna plausibilidad. No creo que nadie proponga seriamente una explicación de este tipo como válida para el hambre, como deseo de comer. La afirmación Butleriana de que el placer supone deseos dirigidos a objetos externos y no al placer mismo, se interpreta entonces como una tesis sobre la plausibilidad relativa de dos explicaciones contrarias sobre el origen de nuestros deseos. E incluso un compromiso entre estas teorías contrarias, que concedería que el deseo de obtener placer y de evitar dolor es primario e independiente de procesos reflexivos cuando se trata exclusivamente del placer y dolor físico, sería un triunfo contra el hedonismo, pues se habría de todos modos abandonado la pretensión hedonista de explicar todos los deseos derivándolos del deseo de placer.

El hedonista podría replicar que el concepto de deseos primarios no coincide con el de deseos últimos, y que no importa qué deseos primarios tengamos por naturaleza, ya que siempre



podemos proponernos el placer como fin último. Esto puede concederse, pero nos sitúa fuera del hedonismo como doctrina sobre nuestra naturaleza psicológica. Quizás nuestra psique sea flexible, y podamos decidir qué deseos queremos como últimos y cuáles no, y de hecho tendría que ser así, si el hedonismo como doctrina normativa ha de tener sentido. Pero si la teoría Butler–Hume sobre el origen de los deseos es correcta, entonces el placer no nos ha sido dado por naturaleza como único fin último, aunque podamos reflexivamente convertirlo en tal.

El argumento de Nagel en favor del altruismo (*The Possibility of Altruism*, 1970) merece un comentario más extenso del que puede recibir aquí, e incluso del que recibe en el libro. Nagel se propuso demostrar que, si una persona *p* tiene una razón para una acción determinada, entonces toda persona tiene una razón para querer que *p* haga esa acción y para promoverla o no impedirle. En otras palabras, Nagel quiso mostrar que las razones son objetivas, en el sentido de ser válidas a través de las fronteras entre las personas. Su argumento no es fácil de seguir. Hace ya algunos años, Nicholas Sturgeon publicó una reconstrucción y análisis del mismo con intención demoleadora (*Philosophical Review*, 1974). Lo comenzó, además, con un veredicto lapidario sobre su conclusión: “El argumento de Nagel es digno de atención, porque su conclusión es tan extravagantemente implausible” (1974, 375). Esta rotunda declaración es confirmada por Sober y Wilson (357, n.4),

quienes acaso por ello le dedican poco espacio al argumento mismo. Sostienen que cuando Nagel afirma que una persona que no reconoce la objetividad de las razones es irracional, parece emplear, o necesitar, un sentido específicamente *moral* del término: “puede ser verdad que los egoístas son irracionales, pero con ello sólo se dice de otro modo que su atención exclusiva a razones egocéntricas es moralmente errónea” (281). Pero esto es un juicio evaluativo, que no puede contribuir en nada a demostrar la realidad del altruismo psicológico, que es la tesis que interesa a Sober y Wilson (p.357, n.4). Su parco tratamiento de Nagel sería inobjetable, si hubiesen dado con el meollo de su argumento. Pero opino que este no es el caso. Nagel pretendía mostrar que el precio del desconocimiento de la objetividad de las razones es el solipsismo, y no la inmoralidad. Esta quizás ni siquiera sea un precio que haya que pagar en ningún caso. Pues Nagel reconocía formas de moralidad apoyadas en la benevolencia, o incluso en el miedo a la muerte, que no requieren de la objetividad de las razones en el sentido explicado arriba. En ellas, el reconocimiento que un individuo concede a los intereses de otro se apoya en una motivación egoísta, o en un deseo contingente de su bienestar. Aun así, Nagel no congeniaba con explicaciones psicológicas de la moral que simplemente terminasen postulando un deseo altruista. Pensaba más bien en una psicología metafísica, que explicaría en dónde radica la posibilidad de tener deseos altruistas; que hay en

la psicología humana elementos estructurales, y no meramente contingentes, que nos capacitan para interesarnos por el bienestar de otras personas, y que ellos son accesibles a una reflexión *a priori*; y pensaba haber encontrado uno de esos elementos claves en el hecho de que no somos solipsistas. Es justo reconocer, sin embargo, que su argumento carece de fluidez intuitiva y requiere de una reelaboración. Aquí tan sólo pretendo rebatir el veredicto de implausibilidad que pesa sobre su tesis, y abogar así por mantener vivo el interés en el mismo.

El veredicto de implausibilidad es curioso, porque a Nagel, en todo caso, el altruismo le parecía un fenómeno cotidiano, al menos en el sentido en el que él lo tomaba. “Por altruismo no entiendo solamente los comportamientos de noble autosacrificio... Entiendo cualquier comportamiento motivado por la sola creencia de que alguien obtiene por él un beneficio o evita un daño. Nuestras vidas están llenas de tales comportamientos. En su mayoría se trata de la consideración mundana que nada nos cuesta, y que no involucra ni autosacrificio ni nobleza —como cuando le decimos a alguien que tiene una llanta desinflada...” (PA, 16, n.1). El elemento esencial en su concepto del altruismo es la capacidad de ser motivado directamente por los intereses o necesidades de otra persona, sin intervención de un sentimiento como la benevolencia. O, para ser más precisos, Nagel sostuvo en su libro que sentimientos como la benevolencia se construyen y apoyan sobre un fenóme-

no más originario, precisamente la objetividad que concedemos a los intereses de otras personas. Pero igualmente importante para su concepción del altruismo es que hubiese dado ejemplos tan poco espectaculares del mismo. Esto le permite entender la fuerza del altruismo como una variable independiente de su elemento esencial. En el límite, el reconocimiento de los intereses del otro es tan débil, que sólo se expresa en la acción cuando ella no nos cuesta nada. Por esta razón, una persona que usualmente llamaríamos egoísta, porque no tiene en cuenta los intereses de los demás cuando el costo para él es más que insignificante, puede ser un altruista en el sentido de Nagel, si se deja motivar directamente por un interés ajeno cuando el costo es nulo. Si este “egoísta” pisa inadvertidamente, en una aglomeración, los dedos gotosos del pie de un prójimo desconocido por quien no siente nada, y retira inmediatamente su talón, tenemos lo que Nagel llama un caso de altruismo. Pensando en estos ejemplos es que Nagel dice que resulta improbable que existan especímenes puros y genuinos de egoístas, de individuos que en ningún caso expresan su reconocimiento de la objetividad de las razones (84-85). Pero en los casos en que los intereses del otro entran en conflicto con los propios, el altruismo en el sentido de Nagel no exige que la decisión final sea distinta a la que es conforme con los propios intereses. Nagel sostuvo que la objetividad de los intereses ajenos sólo nos provee de razones *prima facie* (125, 133). Estas pueden ser ven-

cidas y anuladas por otras razones, sin menoscabo de su objetividad. Para los casos de conflicto, es preciso contar, además de las razones objetivas *prima facie*, con principios combinatorios que nos permitan llegar a una decisión. Estos plantean un problema difícil, respecto del cual Nagel dio sólo algunas breves sugerencias, pero en su opinión es un problema distinto al de la objetividad misma de las razones.

Lo mismo que Sturgeon, Sober y Wilson basan el veredicto de la implausibilidad en los casos usuales de conflictos de intereses. (Sturgeon menciona también los casos de competencia en los juegos, pero es más fácil acomodar estos a la objetividad de las razones. Cfr. las observaciones del mismo Nagel en PA, 131-132). Si dos individuos A y B tienen un conflicto de intereses, la tesis de Nagel parece implicar que A tiene razones para actuar al mismo tiempo en su favor y en favor de B; y lo mismo vale para B. Si el conflicto es que A tiene razones para dañar a B y B obviamente tiene razones para no ser dañado por A, entonces se seguiría, si las razones son objetivas, que A tiene razones para dañar y no dañar a B, y B razones para no ser y ser dañado por A, lo cual realmente parece absurdo. La objeción de implausibilidad también puede expresarse diciendo que la tesis de la objetividad de las razones nos obliga a describir los conflictos *interpersonales* como si implicasen, o debiesen implicar, al mismo tiempo, conflictos *intrapersonales*. Esta es una consecuencia contraintuitiva de la tesis de Nagel. Pues ella no

corresponde a los hechos, ya que no es así como experimentamos los conflictos interpersonales. Tampoco sería plausible afirmar que deberíamos, moralmente hablando, experimentarlos de ese modo. Sin embargo, la explicación que acabo de dar de la tesis Nageliana del altruismo —y por tanto de la objetividad de las razones—, permite dar una respuesta a esta objeción. La tesis de la objetividad de las razones implica que las razones del otro son para mí sólo razones *prima facie*. Esto no tiene por qué arrastrarme a constantes conflictos *intrapersonales*, si utilizo principios combinatorios que dan prioridad a mis razones por sobre las del otro, como puede ser usual en la mayoría de casos. La tesis de la objetividad no nos obliga a describir los conflictos interpersonales como implicando de hecho, o de derecho, conflictos *intrapersonales*. Todo depende de los principios combinatorios, que son, en la teoría de Nagel, un problema aparte de la tesis de la objetividad.

Como ya se dijo, Nagel pensaba que el altruismo, o reconocimiento de la objetividad de las razones, es una realidad más o menos obvia, obscura acaso por las asociaciones pretensiosas del epíteto, y por ciertas doctrinas filosóficas que no distinguen entre razones *prima facie* y razones netas para actuar. Sober y Wilson no caen en estas trampas conceptuales y filosóficas en su discusión del altruismo, razón de más para recriminarles el no haber representado con fidelidad la tesis Nageliana. Destacan que un comportamiento egoísta no permite inferir

la existencia de deseos exclusivamente egoístas. “Cuando dos deseos están en conflicto, el más fuerte determina el comportamiento resultante” (242) Y agregan en una nota: “Los deseos individuales son a las disposiciones conductuales, como las fuerzas componentes son a las fuerzas netas. Así como una atención exclusiva a las causas netas nos da un cuadro empobrecido de los hechos causales, así también una atención exclusiva a las disposiciones conductuales nos da una representación limitada de los motivos psicológicos” (354, n.24). Estas precisiones hubieran podido ser utilizadas por Nagel (aunque sustituyendo el concepto de deseos por el de razones), para señalar que el comportamiento egoísta usual de las personas en casos de conflicto no desconfirma el reconocimiento de la objetividad de las razones, ni la existencia del altruismo en el sentido “débil” asumido por él. Pero sería igualmente ilegítimo, para Sober y Wilson, apoyarse en esas precisiones y en el sentido común para considerar que el caso está cerrado en favor del altruismo: ellas tan sólo sirven para mostrar que ciertos argumentos en su contra son lógicamente inválidos. Nagel, al parecer, va demasiado lejos, pues en cierto modo asume la existencia del altruismo y busca tan sólo una explicación del mismo. No obstante, pienso que sería apresurado sostener que su explicación no es pertinente para quien considere, como Sober y Wilson, que su realidad no es un hecho, y que lo que se requiere son pruebas en uno u otro sentido. Pues si el argumento de Nagel, o

uno inspirado en él, lograra su objetivo de mostrar cómo es posible el altruismo, es decir, si lograra explicar su existencia a partir de premisas tan incontrovertibles como las que él utilizó —a saber, el carácter motivante de las razones egoístas y la creencia de cada cual de que las demás personas son tan reales como él mismo—, entonces su explicación podría igualmente valer como una prueba de su realidad. En todo caso, no hay nada de implausible en su tesis de la objetividad de las razones, aunque su explicación está lejos de ser intuitivamente clara.

Volviendo a Sober y Wilson, concluyen que los argumentos filosóficos y psicológicos en favor y en contra del altruismo, ni establecen su realidad, ni la refutan. La hipótesis altruista y la egoísta se enfrentan ahora en pie de igualdad, y el derecho de fuerza del que ha gozado la hipótesis egoísta en las ciencias sociales contemporáneas queda sin respaldo legítimo (295). Con esta pequeña victoria se vuelcan a la tarea de presentar su argumento evolucionista en favor del altruismo *psicológico*. La primera parte del libro mostró que los comportamientos de tipo cooperativo, en los que un individuo beneficia a otro, son biológicamente altruistas y requieren de la selección de grupos para adquirir un valor adaptativo y evolucionar. La ambición de los autores es mostrar ahora que consideraciones evolucionistas sobre el valor adaptativo pueden ayudar también a decidir la cuestión acerca de los mecanismos psicológicos que los producen.

Es importante precisar el sentido

en el que puede preguntarse sobre el origen evolutivo de comportamientos específicos y de los mecanismos psicológicos que los sustentan. La incursión de la biología evolucionista en la explicación del comportamiento parece usurpar los derechos de las explicaciones culturales y psicológicas. Especialmente en relación con las explicaciones culturales, pueden darse a menudo conflictos de competencia. Pero dado que el comportamiento humano incluye hechos variados y de muchas clases, puede haber campo para explicaciones de ambos tipos y los autores se inclinan por un pluralismo. Cuando se trata de comportamientos universales, constantes en todas las culturas y de los que se sospecha un carácter innato, la presunción inicial es la de una explicación de carácter evolutivo. Es el caso, por ejemplo, de por qué los humanos consumimos preferentemente alimentos con alto contenido de azúcar y grasas. El gusto por grasas y azúcares fue seleccionado en un entorno evolutivo ancestral, en donde había escasez de esos nutrientes. Esto es lo que sugiere un examen de restos de huesos y dientes de nuestros ancestros. En condiciones de escasez, los individuos con ese comportamiento alimenticio ingerían una mayor cantidad de calorías en tiempos de abundancia, y tenían, en ese aspecto, una ventaja adaptativa frente a los demás. Cuando se trata de comportamientos que varían dentro de la especie, la presunción inicial, en cambio, es hacia explicaciones culturales. Pero también hay casos en donde la explicación evolutiva y la cul-

tural o relativa al entorno se complementan. Supongamos que se da la siguiente explicación de las diferencias en el uso de condimentos en las comidas. Los humanos hemos evolucionado para evitar los alimentos descompuestos (no es así en todas las especies). Como al mismo tiempo tenemos la capacidad de ajustar nuestro comportamiento a las condiciones del entorno, usamos más condimentos sobre las comidas en climas calientes, como medio para preservar los alimentos de la descomposición. La explicación hace referencia, al mismo tiempo, a una causa relativa al entorno —la diferencia de climas— y a una preferencia evolutivamente explicable, de modo que ambas se complementan en la explicación de este comportamiento (300).

La relación con las explicaciones psicológicas es distinta. Aquí no hay razón para pensar que estén en conflictos de competencia unas con otras. Pues cuando se ofrece una explicación evolucionista para un comportamiento, se da tan sólo una causa remota del comportamiento, y nada se dice sobre su causa próxima, es decir, sobre el mecanismo psicológico que lo produce. Y sin embargo, es inevitable que, cuando un comportamiento evoluciona por selección natural, también evolucione con él un mecanismo psicológico (199-200). La selección afecta tanto al comportamiento como al mecanismo causal, y esa es la razón por la cual consideraciones sobre el valor adaptativo también son aplicables a los mecanismos. Los autores no pretenden defender que hay explicaciones evolu-

tivas para todo caso o tipo de comportamiento benefactor, pues en principio es posible que algunos comportamientos benefactores sean sólo culturalmente explicables, en cuyo caso no cabría hacer hipótesis evolutivas sobre su motivación. Ellos escogen un comportamiento benefactor para el que los biólogos asumen un origen evolutivo en todas las especies. Se trata del cuidado que los progenitores dedican a su prole (*parental care*) (301). Es un rasgo que por lo general se explica por selección individual, por lo que se trata de un rasgo biológicamente egoísta. Aquí se ve la importancia de haber separado las dos partes de la investigación. En el estado actual del conocimiento biológico sobre los comportamientos que tienen un origen evolutivo, es más fácil defender la tesis del altruismo psicológico respecto de un comportamiento que es biológicamente egoísta. (Los autores especulan, no obstante, sobre la posibilidad de que en los humanos la evolución del cuidado de los hijos tenga también un componente de selección de grupos. Ello se debe a que en los humanos ese rasgo implica un vínculo fuerte de pareja, que a su vez supone una donación altruista entre individuos que no están genéticamente emparentados cfr. 348, n.3 y 360, n.6).

La tesis de que el tiempo que los padres dedican al cuidado de los hijos ha evolucionado por selección natural, significa que hay una predisposición genética para el mismo. Pero ello no implica un determinismo conductual que excluya la posibilidad de compor-

tamientos divergentes. La predisposición genética puede chocar con otras predisposiciones, ya sea genéticas o culturales, en cuyo caso el comportamiento resultante depende de todos estos factores y sus fuerzas relativas (301-303). Lo importante es que, si el comportamiento ha evolucionado por selección natural, lo mismo debe decirse de los mecanismos psicológicos que lo sustentan (308). Esto abre nuevas posibilidades en el debate en torno al altruismo. Ya que en el estado actual de nuestra psicología y anatomía, los mecanismos psicológicos no pueden ser inferidos, ni del comportamiento, ni de la estructura anatómica, quizás podamos decir algo acerca de la plausibilidad relativa de que evolucione uno u otro mecanismo. Esto exige elaborar unos criterios para evaluar la plausibilidad evolutiva. Cuando un mecanismo evoluciona por selección natural, se soluciona un problema de diseño, que consiste en producir un comportamiento precisamente cuando el organismo está frente a cierto rasgo del entorno. Por lo general, esto supone mecanismos de detección, y, en principio, varios de ellos podrían producir el comportamiento requerido. Las diversas posibilidades provienen de que muchas veces, o así lo asumen los autores al formular sus criterios, la detección no siempre se conecta al rasgo directamente relevante para la aptitud del organismo, sino a un rasgo que está más o menos correlacionado con él. Esto se debe a que algunos rasgos son más fáciles de detectar que otros. Por otro lado, suele suceder que los rasgos

que son fáciles de detectar tienden a tener un grado menor de correlación con los rasgos directamente relevantes (311). El primer criterio, entonces, es la confiabilidad de un mecanismo para producir un comportamiento adaptativo en presencia de cierto rasgo del entorno. La confiabilidad es una función de dos factores: el grado de correlación entre el rasgo detectado y el rasgo del entorno en cuya presencia el comportamiento es adaptativo, y la precisión con que se detecta el rasgo (306). Por ejemplo, si una bacteria acuática necesita evitar el oxígeno para sobrevivir, no necesariamente lo hace por medio de un detector de oxígeno. Dado que las altas concentraciones de oxígeno tienden a estar cerca de la superficie del agua, la bacteria puede evitarlas siendo capaz de distinguir entre superficie y fondo, y nadando siempre hacia el fondo, es decir, hacia abajo. Muchas bacterias pueden hacer esto detectando el campo magnético terrestre. Un buen detector del campo magnético terrestre puede ser mejor que un mal detector de oxígeno. Pero si la correlación entre elevación y concentración de oxígeno es menos que perfecta, entonces un detector de oxígeno con una precisión equivalente al mecanismo que detecta la elevación, sería más confiable para alejar a la bacteria del oxígeno. (305-306). En los organismos simples, la producción de comportamientos adaptativos sucede sin que tengamos que suponer mecanismos específicamente psicológicos, ligados a creencias y deseos.

Para investigar el caso específico

del altruismo, los autores asumen que los mecanismos en cuestión están esencialmente ligados a creencias y deseos, entendidos ambos como actitudes proposicionales (208-211). La detección está realizada en los humanos generalmente por procesos cognitivos, y estos se acoplan a procesos volitivos o deseos para producir un comportamiento. Los autores utilizan aquí una ficción para enfocar mejor los problemas de diseño que se plantean en organismos con capacidades cognitivas. Se imaginan un continuo que va desde organismos simples, que sólo pueden representarse sus estados de placer y dolor, hasta superorganismos con capacidades cognitivas ilimitadas, y se preguntan qué deseos últimos maximizarían su aptitud biológica. Para el caso de los organismos simples, el único deseo posible es el deseo hedonista de maximizar el placer y minimizar el dolor, pues este es el único contenido representacional del que disponen estos organismos. En cuanto a los superorganismos, no les podría ir mejor que teniendo el único deseo último de maximizar su aptitud. Dados sus recursos cognitivos ilimitados, son capaces de conocer todo lo que se necesite para saber qué cursos de acción maximizan su aptitud biológica. Los humanos estamos en algún punto entre esos dos extremos. Estaríamos mal diseñados, si tuviésemos el único deseo último de maximizar nuestra aptitud, pues nuestros recursos cognitivos son limitados, y no sabríamos qué hacer si nuestra única directiva de acción fuese la de hacer todo lo que conduce a nuestro



éxito reproductivo. Es de esperarse, entonces, que organismos como nosotros no persigamos como único fin último la aptitud biológica (los autores no dicen, pero sugieren, que ni siquiera es uno de nuestros fines últimos, cf. 361, n.12. Este es un concepto demasiado abstracto para que figure en el contenido proposicional de nuestros deseos naturales. No debe confundirse, por ejemplo, con el deseo de tener hijos). La única alternativa viable para nosotros es que nuestros deseos últimos se enfoquen sobre estados de cosas que, aunque distintos de la aptitud biológica, estén suficientemente bien correlacionados con ella. Por hipótesis, el bienestar de los hijos es uno de esos estados de cosas, y los comportamientos destinados a producirlo fueron objeto de la selección natural. La pregunta es, entonces, qué mecanismo psicológico es más apto, es decir, más confiable para asegurar esos comportamientos. Si la tesis egoísta está en lo correcto, tendría que ser en este caso un mecanismo hedonista, que no tendría el bienestar de los hijos como fin último, sino como medio para un placer psicológico.

Para evaluar la confiabilidad relativa de los mecanismos rivales, hay que entender su funcionamiento. ALT, el mecanismo altruista dirigido al bienestar mismo de los hijos, desencadena un comportamiento benefactor cuando se detecta que los hijos necesitan ayuda. Es un mecanismo directo, y su confiabilidad depende sólo de la precisión con que se detecte el bienestar/malestar de los hijos. HED, en cambio, como

mecanismo egoísta-hedonista, no desencadena un comportamiento benefactor cuando se detecta que los hijos necesitan ayuda, sino sólo cuando se detecta el placer/dolor asociado a la creencia en el bienestar/malestar de los hijos. HED es entonces indirecto. Requiere, al igual que ALT, de la detección del bienestar/malestar de los hijos, pero no conecta el comportamiento directamente a ella, sino al placer/dolor asociado a ella. Su confiabilidad depende entonces de otros dos factores adicionales: de la precisión con que se detecte el placer/dolor, y de la correlación entre placer/dolor y la creencia en el bienestar/malestar de los hijos. Suponiendo que HED detecta el placer/dolor con la misma precisión con la que ALT detecta el bienestar/malestar de los hijos, y que la correlación del bienestar de los hijos con la aptitud biológica es positiva, la evaluación comparativa de confiabilidad va a depender de lo que se pueda decir del grado de correlación entre placer/dolor y la creencia en el bienestar/malestar de los hijos (317). Si la correlación fuese perfecta, la confiabilidad de ambos mecanismos sería idéntica, y no habría modo de decidir entre ellos, al menos con base en la confiabilidad. Pero los autores piensan que hay razones en contra de una correlación perfecta. Las razones, sin embargo, no están muy claras. Sostienen que, si el grado de correlación que existe entre el dolor físico y el daño corporal no es perfecto, no hay razones para pensar que la correlación entre el placer/dolor psicológico y la creencia en el bien-

tar/malestar de los hijos sea perfecta (315-316). Confieso, no obstante, no ver bien en qué se fundamenta esta analogía.

Hay una línea de reflexión, sugerida en un párrafo, que me parece más promisorio para explicar por qué la correlación no es perfecta. Para que haya una correlación perfecta entre creer que los hijos están bien/mal y sentir placer/dolor, es necesario que el placer derivado de esa creencia no pueda ser contrarrestado por otros placeres (315). Pero ¿por qué podría serlo? Esto es lo que requiere aclaración. La razón es que, en un mecanismo de tipo hedonista, el bienestar de los hijos y el placer tienen un valor relativo diferente: el valor que se le asigna a medios y fines respectivamente. Para HED, el placer es indispensable, mientras que el bienestar de los hijos no lo es. Hay muchos otros medios para el placer, y cuando el objetivo es el placer mismo, todos son, por así decirlo, intercambiables. En circunstancias en las que puedo obtener mayor placer por otros medios, la creencia de que los hijos están bien deja de estar correlacionada con el placer. Hay que notar que un mecanismo altruista no excluye la posibilidad de sentir placer en el bienestar de los hijos. Sostener lo contrario sería implausible, pues, como cualquier deseo, la satisfacción del deseo por el bienestar de los hijos produce placer. Pero en este caso, el placer no es intercambiable con otros, y es plausible que esté aquí perfectamente correlacionado con la creencia de que los hijos están bien.

Aunque la confiabilidad de ALT sea superior a la de HED, es todavía posible que HED evolucione en lugar de ALT. Esto sucedería si ALT no hubiese estado disponible para la selección natural (segundo criterio), o si fuese energéticamente menos eficiente (tercer criterio), es decir, si comparado con HED emplease demasiada energía y disminuyese por ello la aptitud biológica del organismo. Pero no hay razones para pensar ninguna de las dos cosas. ALT no necesita de más elementos psicológicos que HED. Ambos requieren de mecanismos para sentir placer/dolor y de mecanismos para representar deseos y creencias. No puede decirse que ALT requiera de más creencias que HED, pues también éste requiere de creencias respecto del bienestar de los hijos, sólo que HED pone este contenido proposicional en la "caja" psicológica de los deseos instrumentales, mientras que ALT lo pone en la de los deseos últimos. Tampoco es plausible decir que se requiera de más energía para ponerla en una u otra caja. De estas observaciones se deduce que si HED está disponible, también ALT lo está; y que todo parece indicar que ALT no requiere más energía que HED para su construcción.

La aplicación de estos tres criterios conduce a un caso positivo en favor de la evolución del altruismo psicológico. Los autores no entienden esta tesis en el sentido de que necesariamente el mecanismo altruista actúe sólo, sin concurso del mecanismo egoísta, para producir el comportamiento benefactor. Es posible, e incluso mejor desde el pun-

to de vista de la aptitud biológica, tener dos mecanismos para este comportamiento en lugar de uno (319-320). Pero esto refuta igualmente la tesis egoísta, pues es ella la que es exclusiva con respecto a su rival, y no al contrario (225, 228).

Concluyo aquí esta reseña ya demasiado extensa, esperando haber dado un cuadro completo y claro de los argumentos centrales. Los interesados en el tema encontrarán que la obra es más rica de lo que aquí se alcanzó a mostrar. Se trata de un tema nuevo, con muchos ángulos que requieren de clarificación y discusión. Los autores han hecho un trabajo imponente en abordarlos con cuidado, rigor y claridad; e incluso en los puntos en que no obtengan el asentimiento del lector, su presentación es una invitación a continuar pensando y desarrollando una concepción del altruismo en la que la biología, la filosofía y la ciencias sociales den un ejemplo vivo de cooperación interdisciplinaria.

ALEJANDRO ROSAS  
UNIVERSIDAD NACIONAL DE COLOMBIA

**William A. Rottschaefer.** *The Biology and Psychology of Moral Agency*, Cambridge: Cambridge University Press 1998. 293 p. ISBN 0-521-59265-8

La popularidad del enfoque naturalista en filosofía no es un fenómeno

exclusivo de la epistemología, ni de la filosofía de la mente. También se observa en la filosofía moral, como lo atestigua un inmenso volumen de publicaciones. Paralelamente al impulso que las ciencias de la cognición dieron a la naturalización de la epistemología, la filosofía moral también recibió sugerencias naturalistas provenientes de las ciencias; pero las más controversiales, sorprendentes y quizás fructíferas llegaron, no de los psicólogos, sino de los biólogos, y alcanzaron gran difusión pública cuando Eduard Wilson – el internacionalmente célebre entomólogo de Harvard–, escandalizara a humanistas de diversas tendencias con la declaración de que había llegado la hora de *biologizar* la ética.

Detrás de lo que puede parecer una banal deformación profesional, hay una preocupación genuina y seria por superar el abismo que separa lo que Wilson ha llamado “dos culturas”: la científico-natural y la humanista, y lograr la gran síntesis entre ambas. Ella es compartida por Rottschaefer, quien hizo su entrenamiento profesional en filosofía de la ciencia y de la física en la Universidad de Boston con Abner Shimony, y aprendió, leyendo a Sellars, que una de las tareas de la filosofía es la de articular las conexiones entre la imagen científica del hombre, por un lado, y la imagen “manifiesta” por otro (celosamente formulada y preservada en la cultura humanista). A diferencia de quienes piensan que es mejor mantener a la ética y a las ciencias naturales en compartimentos separados, y que las conexiones entre ellas son antagó-