

## EL PROBLEMA DE LA VOLUNTAD LIBRE

**Por: Ernst Tugendhat**  
Universidad de Berlín

El problema de la voluntad libre no es si existe o no, sino en qué consiste. Según una opinión popular, la voluntad libre consiste en poder hacer lo que uno quiere. Por ejemplo, puedo ahora, si quiero, alzar mi mano. En relación con esa opinión cabe hacer dos observaciones. Primero: que la acción depende de mi voluntad no es libertad de voluntad, sino de acción. Segundo: recientemente se ha mostrado en la fisiología cerebral que cada vez que uno quiere mover su mano, este acto de la voluntad es anticipado una fracción de segundo por un movimiento en el cerebro. Pero aún si eso no fuera así, tendríamos que suponer que el acto de la voluntad está determinado por antecedentes psicológicos, ya que la alternativa sería que la persona produce el acto de voluntad mágicamente desde la nada.

El problema que ha preocupado la tradición filosófica no ha sido el que se puede ejemplificar con el movimiento de una mano, sino el problema de la responsabilidad. ¿Cómo entender que nos podemos responsabilizar y reprochar recíprocamente, como también cada uno a sí mismo? Una tal responsabilización implica que la persona puede controlar lo que ella quiere. Depende del querer de la persona cuáles de sus deseos han de ser efectivos. Es aquí donde tiene sentido hablar no sólo de libertad de acción, sino de libertad de voluntad. Depende de mí si me decido a actuar según un querer o según otro. El querer de la persona misma es un querer de segundo nivel, un querer reflexivo, cuyo objeto son los deseos y quererres que uno encuentra en sí. Esto obviamente es un fenómeno específicamente humano. También en el caso de los animales se puede hablar de libertad de acción, pero no consideramos los animales como responsables, aunque también pueden mover sus patas cuando quieren.

El problema de la voluntad libre no es entonces un problema que concierne a la voluntad en general, sino específicamente a esta capacidad humana de una voluntad reflexiva que está implicada cuando se dice "depende de mí". Esto ha sido claramente visto en la tradición filosófica. Aristóteles hizo la distinción entre un querer sensitivo y un querer racional. El querer sensitivo es un querer simple, un querer que es simplemente un hecho de la conciencia. El querer racional es el querer que está basado en deliberaciones. Cuando deliberamos preguntamos por razones. Esta capacidad de deliberación es vista por Aristóteles como lo que distingue a los

hombres de los animales. Al mismo tiempo Aristóteles vincula esta capacidad con el hecho de tener conciencia del futuro: si los humanos no se refiriesen explícitamente a fines y a su vida futura, tendrían poca ocasión de deliberar. En Kant encontramos una caracterización de la libertad humana en contraste con la voluntad de los brutos muy similar a la de Aristóteles. Para John Locke el concepto central en este fenómeno es el de la capacidad de suspender un deseo. Es obvio que la deliberación sólo puede ser efectiva si la persona tiene la capacidad de suspender sus deseos inmediatos. Tenemos la capacidad de suspender nuestros deseos, dice Locke, para poder preguntarnos si son provechosos y buenos. Podemos añadir: reprochar a alguien sólo tiene sentido en el grado en que se presupone que tiene esta capacidad de autocontrol.

El problema central con relación a la libertad de la voluntad, siempre ha sido la pregunta sobre si este fenómeno es compatible o no con el determinismo causal. Muchos no han entendido que esto no es un problema que surge en relación con el fenómeno del querer en general, que se puede ejemplificar con mover una mano, sino sólo en relación con la voluntad deliberativa. En el caso de un querer simple el determinismo no causa problemas. La acción es libre, es decir, voluntaria cuando no es coaccionada, y parece obvio que la voluntad a su vez es determinada causalmente. Es un error pensar que ser determinado causalmente significa ser coaccionado. Pero el caso es diferente con la voluntad reflexiva. Cuando hablamos de autocontrol y responsabilidad se supone que la acción depende no sólo de los deseos de la persona, sino, por lo menos en parte, de la persona misma, y es este fenómeno el que puede parecer incompatible con el determinismo. El primer paso, antes de podernos confrontar con el problema del determinismo, tiene que ser la aclaración del fenómeno mismo, pues no parece claro qué significa hablar de la persona misma, de mí mismo.

Como punto de partida se presta el concepto opuesto al de la libertad, el concepto de compulsión. Se habla de compulsión externa y también interna, pero en los dos casos la palabra compulsión no significa lo mismo. La acción es libre si el individuo no está bajo una compulsión externa: el individuo hace lo que él quiere. En cambio, un individuo es responsable sólo si además no se encuentra bajo una compulsión interna. Estar bajo una compulsión interna significa que el individuo es incapaz de suspender sus deseos. Una compulsión interna no impide la libertad de acción, pero sí la libertad de voluntad. Si uno se encuentra bajo una compulsión interna, no tiene la flexibilidad para poder ajustar su comportamiento a sus razones, hace lo que quiere, pero no puede controlar su querer. Mi deliberación y mi juicio sobre lo que me parece mejor no tienen efecto. Mi juicio sobre lo que es mejor; ahí tenemos los dos polos claves: a mí de un lado y a la perspectiva de lo que me



parece bueno por el otro lado. La perspectiva de lo que uno piensa que es mejor es igualmente constitutiva para el autocontrol como el polo del yo. La suspensión de deseos no acontece al azar o en un espacio vacío, esto no tendría ningún sentido, sino siempre en función de algo que es considerado como bueno, en función de algún fin o alguna razón práctica. Para poder entender la relación entre estos dos polos, entre mí y un bien, debemos recordar lo que nos decía Aristóteles sobre la conciencia del futuro. La conciencia del futuro tiene como consecuencia que los humanos siempre se encuentran en lo que quiero llamar espacios de acción.

Estos espacios de acción aparecen en dos tipos. El primero es el espacio de deliberación y elección: deliberamos sobre cuál es el mejor camino que conduce a un fin, y también a qué fines nos queremos dirigir. Una vez que hemos elegido y nos dirigimos a un cierto fin, nos encontramos delante un segundo tipo de espacio. Este consiste en poder ser más o menos atentos hacia el fin. Nos podemos esforzar más o menos. En ambos tipos de espacios prácticos estoy consciente de que depende de mí, y en ambos estoy dirigido a algo considerado como bueno: en el primero se trata de elegir lo que es mejor, en el segundo de atenerme a mi fin. Ambos son formas de autocontrol, ambos exigen que uno logre suspender o contener los afectos contrarios, en ambos estamos expuestos a reproches. Otros me reprochan o yo mismo me reprocho que haya elegido mal o, en el otro caso, que haya sido negligente en la persecución de mi fin.

Ahora bien, en relación con estos espacios que se abren entre mí y mi idea de un bien, se pueden distinguir varios escalones, es decir, varias dimensiones de libertad de voluntad. En cada dimensión la meta de lo bueno y con esto el espacio de razones en relación a los cuales uno tiene que suspender sus deseos anteriores, es diferente.

Como primer escalón se puede considerar la necesidad para un ser humano de ocuparse de su propio futuro. Todo ser humano tiene que aprender en su infancia a tener que renunciar a gratificaciones inmediatas en favor de su propio bien prudencial. Este es un primer escalón de autocontrol y responsabilización. El segundo escalón me parece consistir en subordinar no sólo los sentimientos del presente a los sentimientos del futuro, sino en someter todos sus deseos inmediatos —tanto del futuro como del presente— a valores. Esto marca un segundo sentido del bien prudencial; es lo que en la filosofía antigua se ha llamado la pregunta por la vida buena, o por lo que es bueno para uno vivir. La moral se puede ver como un tercer escalón, en él se trata de integrar las razones de los demás a los valores propios. Así visto, el tercer escalón, la moral, aparece como una modalidad del segundo escalón, de la pregunta sobre cómo creo que es bueno vivir. Finalmente, surge un cuarto escalón que por el hecho de que los escalones dos y tres se encuentran en una

ambigüedad entre lo convencional y lo autónomo. Tanto los valores prudenciales como los morales pueden ser aceptados en forma convencional o bien la persona puede preguntar si son verdaderamente buenos.

Propongo llamar estos cuatro pasos “alternativas fuertes” para distinguirlos de las alternativas comparativas que se encuentran en las deliberaciones dentro de cada uno de ellos. Cuando comparamos uno de estos escalones con el anterior se trata siempre de una alternativa profunda. Correspondientemente al hablar de responsabilidad e irresponsabilidad adquiere cada vez un sentido diferente: ¿cómo puedes ser tan irresponsable, podríamos reprocharle a una persona, de no tener en cuenta las consecuencias que tu comportamiento tiene para tu futuro? O en referencia al tercer escalón, se diría: ¿no es irresponsable entenderse de una manera puramente egoísta? Finalmente, en el cuarto escalón nos encontramos delante de la pregunta de Sócrates: ¿no es irresponsable vivir sin preguntar si las razones que tenemos para nuestros valores son válidas?

Mi razón para indicar estas distinciones ha sido que fácilmente se tiende a pensar que la única alternativa profunda es aquella entre el egoísmo y la moral. Así se simplifica la complejidad del panorama. Creo que debería hacer una anotación aquí: muchos creen que la responsabilidad que tenemos delante de la ley es una responsabilidad moral, en mi opinión no lo es, es suficiente lo que he llamado el primer escalón, si la persona puede distinguir y tiene la capacidad de comportarse según su razonamiento en relación a su propio bien y mal, tiene que tomar en cuenta los castigos en que incurre cuando actúa contra la ley. La persona no tiene que tener una conciencia formal ella misma. Esos cuatro escalones son por eso la razón. Con eso creo que he esbozado más o menos la variedad de posibilidades, me queda probablemente otra que no se detiene tanto en la práctica, sino también podemos decir por qué no ser responsabilizados por nuestras opiniones; en ese caso la meta no es una de esas cuatro concepciones de lo bueno, sino lo que reflejo en contraste con lo que estoy opinando simplemente.

Ahora que el fenómeno está más o menos esbozado, nos podemos confrontar con el problema del determinismo causal. El fenómeno central consiste en los dos tipos de espacio que he distinguido, tanto en los espacios de poder deliberar como en los espacios de poder esforzarse, siempre podemos decir: depende de mí. Depende, en el primer caso, de cómo evalúo las alternativas y, en el segundo, depende de mí en qué grado me esfuerzo. Correspondientemente, cuando se me responsabiliza, se me dice “habrías podido sopesar las alternativas mejor, habrías podido atenerte mejor a tu fin; dependía de ti”. Este “dependía de ti” parece implicar que la corriente causal de motivos se ve interrumpida por esta referencia a mí. Yo



había hecho tal y cual cosa, y cuando se pregunta “¿por qué pasó eso?”, se me hace responsable a mí *en vez de* a las condiciones motivacionales. Es cierto que hubo condiciones motivacionales, sin embargo la manera como yo me conduje dentro del espacio que yo tenía parece ser lo suficientemente decisiva como para que me haga responsable a mí. Esto aparece como una detención de la corriente causal, cosa que puede parecer extraña, así que debemos estar pendientes de que se pueda tratar de una ilusión; pero es eso lo que está implicado tanto en la experiencia propia como en la imputación de la responsabilidad.

Obviamente es este fenómeno el que está a la base del incompatibilismo, es decir, la libertad de la voluntad no es compatible con él, pero no por determinismo causal. Cuando decimos “depende de mí”, puede parecer como si algo como el yo fuera el punto de arranque de la acción. En su forma más intrépida tal posición fue sostenida por el filósofo norteamericano Roderick Chisholm. Mientras que todos los otros eventos dependen causalmente de otros eventos, el evento de una acción humana depende, según Chisholm, del actor. Los seres humanos, dice Chisholm, tenemos cuando actuamos la prerrogativa que en general le es otorgada sólo a Dios, la de mover sin ser movido.

La crítica de teorías de este tipo se concentra normalmente sobre el punto de que una tal concepción se encuentra en contradicción con la estructura normal de la causalidad: un acontecimiento sólo puede ser causado por otros acontecimientos, no por una entidad o algo ajeno a su cadena. Sin embargo, a esta crítica se puede responder (y Chisholm ha respondido así) con la pregunta de por qué no puede haber más tipos de causalidad. La verdadera dificultad radica más bien en el concepto de un ente que mueve sin ser movido. Quizás se pueda entender esto en el caso de un Dios creador (aunque tampoco ahí la idea resulta fácil), pero ¿qué sentido puede tener decir de un ser humano que es una causa no causada? ¿Se puede decir que en el ente que soy yo, es decir en mí, hay un segundo ente que se puede llamar el yo o la mismidad, algo como un núcleo de mí? Aun si suponemos que esto tenga sentido, la idea de que esto sea algo que mueve sin ser movido parece absurda. Pues esto significaría que los actos de voluntad saldrían como burbujas de ese núcleo, y una tal idea no sólo resulta incomprensible en sí, sino además estaría lejos de la base fenomenológica que se debía explicar, es decir, del fenómeno de que yo tengo la capacidad de evaluar alternativas y la capacidad de atenerme con más o menos atención a un fin.

En mi opinión el incompatibilismo debe su atractivo sólo a las dificultades que aparecen en la posición contraria, en el compatibilismo. Tan pronto se intenta describir una posición incompatibilista positivamente, como lo hace Chisholm,

ella se derrumba. Pero esto no puede significar que debemos renunciar al fenómeno que se articula en la frase “depende de mí”, ni tampoco al de la detención de la corriente causal que implica. Debo, por consiguiente, antes de dirigirme al problema del compatibilismo, poner de relieve algunos aspectos que pertenecen al fenómeno del decir “yo”.

Para un ser humano, cuando ha aprendido a hablar, es característico que no se encuentra en estados de conciencia como por ejemplo sentir, querer, creer, tal como lo suponemos también en los otros animales, sino que tiene además una conciencia de que se encuentra en estos estados de conciencia. Esto lo expresamos en frases como “yo sé que yo creo o quiero tal y tal cosa”; igualmente podemos decir simplemente “yo creo, etc.”; “yo” es la expresión que se usa cuando se habla de sí mismo. Ahora bien, nos podemos dar cuenta de que en esta conciencia de mí mismo surge en varios respectos una conciencia de un espacio de posibilidades, y una tensión en la que me encuentro con relación a él. Primero: yo, quien sé que tengo este deseo o esta creencia, sé también que tengo estos otros deseos y otras creencias. En el decir “yo” la multiplicidad de mis estados subjetivos se reúne en *una* conciencia, y esto tiene la consecuencia de que puedo ver, por ejemplo, un deseo a la luz de mis otros deseos. Segundo: cuando creo o quiero que *p* sea alguna cosa, soy consciente de que se podría también creerlo o quererlo para *no-p*; me encuentro implícitamente en una suspensión entre estas posibilidades, suspensión que se hace explícita en la pregunta, en la duda y en particular en la deliberación. Tercero: tengo la conciencia de ser el mismo ahora que después, y sé que también después voy a tener deseos, etc. Cuarto, y basado en el punto anterior: me puedo relacionar en mi querer con fines que están en el futuro y también con fines que se basan en ideas de valor como quiero ser. Por esto, en el caso particular de los deseos, se abre otro espacio entre mis deseos inmediatos que tengo porque los siento (son los deseos sensibles de Aristóteles), y los deseos que están dirigidos a fines y a valores; deseos de este segundo tipo no los tengo porque los siento sino porque me propongo un fin, es decir, porque los considero como racionales y buenos. A mis fines me tengo que dirigir; por eso estos deseos dependen de una manera enfática de mí, en contraste con los deseos que simplemente encuentro en mí por sentirlos.

En ninguno de estos cuatro puntos —ni en la conciencia de sí, ni en la unidad de la conciencia, ni en la conciencia de encontrarme en un espacio de deliberación, ni en los deseos en que yo me dirijo a algo— está implicado un yo, sino que se trata simplemente de mí como encontrándome en diferentes comportamientos, en los cuales surgen ciertos espacios prácticos que resultan ser espacios para mí; me encuentro en una tensión con respecto a ellos. Estamos confrontados una vez más con los dos espacios de los que hablé anteriormente (con el espacio de las razones



y con el espacio de la tensión que resulta por encontrarme entre deseos que me afectan porque los siento, y deseos a los cuales me tengo que atener). Y para estos espacios parece característico que nos relacionemos con ellos mediante un poder que parece ser irreductiblemente un “yo puedo”.

Todos sabemos de nosotros mismos y de los otros que cada cual —y esto significa cada vez “yo”— puede sopesar razones y que puede esforzarse más o menos. Aquí nos topamos con algo que siempre se ha visto como fundamental para la libertad de la voluntad, y es cuando se puede decir “podría haber actuado de otra manera”, pues este “podría” no es simplemente una posibilidad cualquiera: la conciencia de que yo puedo de otra manera es siempre una conciencia de poder mejor, y es un “yo puedo” en el sentido de que depende de mí, de mi esfuerzo, si lo voy a hacer mejor o no. Tanto la expresión “podría de otra manera” como también el concepto de una elección son demasiado formales. También los animales pueden, en ciertas condiciones, elegir entre diferentes opciones. Y cuando nosotros tenemos que elegir, por ejemplo, entre diferentes tortas o diferentes helados, no se trata de libertad de voluntad. En tal elección lo que decide es mi gusto y no se trata de que yo debería hacer algo mejor.

Si ahora nos confrontamos con el determinismo, esto se puede hacer en dos pasos. En un primer paso se puede ver así: en el hecho de que depende de mí efectivamente ocurre una detención en la corriente causal; en vez de la corriente causal soy yo quien es responsabilizado, puesto que yo he intervenido en la corriente causal cuando he suspendido los motivos inmediatos hacia un fin o los podía haber suspendido. Por otro lado parece plausible ver este acto de suspensión como a su vez condicionado causalmente. Se puede aclarar esto con la imagen de una cuerda en la que hay un nudo. La cuerda representa la corriente de causalidad. Por medio del nudo, que representa la acción del yo en los dos espacios, la causalidad se ve interrumpida y sustituida por mi actividad y, sin embargo, también el nudo pertenece a la cuerda. Aunque no podemos probar que la acción del yo está determinada causalmente, tampoco hay una buena razón para negarlo. ¿Por qué, se podría preguntar, no va a estar causalmente determinado el grado en que doy peso a las diferentes razones, y por qué no será determinada la fuerza del yo que está a mi disposición en perseverar en mi fin? Así que la detención de la corriente causal y el estar determinado no se contradicen (el nudo intercepta la cuerda, pero también es de la cuerda). Sólo debemos evitar la idea de que la persona pudiera decir: “puesto que está determinado, no tengo que deliberar o no tengo que esforzarme”. Esto sería el malentendido que se llama fatalismo, según el cual, el resultado no sólo está determinado sino que está determinado independientemente de mi esfuerzo, mientras que la conexión causal tiene que entenderse como pasando por dentro y

no por encima de la acción del yo. La corriente causal ocurre a mis espaldas, y esto no tiene que molestarme siempre que no confunda mi perspectiva interior con la perspectiva desde afuera.

Pero esta dificultad conduce a un segundo paso. Parece evidente que las frases subjetivas, cuando decimos "depende de mí esforzarme", no pueden ser traducidas a un lenguaje objetivo, es decir, a un lenguaje sin la palabra "yo". Igualmente no sabemos cómo traducir el lenguaje de razones en un lenguaje de causas. Una razón siempre es una razón para alguien que pueda decir "yo". La frase "depende de mí" parece irreduciblemente subjetiva, y cuando decimos de otra persona que depende de ella, suponemos que ella tiene igualmente la conciencia de que su esfuerzo depende de ella.

Así lo que finalmente resulta no es un yo que sería acausal, sino que nos encontramos diciendo "yo", y sabiendo que depende de mí (no de un yo), en la perspectiva de que no se puede traducir a un lenguaje objetivo, a un lenguaje en que se habla de objetos. Nos encontramos en una perspectiva que está igualmente alejada de la causalidad como de la acausalidad.

Mientras que el incompatibilismo contiene una tesis dogmática, lo característico del compatibilismo es que solamente niega la tesis del incompatibilista, no sostiene a su vez una tesis positiva. De hecho, había dentro de la tradición del compatibilismo también la variante de que la voluntad libre implica necesariamente el determinismo, esta fue la tesis de David Hume, pero para defenderla sostenía que la imputabilidad implica un carácter psicológico determinado, pero este argumento no tiene plausibilidad. Exigimos los unos de los otros y cada uno de sí mismo que actuemos también en oposición a nuestro carácter.

Debemos distinguir entre carácter y persona. No responsabilizamos al carácter, sino al que dice yo. Si el carácter determina una acción con necesidad, lo que es posible, entonces eso no es libertad, sino compulsión. Así que según el compatibilismo, si se lo entiende bien, la voluntad libre no implica el determinismo ni lo contrario. Poniéndolo en la metáfora de la cuerda con el nudo, podemos dejar abierto si lo que pasa dentro del nudo es reducible a conexiones causales o no: para la estructura de la voluntad libre nada depende de esta alternativa.

La razón para sostener una posición compatibilista no es un dogmatismo, sino que lo único que se puede pensar en el mundo objetivo que está afuera de la conexión causal sería el azar. Lo que hace el incompatibilismo tan inatractivo es que intenta pensar algo dentro del mundo objetivo que no es ni lo uno ni lo otro, ni causalidad ni azar. Mientras que el compatibilismo sólo tiene que reconocer la



dificultad de que el lenguaje en que experimentamos el actuar desde adentro no parece poder ser traducido al lenguaje causal, el incompatibilismo es el intento de darle a la perspectiva interior un fundamento objetivo que no parece tener sentido.

Lo que llamé perspectiva interior es obviamente una perspectiva que se comparte intersubjetivamente. Nos relacionamos recíprocamente uno con el otro de tal manera, que cada uno supone que el otro tanto como él mismo tiene la capacidad de suspender sus deseos en favor de fines y razones. La esfera de razones es intersubjetiva: así como aprendemos el lenguaje intersubjetivamente, también aprendemos intersubjetivamente como se actúa por razones, así que por medio de los reproches que se nos hacen, es como aprendemos la práctica con razones.

Mi opinión de que el compatibilismo no puede explicar la práctica de reproches y de reprobación es errada. Esta objeción valdría más bien contra el incompatibilismo, pues ¿qué sentido podría tener reprender a un ser incondicionado? En un reproche se combinan dos cosas: que se considera una acción haber sido mala y se le dice esto al actor de una manera que presupone que habría podido actuar mejor, y esto corresponde exactamente al fenómeno de la libertad de voluntad. El que piensa que el compatibilismo no permite verlo así, supone que el segundo de estos factores —que la persona había podido actuar mejor— contradice el determinismo. Supuestamente la conexión causal impide que la persona pueda haber actuado de otra manera. Una vez más se confunde determinismo y compulsión. Vimos que la conexión causal no determina la acción directamente, es decir, con independencia de mí; la responsabilidad es mía, pero esto puede ser a su vez causalmente determinado. Lo que llamé la detención de la corriente causal es realmente su desplazamiento a mi interior. El determinismo no conduce a dudar de que la persona hubiera estado en un espacio abierto, y así realmente podía haber actuado mejor.

Sin embargo, siempre es posible que la persona no haya podido pueda actuar de otra manera. Pero esto ocurre cuando el querer de la persona se encuentra bajo una compulsión interna, es decir, cuando no tenía la capacidad de controlar su acción. Así que nuestro derecho de hacer reproches es realmente limitado, pero no por el determinismo, sino porque una persona sólo puede ser considerada como responsable, si realmente tenía la capacidad de suspender sus deseos y actuar según razones, y si esto es así, en que medida, no está siempre claro. Si tuviéramos sólo el caso individual, no podríamos decidir cuándo una persona no actúa como debe, si fue porque no quería o porque no podía. Es sabido que este problema de en qué grado la persona es imputable, es especialmente grave en el juicio penal. La existencia

de una capacidad sólo se puede verificar de modo general, tomando pruebas, es decir, en nuestro caso, por observación sobre si y en qué medida la persona y también uno mismo está en general en la capacidad de cambiar su comportamiento por razones y esfuerzo. Debemos entonces distinguir entre dos preguntas: la pregunta acerca de qué está implicado en el concepto libertad, y fue sólo eso de lo que he tratado en esta ponencia, y la pregunta acerca de hasta dónde llega. Muchas veces esta capacidad no llega tan lejos como se cree, y esto debe llevarnos a no usar de manera tan fácil este instrumento tan cruel de reproches y autorreproches. Esto sería a su vez irresponsable.