

La moral sin personas

Jorge Mínguez

Derek Parfit

RAZONES Y PERSONAS

Trad. de Mariano Rodríguez González

Antonio Machado Libros, Madrid 908 pp. 45 €

El problema de la identidad personal aparece por primera vez de modo explícito en la historia de la filosofía de la mano de John Locke. Locke es también el primero en defender una teoría psicologista de la identidad personal, teoría que ha sido dominante en la tradición filosófica angloamericana, a la vez que inaugura, con el caso imaginario del intercambio de cuerpos entre un príncipe y un zapatero, una metodología que ha hecho fortuna en la literatura sobre este tema. El problema de la identidad personal es, dicho en pocas palabras, el de determinar en qué consiste la continuidad de las personas a lo largo del tiempo. Lo que a primera vista pudiera parecer uno más de los pasatiempos ociosos a los que la metafísica nos tiene acostumbrados se ha revelado como un tema apasionante, no sólo por sus ramificaciones profundas en varios campos de la filosofía -desde la epistemología hasta la filosofía de la mente, pasando, cómo no, por la ética-, sino también por cómo ha traspasado las fronteras de la academia, inspirando el guión de más de una película comercial[1].

La tercera parte de *Razones y personas*, de Derek Parfit, constituye quizá la defensa más importante de la posición psicologista respecto a la identidad personal que se ha hecho desde que el tema resurgió con fuerza a finales del siglo XX, en parte como consecuencia de su publicación en 1984. Vistas las cosas desde la perspectiva que nos dan los años transcurridos desde su publicación, ésta es la sección del libro que ha tenido una repercusión mayor. Pero, aunque centraré mi comentario en esta parte, el libro merece ser leído al completo. En particular la parte IV, sobre el estatus moral de las personas no nacidas, es francamente apasionante. El libro, además de tener un extraordinario interés filosófico, es una obra maestra por la claridad de su argumentación y por la increíble imaginación de los muchos casos que utiliza como bombas de intuición (en expresión de Daniel Dennett). Su prosa, desnuda de todo adorno y elegantemente vertida al español por Mariano Rodríguez González, se entiende y atrapa, sin presuponer nunca un lector especialista en filosofía analítica.

El tratamiento parfitiano de la identidad personal, al insertarse en una argumentación más amplia de carácter moral, va mucho más allá de una teoría sobre el tipo de cambios a los que una persona puede sobrevivir. Porque *Razones y personas* es, ante todo, una larga defensa de la moral utilitarista. El utilitarismo es una teoría moral central en la tradición anglosajona, cuya exposición clásica debemos a Stuart Mill y que sigue contando con defensores importantes, pese a haber sido objeto de ataques tan formidables como el que John Rawls formula en su *Teoría de la justicia*. De acuerdo con el utilitarismo, debemos hacer lo que produzca un mayor beneficio al conjunto de las

personas. Tener un sesgo a favor de nosotros mismos o de quienes nos son cercanos puede ser una tendencia natural, pero no es moral. Como muestra de la radicalidad de esta posición baste esta cita de Parfit: «Debemos hacer lo que sería lo mejor para los hijos de todos, imparcialmente considerados. Al decirnos que ignoremos nuestras relaciones con nuestros propios hijos, R [la moral utilitarista defendida por Parfit] nos dice que ignoremos la que puede ser la más fuerte de todas nuestras relaciones personales» (p. 751). El principal escollo a la hora de guiar nuestras acciones por la moral así entendida es nuestro egoísmo, el cual incluye, como vemos, intereses autorreferenciales como la preocupación por nuestros propios hijos. Parfit dedica buena parte de su libro a explicar por qué el egoísmo se asienta sobre una teoría de la racionalidad incoherente, pero el argumento más importante es el que hace derivar esta incoherencia de una concepción equivocada de la persona. La identidad de las personas es «menos profunda» de lo que tendemos a creer, como también es menos profunda la separación entre unas personas y otras. En la visión de Parfit, las repercusiones prácticas de este concepto revisado de identidad personal son cataclísmicas. Una vez hayamos ajustado nuestras actitudes a nuestra nueva noción metafísica de identidad debemos dejar de tener en cuenta quién es el beneficiario de nuestra acción y preocuparnos sólo de cuánto beneficio (o daño) causaremos con ella. Así, la compensación intrapersonal deja de anteponerse al objetivo de maximizar el beneficio agregado para las personas. Con ello se borra, por ejemplo, la diferencia moral entre imponer un sacrificio a un niño para su propio bien futuro y hacerlo en beneficio de su hermano. Deja asimismo de estar justificado el interés natural que tenemos por nuestro propio futuro y por nuestra supervivencia. Como podemos leer en el último párrafo de esta tercera parte del libro: «Encuentro la verdad liberadora y consoladora. Me hace preocuparme menos por mi propio futuro y por mi muerte, y más por los demás. Doy la bienvenida a esta ampliación de mi interés» (p. 603).

Todos tenemos, siquiera implícitamente, una concepción natural o espontánea de lo que somos. Si no me equivoco, lo que caracteriza principalmente a las personas en esa concepción de sentido común es que tenemos una perspectiva del mundo, un punto de vista: somos seres conscientes. Cada uno de nosotros es un centro de gravedad en torno al cual se organizan todas las cosas. Esta perspectiva es, en principio, cognitiva: perceptiva y conceptual. Lo que percibo y conozco del mundo está estructurado en torno a mí. Pero también se trata de una perspectiva práctica: cada persona es un foco de intereses. Lo que primariamente le interesa a una persona son cosas relacionadas consigo misma. Esta preocupación por uno mismo –que compartimos con todos los seres vivos– cobra en el caso de los humanos una profundidad especial debido a nuestra capacidad de recordar el pasado y de anticipar el futuro. La conciencia, y su extensión en la memoria, proporciona una unidad a nuestra vida que va más allá de la que poseen el resto de los objetos. No sólo nos preocupa lo que nos afecta ahora, sino también nuestro futuro y nuestro pasado. Esto hace que la identidad personal a lo largo del tiempo sea siempre, para el sentido común y para la concepción metafísica que intenta ser fiel a él, una cuestión determinada. Como escribió Thomas Reid en 1758: «La identidad, cuando es aplicada a las personas, no tiene ambigüedad, y no admite grados, ni un más o un menos»[2]. Necesitamos que sea determinada, entre otras cosas, para poder justificar nuestras prácticas morales. Para nuestra moral de sentido común, la persona es la unidad relevante a la hora de juzgar acerca de cosas como la culpa, el bienestar o la redistribución. Es el propio Reid quien escribe, en la continuación del párrafo recién citado: «[La identidad] es el fundamento de todos los

derechos y obligaciones, y de toda responsabilidad; y su noción es fija y precisa». La cuestión es si esta visión de sentido común está metafísicamente fundada, o si es, por el contrario y como defiende Parfit, efecto de una ilusión.

Según sostiene Parfit, la concepción metafísica de la persona que subyace a la visión de sentido común es el no reduccionismo, es decir, la idea de que somos *algo más* que nuestras características objetivas, algo más que nuestro cuerpo o nuestras experiencias considerados como entidades observables. Aunque Parfit no es explícito en este punto, creo que la razón principal que le lleva a rechazar el no reduccionismo tiene que ver con nuestra concepción general del mundo, en la que no hay cabida para entidades ajenas al lenguaje de las ciencias naturales. Frente a esta visión de la persona como algo distinto de sus propiedades empíricas, Parfit defiende un reduccionismo de corte humeano, según el cual «la existencia de una persona no consiste en otra cosa que en la existencia de un cerebro y de un cuerpo, y en la ocurrencia de una serie de sucesos físicos y mentales interrelacionados» (p. 389). Además, en el reduccionismo todos estos hechos básicos pueden describirse de modo impersonal: «Aunque las personas existen, podríamos dar una descripción *completa* de la realidad *sin* afirmar que existan personas» (p. 391). Hasta aquí, el programa del reduccionismo ontológico. No existen entidades misteriosas, sino seres naturales que tienen relaciones causales entre sí. Pero Parfit añade una especificación importante al tipo particular de reduccionismo que él defiende: el criterio psicologista de la identidad personal.

Parfit intenta dar plausibilidad al reduccionismo psicologista sobre la identidad personal con su célebre experimento mental del teletransportador. Se trata de un peculiar transportador de personas, del estilo del que puede verse en la serie *Star Trek*, que funciona copiando los estados celulares de la persona en el punto de partida y enviando los datos al lugar de destino, donde éstos son utilizados para informar un cuerpo virgen. La persona original es destruida en el proceso. Pese a que es el cuerpo entero lo que es escaneado, la fase crucial del proceso es el copiado de los estados neuronales, que subyacen a los estados psicológicos. Por ejemplo, si el viajero tiene, antes de utilizar el teletransportador, la intención de visitar a un amigo cuando llegue a su lugar de destino, esa intención corresponde a un estado físico concreto de su cerebro, el cual, al ser copiado en el nuevo cuerpo en el lugar de destino, servirá de soporte para un estado psicológico idéntico al que tenía la persona antes de viajar. No hace falta decir que, pese a ser una copia cualitativamente idéntica, el cuerpo del lugar de destino no es *el mismo* que el del punto de origen, pues no se cumplen los requisitos mínimos de continuidad material exigibles para la identidad intertemporal de los cuerpos en general. Pero al copiar los estados neuronales se consigue crear una relación psicológica entre los dos individuos similar a la que existiría entre dos momentos sucesivos de una misma persona. Es decir, se consigue que haya conexión psicológica entre ambos individuos. Estas conexiones directas son la base de la continuidad psicológica de la persona en el tiempo. Conexión y continuidad psicológicas constituyen la identidad personal, según Parfit, que al referirse a ellas utiliza la expresión «relación R». Supongamos que experimento algo antes de ser escaneado, y que recuerdo esa experiencia después, ya en el lugar de destino y con mi nuevo cuerpo y mi nuevo cerebro recién «configurado». Según Parfit, nuestra intuición nos dice que, en un caso así, yo sería la misma persona con un cuerpo diferente. Cierto que la relación R no tiene aquí su causa normal, que es la continuidad del cuerpo. Pero tiene una causa, y eso es suficiente para conservar la identidad. Ahora bien, la posición

de Parfit es en este punto algo ambigua. Su método ha consistido en preguntarnos qué es lo que realmente nos importa cuando nos preocupamos por nuestro futuro y nuestra supervivencia. Hemos contestado que lo que nos importa es la relación R, y él ha tomado esa respuesta como un indicador del criterio correcto de identidad personal. Pero hay casos en que la identidad personal *no puede* adaptarse a ese criterio. Son los casos de fisión. Éstos le llevarán a modificar su argumento, haciendo bascular la carga que soportaba la concepción metafísica de la identidad personal sobre una nueva cuestión, ya genuinamente moral, que es la de qué nos importa realmente cuando nos preocupamos por la identidad personal.

Imaginemos, siguiendo al propio Parfit, que el teletransportador funciona mal y no destruye el cuerpo original. Si antes habíamos aceptado que la persona copiada en el lugar de destino era la misma que entraba en la máquina en el punto de partida, tendremos que reconocer que ahora el original tiene dos continuadores, uno con el mismo cuerpo y otro con un cuerpo distinto. Pero esto viola las leyes de la identidad, pues es evidente que las dos personas resultantes no son la misma. Si hay fisión, no hay identidad. Si continuáramos convencidos de que lo que nos importa es la identidad, es decir, que haya alguien en el futuro con quien seamos numéricamente idénticos, nos encontraríamos con una extraña consecuencia: la persona copiada deseará que al menos una de las dos –presumiblemente *la otra*– sea destruida, porque, si no, ella no sobrevivirá. Cuando el hombre que ha sido escaneado y replicado comprueba que no ha sido destruido, podrá lamentarse porque su no destrucción (y su consiguiente duplicación, aquí y en el lugar de destino) representa, si no su propia muerte, sí la muerte de quien, según todas las apariencias, era él mismo hace sólo unas horas. Esto, según Parfit, es absurdo, por lo que concluye que lo que nos importa no es tanto la identidad como tener una relación R (conexión y continuidad psicológicas) con alguien. La solución propuesta por Parfit para los casos de fisión es aceptar que en estos casos la relación R no preserva la identidad, pero a la vez insistir en que preserva lo que en condiciones normales nos importa realmente en ella. La duplicación es sólo un inconveniente menor, que sin duda llevará asociadas sus propias oportunidades. La identidad personal, por tanto, tiene sólo una importancia relativa, como medio normal de conservación de lo que realmente nos importa, que es la relación R. De aquí extrae Parfit una defensa del utilitarismo y de cierta indiferencia ante la muerte. Lo que tenemos que ver es si esta intuición, la de que lo que nos importa es la relación R, tiene el peso que Parfit le otorga en su teoría.

Volvamos ahora a la repercusión que tiene el reduccionismo, según Parfit, sobre nuestras actitudes prudenciales y morales. La idea de fondo es que nuestras actitudes normales están basadas en una concepción metafísica falsa. Dos tipos de casos, los de indeterminación y los de fisión, revelan la incapacidad de la vieja concepción metafísica para servir de guía fiable para nuestras actitudes. La teoría reduccionista de la identidad personal, en cambio, no hallará dificultades en estos casos y dará lugar a nuevas actitudes, que nos harán ver nuestro propio futuro como algo menos importante y el bien de los demás como algo más importante. Veamos primero el caso del espectro psicológico, que Parfit utiliza para ilustrar el problema de la indeterminación.

Imaginemos que existe la tecnología (y la voluntad experimentadora) necesaria para sustituir una parte tan grande como se quiera de los contenidos psicológicos de una persona por los de otra. Veamos lo que ocurre contemplándolo desde el punto de vista

de la primera persona. Cuando sólo un pequeño porcentaje de mis contenidos psicológicos es sustituido, puedo imaginar que yo seguiría siendo el mismo, sorprendido quizá de recordar cosas que no cuadran con el resto de mis recuerdos. Si la sustitución es masiva, el criterio psicologista nos dice que mi cuerpo sería ahora el soporte para un yo distinto, el que ha aportado los contenidos psicológicos injertados en mi cerebro en sustitución de los míos. Pero, ¿y en un caso intermedio, en el que son sustituidos la mitad de mis contenidos mentales? En nuestra visión de sentido común, no reduccionista, donde yo soy un alma cartesiana o al menos algo distinto de mis propiedades empíricas, la cuestión de si quien ocupa mi cuerpo actual soy yo o no lo soy es siempre determinada. Por mucho que uno cambie, uno sigue siendo el mismo; por mucho que otro se parezca a mí, siempre será otro. Pero no sabemos cómo podemos abordar la cuestión de si sobrevivo o no en un caso como el del espectro psicológico. En la visión reduccionista, en cambio, podemos aceptar que la cuestión no tiene una respuesta determinada, que nuestra supervivencia es una cuestión de grado: «La pregunta es aquí *vacía*. En cada uno de estos casos yo podría saber hasta qué grado estaría yo psicológicamente conectado con la persona resultante. Y yo podría saber qué conexiones particulares se darían o no se darían. Si conociera estos hechos, lo sabría todo» (p. 424).

Normalmente doy por hecho que la perspectiva que tengo ahora sobre el mundo –cuya posesión me convierte en una persona determinada– se mantiene incólume en el tiempo mientras sigo existiendo, y desaparece completamente cuando dejo de existir. No hay término medio. Por eso la posibilidad de la pérdida de memoria o del deterioro de mis facultades mentales no me parece en absoluto un mal menor, como habría de parecerme sin duda de creer yo que ese deterioro psicológico iría acompañado de una ruptura de la identidad entre mi yo actual y ese desgraciado yo futuro, sino que me aterra incluso más que otro tipo de males. Yo seguiré viendo el mundo desde esa perspectiva, por muy deteriorada que esté. En cambio, la visión reduccionista sostiene que el grado de conexión psicológica que tengo con «mis» yoes futuros (que en realidad son sólo, en expresión de Nozick, «candidatos» a ser yo) varía con las circunstancias, y en consecuencia varía con ella la identidad personal que me une a esos yoes supuestamente míos. O, al menos, varía el grado en que es razonable que lo que les ocurra a ellos haya de importarme. (La continuidad psicológica, que forma parte de la relación R que caracteriza a la identidad personal, tiene menos peso a la hora de determinar lo que me importa en la identidad.) La idea de fondo es clara: lo que debe importarme depende de mi grado de conexión psicológica con esos supuestos yoes futuros, y no de una misteriosa conexión que se dé como una cuestión de todo o nada. En situaciones normales, el grado de conexión psicológica entre dos estadios de una misma vida estará en función del tiempo transcurrido entre ambos y de los cambios psicológicos acaecidos.

En los casos de fisión o ramificación la visión espontánea de la identidad vuelve a fracasar estrepitosamente. Recordemos que aquí tenemos una persona que se escinde en dos, de modo que no puede considerarse que ninguna de las personas resultantes sea idéntica a la primera. La visión no reduccionista, enfrentada a una situación de este tipo, se ve abocada a equiparar fisión y muerte. El reduccionista, en cambio, puede aceptar la descripción de los hechos, dejar de lado la identidad personal y evaluar directamente en qué medida la fisión satisface lo que realmente le interesa de la supervivencia.

El reduccionismo de Parfit defiende que las personas pueden describirse en términos de unidades más simples, que estas unidades son estados psicológicos y que los estados psicológicos pueden describirse sin hacer referencia a la persona en que se presentan, es decir, impersonalmente. Esto le lleva a rechazar incluso el término habitual, «estado mental», por parecerle que los átomos psicológicos no deben presentarse como propiedades de un sujeto preexistente, ya que se trata justo de lo contrario: esos átomos psicológicos son el material con que está construido el sujeto. De ahí su preferencia por el término «suceso»: «Mientras que un estado tiene que ser un estado *de* una entidad, esto no ocurre con un suceso» (p. 389). Parfit reitera de este modo la crítica de Lichtenberg a Descartes: de mi certeza de que pienso no se sigue que exista un sujeto que posea esos pensamientos, sino simplemente que hay pensamientos. La persona –que es, como dijimos antes, esencialmente una perspectiva– está, así, constituida por elementos que no pertenecen necesariamente a un sujeto. Los recuerdos, por ejemplo, deben describirse sin presuponer que pertenecen a la misma persona que tuvo la experiencia recordada. Así redefinidos, los recuerdos son cuasirrecuerdos, recuerdos aparentes. Estos cuasirrecuerdos, junto con otros estados mentales redefinidos de la misma manera, constituyen auténticos átomos con los que se construye la persona. Lo que nos une a nuestro propio pasado no es que seamos un mismo sujeto, una misma persona, sino la conexión causal entre las experiencias vividas por alguien antes y las experiencias recordadas por mí ahora. Es esa conexión la que me convierte en una misma persona, y no al revés. Y, en todo caso, es esa relación, la relación R, la que justifica mis actitudes morales ante la persona que tuvo las experiencias. A medida que mi conexión psicológica con ella se debilita como consecuencia de cambios radicales en mi carácter o, simplemente, por el paso del tiempo, mi responsabilidad por sus actos, por ejemplo, se debilita también.

Esta noción atómica de los estados psicológicos ha sido atacada, a mi entender con razón, desde varios frentes. Desde luego, uno de los problemas que no puede resolver es el planteado por la existencia de pensamientos en primera persona en general[3]. Sin embargo, aquí voy a plantear una cuestión que, aunque está también relacionada con los pensamientos en primera persona, tiene un cariz decididamente moral[4]. Me refiero al estatus que los deseos egoístas tienen en la descripción impersonal propuesta por Parfit. Pondré un ejemplo. Cuando, siendo un chaval, Lance Armstrong tenía el deseo de ganar el Tour de Francia, no le habría costado prever que, en caso de conseguir su propósito, su vida iba a cambiar radicalmente. Si entiendo correctamente la propuesta de Parfit, lo que debería haber hecho Armstrong es descontar del bien esperado por ganar el Tour la desconexión psicológica prevista provocada por esos mismos triunfos. «Si gano –debía haberse dicho– seré más feliz, pero seré menos yo». Además, Parfit sostiene que la relación de conexión psicológica es *menos profunda* que la relación de identidad de sentido común, de modo que Armstrong debería haber aplicado un nuevo descuento, diferente del anterior. «Si gano el Tour varias veces –podría haberse dicho– la persona que ganará seré yo mismo, pero la diferencia entre que sea yo quien gane y que gane cualquier otro, quizá Jan Ullrich o Marco Pantani, es menor de lo que tiendo a creer». Para terminar, un Armstrong completamente convertido al parfitianismo podría añadir: «Lo que realmente importa no es que yo gane el Tour, sino que *alguien* lo gane, que alguien tenga esa experiencia». Pero, al redefinir así nuestros deseos egoístas, cambia enteramente su naturaleza. La persona que resulta de combinar estos estados mentales, revisados para eliminar de ellos el

elemento personal, es una persona totalmente distinta de las personas normales. En ella faltan muchas cosas que consideramos importantes y, desde luego, dudo mucho que nadie pueda ganar nunca el Tour si sólo tiene deseos impersonales de este tipo. Pero quizá lo más destacado es que, al estar construida a partir de estados psicológicos desprovistos del elemento personal, no puede conservar el deseo de ser libre, porque ese deseo tiene un elemento personal intrínseco. Cuando Lance Armstrong decide que quiere intentar ganar el Tour, su deseo conlleva el deseo de un cambio notable en su vida, pero esos cambios no ponen en entredicho su identidad personal, porque son cambios deseados y producidos libremente por él. La libertad entraña esta apropiación de las experiencias y de los cambios personales que uno mismo ha buscado o que no son impuestos desde fuera, así como un rechazo de toda autoridad externa no reconocida. El valor de la libertad es, por ello, ininteligible en una concepción reduccionista parfitiana de la persona, donde se ha eliminado totalmente el elemento personal de las entidades atómicas en cuyos términos se realiza la reducción. Y esto afecta también a los deseos propiamente morales. Una de las cosas que queremos cuando deseamos respetar y ayudar a los demás es que vivan vidas autónomas, que puedan dirigir su propia vida y abrirse paso entre el universo de valores para crear para sí mismos vidas virtuosas y con sentido. Sería risible que lo que quisiéramos para ellos fueran experiencias buenas, impersonalmente concebidas, pues sería querer para ellos una vida profundamente empobrecida, en el límite de lo que merece la pena vivirse. La imagen parfitiana no sólo acaba con conceptos como los de merecimiento, culpa y responsabilidad. Acaba también con el bien humano, concebido al modo aristotélico como actividad libre y virtuosa de la persona. Si dejamos de lado a las personas, nada tiene importancia: ni los hijos propios ni los de los demás.

[1] Arnold Schwarzenegger protagoniza dos de ellas, *Desafío total* (que les recomiendo) y *El sexto día*.

[2] *Essay on the Intellectual Powers of Man*, recogido en John Perry (ed.), *Personal Identity*, Berkeley y Los Ángeles, University of California Press, 1975, p. 1.

[3] Estas críticas al reduccionismo parfitiano no implican que no puedan defenderse otras formas de reduccionismo respecto a las personas y a la conciencia en general, pero se trata en todo caso de un reto difícil para el programa reduccionista.

[4] La objeción que aquí apunto, de raigambre claramente kantiana, ha sido desarrollada por diversos autores, entre ellos Susan Wolf, «Self-Interest and Interest in Selves», *Ethics*, núm. 96 (julio de 1986), pp. 704-720; Christine M. Korsgaard, «Personal Identity and the Unity of Agency: A Kantian Response to Parfit» (1989), recogido en *Creating the Kingdom of Ends*, Cambridge, Cambridge University Press, 1996; y Simon Blackburn, «Has Kant Refuted Parfit?», en Jonathan Dancy (ed.) *Reading Parfit*, Oxford, Blackwell, 1997.