

## PREFERENCIAS PASADAS Y UTILIDAD<sup>1</sup>

BLANCA RODRÍGUEZ LÓPEZ  
*Universidad Complutense de Madrid*  
bmerino@filos.ucm.es

### ABSTRACT

Preference Utilitarianism faces some problems and criticisms about the suitability of assigning weight to every single preference any agent happens to have. Some of these problems arise from time-related factors, up to the point that authors such as Brandt (1979) has proposed to abandon Preference Utilitarianism in favor of a happiness version close to hedonism. This paper focuses on the analysis of time-related preferences and in particular now-for-then preferences, taking into account Hare's and Brandt's viewpoints. We will ponder the arguments in favor and against leaving aside such preferences, discussing their merit depending on the two different moments involved in now-for-then preferences. Given that the then-moment relates these preferences with past preferences, we will present reasons for excluding past preferences from the utility function. As a consequence of this analysis, we face an alternative; on the one hand, we can take now-for-then preferences into account, thus moving away from utilitarianism insofar as we break the connection between happiness and preference satisfaction. And on the other hand, we can exclude them, embracing a hedonistic utilitarianism and rejecting the widely accepted version of Preference Utilitarianism.

**Keywords:** Utilitarianism, Preferences, Now-for-then preferences, hedonism

### RESUMEN

El utilitarismo de la preferencia se enfrenta a algunos problemas y debates acerca de la conveniencia de tomar en cuenta en el cálculo de la utilidad todas las preferencias de los distintos individuos tal y como estas se presentan. Algunas de estas preferencias "problemáticas" son las referidas a acontecimientos situados en un momento temporal distinto de aquel en el que se tiene la preferencia, que plantean problemas de tal magnitud que hacen que Brandt (1979) planteara la conveniencia de relegar la versión preferencialista a favor de una versión de la felicidad muy cercana al hedonismo. Este artículo se centrará en el análisis de esta categoría y en particular de las preferencias presentes acerca

---

<sup>1</sup> Aceptación: 26 de octubre de 2005.

de acontecimientos futuros, teniendo especialmente en cuenta las posiciones de Hare y Brandt al respecto.

Analizaremos los argumentos a favor y en contra de su exclusión en los dos distintos momentos, presente y futuro, involucrados en estas preferencias. Puesto que el segundo momento relaciona la problemática de estas preferencias con la de las preferencias pasadas, sopesaremos los argumentos para excluirlas del cálculo utilitarista. A partir de este análisis plantearemos una disyuntiva: o bien las admitimos, alejándonos del utilitarismo al renunciar a la conexión entre felicidad y satisfacción de las preferencias, o bien las excluimos (salvo en la medida en que producir el acontecimiento objeto de la preferencia pasada satisfaga alguna preferencia presente). Si aceptamos esta última vía, que parece la más defendible, estaremos dando un paso que nos aleja de la popular versión preferencialista y nos acerca a una versión hedonista del utilitarismo.

**Palabras clave:** Utilitarismo, preferencias, preferencias presentes acerca del futuro, hedonismo

En 1954 Ayer publicó, con el título "*El principio de Utilidad*", una revisión crítica del utilitarismo, en la que, tras unas aclaraciones certeras sobre el alcance y correcto entendimiento del Principio de Utilidad, condena las afirmaciones de la teoría conocida como *hedonismo psicológico*, en las que Bentham apoyaba su Principio, a ser o bien empíricamente falsas o bien tautológicas. Esto, sin embargo, no resultaba tan fatal para el utilitarismo como pudiera parecer pues era posible, según Ayer, que la parte esencial del "sistema de Bentham sobreviviera realizando una nueva identificación de la felicidad no ya con el placer y la ausencia de dolor sino con la satisfacción de los deseos. El Principio de Utilidad pasaría a decir que hay que dar a la gente la mayor cantidad posible de aquello que desea.

La desconfianza hacia el hedonismo psicológico no constituía una novedad. Por el contrario, ya desde finales del XIX en la teoría económica de corte utilitarista "subsistía un constante malestar acerca de los fundamentos psicológicos del utilitarismo" que hizo que los economistas, en el plazo de pocos años, prescindieran "de manera consciente de toda dependencia psicológica al tiempo que mantenían la estructura matemática de la teoría utilitarista"<sup>2</sup>. A partir de ese momento, el llamado Utilitarismo de la Preferencia (UP) fue abrazado por la práctica totalidad de los utilitaristas en el convencimiento de que se trataba de un versión superior, que superaba los inconvenientes planteados por la versión hedonista clásica y que, tal como señalaba Ayer, tenía la ventaja de hacer que el Principio de Utilidad pudiera sostenerse con independencia de cualquier supuesto de carácter psicológico.

2 Hollis y Sugden (1993).

Sin embargo, no tardaron en aparecer nuevos problemas y debates acerca de la conveniencia de tomar en cuenta todas las preferencias de los distintos individuos (en el cálculo utilitarista de carácter moral y político) o incluso las propias (en el cálculo utilitarista de carácter prudencial) tal y como estas se presentan. Dicho de otro modo, la lectura del Principio de Utilidad ofrecida por Ayer, dar a la gente la mayor cantidad posible de aquello que desea, empieza a perder su atractivo, y empiezan a surgir versiones conocidas como Utilitarismo de la Preferencia Restringido (UPR) que cuentan en la actualidad con la adhesión mayoritaria dentro de las filas utilitaristas.

Las preferencias que podemos llamar “problemáticas” pueden agruparse en tres categorías:

- Preferencias irracionales.
- Preferencias relativas a acontecimientos situados en un momento temporal distinto de aquel en el que se tiene la preferencia.
- Preferencias externas.

Cada una de estas categorías presenta una problemática propia que hace que se plantee su exclusión del cálculo de utilidad individual y/o general<sup>3</sup> y es habitual que el UPR tienda a excluirlas todas.

La segunda categoría, aunque no sea quizá la más debatida, plantea problemas de tal magnitud que hacen que Brandt (1979) planteara la conveniencia de relegar la versión preferencialista a favor de una versión de la felicidad muy cercana al hedonismo. Mi comunicación se centrará en el análisis de esta categoría.

Preferencias y tiempo: el planteamiento del problema

En su obra de 1979, Brandt defiende que el sistema moral que sostendría un individuo racional es aquel que maximice el bienestar esperado de un grupo de individuos y plantea la necesidad de concretar qué entendemos por bienestar. Esta necesidad surge de dos conclusiones suyas anteriores que parecen dirigir la atención en direcciones contrarias:

- La gente, por lo general, se propone como meta obtener una serie de estados de cosas, producir una serie de acontecimientos, que son aquellos que más desean. Esta afirmación, de carácter empírico, aparece junto a otra de carácter normativo: un individuo enteramente racional no solo querría felicidad sino también podría querer otras cosas (Brandt menciona “éxito” y “admiración de los demás”).

3 En líneas generales, se considera que las Preferencias externas solo han de ser excluidas del cálculo de utilidad general en los razonamientos de tipo moral y político, mientras las preferencias englobadas en las otras dos categorías se excluyen también del razonamiento prudencial para el cálculo de la utilidad individual.

- Un individuo racional y benevolente querría incrementar la felicidad (entendida como goce neto) de las criaturas sensibles (y no la satisfacción de los deseos de las mismas).

Brandt ve con claridad que detrás de estas dos afirmaciones hay dos modos distintos de entender “bienestar”:

Teoría de los deseos: el Principio de Utilidad nos llevaría a procurar la maximización de la satisfacción de los deseos de todos los implicados. Esta es la versión que proponía Ayer y a la que nosotros nos estamos refiriendo con el nombre de Utilitarismo de la Preferencia.

Teoría de la felicidad: en esta versión, el Principio de Utilidad dirigiría la acción hacia la maximización de la felicidad. Esta versión parece la más cercana al hedonismo y a la formulación clásica del Utilitarismo.

Aunque para Brandt la discusión tiene repercusiones prácticas mínimas, dado que habitualmente hay una relación entre lo que la gente desea y lo que la hace feliz y a pesar de afirmar que la versión más extendida en la actualidad, tanto entre filósofos como entre economistas, es la Teoría de los deseos, cree que la teoría de la felicidad tiene una mejor defensa. La defensa de Brandt toma la forma de objeciones planteadas a la Teoría de los deseos, y cuya magnitud es tal que hacen que califique esta versión de “implausible e incluso ininteligible”. Brandt cree que las objeciones son aplicables incluso a las versiones restringidas de la teoría de los deseos, que dejan fuera de consideración algunos deseos, tales como los irracionales, y no es de extrañar, pues los problemas a los que apunta no surgen de deseos o preferencias irracionales, sino de aquellas preferencias que hemos englobado en la segunda categoría y a cuyo análisis dedicamos este trabajo.

Los problemas surgen al considerar qué programa conseguiría tal maximización, es decir, cómo seguir esta versión del PU. El problema no está en seguir el PU con los deseos actuales de la gente, pero aparece cuando consideramos los deseos de los individuos a lo largo de su vida. Bajo esta consideración, la teoría es ininteligible.

Dicha consideración parece ser la requerida si pensamos que el bienestar aumenta cuanto mayor es la satisfacción de los deseos. Parece que la idea es tener en cuenta todos los deseos que tiene un individuo. Y aquí surge el problema: pensamos que algunos deseos no deben ser tenidos en cuenta. No tenemos en cuenta nuestros deseos pasados. Tampoco tenemos en cuenta algunos deseos para el futuro (si tenemos motivos para pensar que ya no existirán cuando llegue el momento de satisfacerlos).

Para Brandt, el problema tiene un doble origen:

- No solo tenemos deseos (presentes) respecto al presente sino también respecto al futuro.
- Los deseos cambian con el tiempo.

Hare en *Moral Thinking*, publicada dos años después de la obra de Brandt, establece una clasificación temporal de las preferencias que se ha convertido en una referencia clásica a la hora de tratar el tipo de problemas que ahora nos ocupa. Básicamente, establece tres grupos:

Preferencias *Now-for-now* (NN): preferencias que se tienen ahora respecto al presente

Preferencias *Now-for-then* (NT): preferencias que se tienen ahora respecto a algún acontecimiento del futuro

Preferencias *Then-for-then* (TT): preferencias que se tendrán en el futuro para el futuro

Utilizando estos términos, el problema planteado por Brandt sería que NT y TT pueden ser contradictorias en el sentido de no poder ser satisfechas al mismo tiempo debido al carácter incompatible de los acontecimientos que las satisfacen. O podría suceder que, cuando llegara el momento de satisfacer una preferencia NT, no hubiera una preferencia TT con el mismo objeto<sup>4</sup>.

Podemos tomar uno de los ejemplos de Brandt para clarificar la cuestión. Supongamos que un padre, al que suponemos enormemente interesado en el bienestar de su hijo de siete años, identifica dicho bienestar con el cumplimiento de sus deseos teniendo en cuenta los que tendrá a lo largo de toda su vida y pretende, como buen utilitarista, maximizar su satisfacción. En el momento presente, el niño tiene un deseo, que podemos suponer tan fuerte como queramos, acerca del modo de celebrar su décimo quinto cumpleaños, por ejemplo haciendo un viaje de toda la familia a Disneyland. Si el padre es partidario de la teoría de los deseos, debería tomar en cuenta este deseo de su hijo a la hora de planificar la maximización del bienestar a lo largo de su vida. Sin embargo, es posible que cuando llegue el momento de dar satisfacción a este deseo, su hijo haya cambiado de parecer y tenga otro deseo incompatible acerca de cómo celebrar su quinceavo cumpleaños. Pero ambos deseos no pueden satisfacerse al mismo tiempo y esto hace imposible una aplicación del PU en esta versión.

Podríamos preguntar qué haría un individuo racional interesado en maximizar su utilidad. Hare afirma que tendríamos dos opciones para decidir la acción racional:

- Examinar la racionalidad de mis preferencias presentes (NT y NN).
- Examinar el curso de acción que maximizará mis preferencias NN y TT (las NT quedan fuera de consideración porque, cuando llegue el momento de satisfacerlas, serán cosa del pasado)

4 Este problema no debe ser confundido con la posible irracionalidad de minusvalorar los acontecimientos futuros por el mero hecho de serlo. El propio Brandt nos previene contra tal confusión.

Ambas bien pueden, como en este caso, llevar a conclusiones diferentes. ¿Cuál de las dos cuestiones es racional plantear? Depende de qué preferencias quieras maximizar: si las presentes, tendrías que elegir la primera opción, si las NN y TT, deberías elegir la segunda.

Para Hare, esta disyuntiva entre opciones alternativas ilustra la diferencia entre las dos versiones utilitaristas de la felicidad y de las preferencias. La felicidad puede definirse en términos de preferencias NN y TT: el hombre más feliz es el que tiene más de lo que quiere en el momento en que lo quiere. La versión preferencialista tomaría en cuenta todas las preferencias. En este sentido, la versión de la felicidad podría reformularse como una versión preferencialista restringida.

Detrás de estas dos opciones y de las correspondientes versiones del utilitarismo se esconden dos teorías distintas acerca de la racionalidad práctica. En efecto ¿cuál es la acción racional? Según Hare, la teoría elaborada por Brandt afirma que la que deriva de nuestras preferencias racionales presentes, es decir, de las preferencias NN y NT. El problema es que entonces la acción racional puede no conducir a la maximización de nuestro bienestar (y por supuesto, tampoco del bienestar general), pues esta se define sobre la satisfacción de las preferencias NN y TT. Ese es el motivo por el que Brandt cree que, aun en contra de su propia teoría de la racionalidad, debemos abandonar la versión preferencialista por una versión de la felicidad o, tal como Hare la llama, una versión preferencialista restringida.

La respuesta de Brandt abriría entonces, tal y como el mismo veía con claridad, una brecha indeseable entre teoría de la racionalidad y teoría moral: seríamos preferencialistas en la primera y partidarios de la teoría restringida en la segunda. Claro que podríamos seguir siendo utilitaristas en ambos ámbitos, pero a costa de un pequeño juego de malabarismo: en el ámbito prudencial, definiríamos la utilidad tomando en cuenta unas preferencias y en el ámbito moral, tomando en cuenta otras. Pero, trucos aparte, y puesto que el preferencialismo no tienen porqué conducir a la felicidad, resultaría que el código moral utilitarista nos impulsaría a procurar la felicidad general mientras que la teoría de la racionalidad individual nos llevaría a no procurar la propia. Los utilitaristas habríamos pasado de basar el hedonismo ético en el hedonismo psicológico (es un hecho que usted busca la felicidad: busque también la felicidad general) a defender un hedonismo ético al tiempo que mantenemos una teoría no hedonista de la conducta racional (es posible que usted no busque la felicidad y es perfectamente racional no haciéndolo, pero busque la felicidad general). Teniendo en cuenta que ahora se trata de dos teorías de carácter normativo, no parece que hayamos ganado con el cambio<sup>5</sup>.

<sup>5</sup> Habría en este caso dos alternativas. Podríamos intentar mostrar que la búsqueda de algo que no nos haga felices es irracional o bien defender que la diferencia entre el fin propuesto por las

Según Hare, un preferencialismo restringido evita las complicaciones de la versión preferencialista aludidas por Brandt, y se muestra partidario de una teoría de este tipo. Pero tal teoría plantea un problema, precisamente debido a su carácter restringido: ignoraría algunas de nuestras preferencias presentes, las preferencias NT. Parecería entonces que estamos renunciando a definir la acción racional sobre (todas) las preferencias presentes. Según Hare el problema se soluciona con un *requisito de prudencia*: debemos tener siempre una preferencia presente dominante por la maximización de nuestras preferencias NN y TT. Esto hará que siempre prefiramos maximizar nuestra felicidad. Permitir que la preferencia NT prevalezca sobre la TT es un tipo de imprudencia (y llama *autofanático* al que muestra una imprudencia de este tipo<sup>6</sup>). La brecha abierta por Brandt entre teoría de la racionalidad y teoría moral se cerraría con lo que Hare llama un "supuesto simplificador" (*simplifying assumption*): en el cálculo moral solo consideraríamos aquellas preferencias de los demás que se mantendrían si fueran prudentes en el sentido anterior.

#### PREFERENCIAS TEMPORALES Y RACIONALIDAD

Las argumentaciones de Hare y Brandt tienen en común no tomar en cuenta preferencias NT (o al menos no todas) en el cálculo moral/político, Hare dejando fuera las preferencias imprudentes y Brandt considerando solo aquellos deseos ajenos cuya satisfacción produce la felicidad de los individuos. Para Brandt, el argumento es claramente referente a la felicidad. El cumplimiento de preferencias NT en esos casos no nos satisface, no nos hace felices. ¿Por qué, entonces, un individuo racional tomaría en cuenta sus preferencias NT? Según Brandt, porque las preferencias TT no tienen (o al menos pueden no tener) la fuerza motivacional suficiente para desbancar a las NT. Para Hare, un individuo prudente sería el que estuviera motivado por sus preferencias TT. La cuestión es si podemos afirmar que un individuo racional sería prudente.

*El requisito de prudencia* (Hare). Este requisito parece añadir algo a la teoría de la racionalidad de Brandt, que por otro lado Hare admite en buena medida. Creo interpretar correctamente a Brandt al afirmar que, según la teoría de la racionalidad por él desarrollada, la imprudencia, en sentido de Hare, no es irracional<sup>7</sup>. Su ventaja añadida reside en que, conservando la teoría de la racionalidad de Brandt que se basa en

---

dos ramas de la teoría se debe al requisito de benevolencia que encontramos en la teoría moral y no en la teoría de la acción racional. Lamentablemente, no me resulta posible desarrollar en este momento tales posibilidades.

6 Tal adjetivo tiene como finalidad resaltar las similitudes del imprudente con el fanático moral. En este sentido, podría decirse que un autofanático no muestra benevolencia consigo mismo (ver nota anterior).

7 Brandt (1979), pp.81 y ss.

los deseos presentes, permite la consideración de las preferencias TT e incluso su dominio sobre las NT. Sin embargo, la ventaja no es tan evidente. Pongamos un caso relativamente fácil: cuando llegue el invierno, preferiré comprar una estufa. Pero ahora prefiero aprender a bailar el tango este invierno. La preferencia TT juega, según Brandt, un papel importante a la hora de criticar la racionalidad de la preferencia NT. Saber que tendré una preferencia determinada en el futuro es parte de los hechos<sup>8</sup>. El propio Brandt apunta en esta dirección<sup>9</sup>. Puede suceder que al representarme con la necesaria viveza mis preferencias TT mis preferencias NT cambien (se hagan equivalentes en su objeto a las TT correspondientes: prefiero ahora para él futuro lo que se que preferiré en ese momento). Hare desestima esta posibilidad y considera, posibilidad también admitida por Brandt, que, tras la crítica racional, me encontraría con dos preferencias tipo NT contradictorias: la NT original, que sigue reteniendo además su grado de fortaleza relativa, más la nueva NT surgida de la crítica racional. Esto sin embargo puede no ser así. De hecho, si la crítica ha dictaminado la irracionalidad de la NT originaria, esta desaparece o se debilita, (al menos según Brandt). De no ser así, mi conjunto de preferencias presentes contiene preferencias contradictorias y tendré que ordenarlas de la forma habitual, atendiendo a su fuerza relativa.

En este caso, el requisito de Hare cumpliría una función, pero operaría como una preferencia de segundo orden, introduciendo las complicaciones habituales. Tiene, además, otro inconveniente y es que no parece que pueda demostrarse que un individuo racional deba ser prudente. Para hacerlo, tendríamos que poder afirmar que a un individuo prudente le va mejor en la vida. Al llegar a este punto, parece que volvemos al principio, pues sería necesario ahora dar una interpretación a la expresión “que la vida de uno vaya lo mejor posible” y tendríamos que optar por una lectura en términos de satisfacción de los deseos o en términos de felicidad<sup>10</sup>.

Una posible solución sería analizar si podemos dar un tratamiento en el cálculo individual de las preferencias NT sin recurrir al requisito

8 Si consideramos que hablar de preferencias TT ya es un poco aventurado, hablar de la fuerza relativa de una TT con respecto a una NT puede parecer excesivo. Sin duda, algo puede decirse dependiendo de muchos factores (cercanía del futuro, lo común de la preferencia etc.), aunque parece que el factor de incertidumbre que inevitablemente rodea todo lo referente al futuro se redobla cuando intentamos precisar grados.

9 Brandt (1998) p. 70.

10 Además, el requisito de prudencia nos hace elaborar nuestra teoría moral no sobre las preferencias reales de la gente sino sobre unas *preferencias hipotéticas* que tendría si fuera prudente según el requisito, cosa que puede resultar problemática. No puedo considerar aquí este problema, que por su importancia requiere mucho más tiempo y espacio. Pero si consideramos que también podemos trabajar en el cálculo moral y político con preferencias hipotéticas en otro sentido, i.e., las que el agente tendría si hubiera sometido sus preferencias a una crítica racional, es fácil darse cuenta de los pasos que estamos dando por el peligroso camino de alejarnos de las preferencias reales de los demás.

de prudencia. Dado que las preferencias NT aluden a dos momentos del tiempo distintos, al preguntar si se deben tomar o no en cuenta podemos ver la cuestión desde dos puntos de vista temporalmente distintos:

**El presente** Momento N (en el que se tiene la preferencia). La pregunta se transforma en ¿debemos poner los medios/realizar las acciones necesarias para facilitar/posibilitar que la preferencia NT se satisfaga en el futuro? El requisito de prudencia de Hare tiene como efecto minusvalorar en el presente las preferencias NT a favor de la preferencia TT correspondiente. Podemos preguntarnos si un individuo racional sería prudente en este sentido. El término “prudencia”, con el que Hare denomina este requisito, lo hace indudablemente atractivo. Sin embargo, plantea algunas cuestiones que considero de difícil solución. Las preferencias NT las tengo yo ahora. Las preferencias TT no las tiene nadie. Yo tengo algún motivo para tomarlas en cuenta bajo la suposición de que serán mías. Esta suposición se debilita si consideramos los casos problemáticos de TT contrarias a las NT. Alguien en el futuro tendrá preferencias contrarias a las que yo tengo ahora. Como no podemos entrar en la interesante pero ardua cuestión de la identidad personal, admitamos que, en algún sentido, seré yo. Pero ahora yo soy yo en sentido pleno. ¿Por qué he de dar menos importancia a mis preferencias presentes? ¿incluso en los casos en los que la fuerza de la preferencia NT es mucho mayor que la de la TT? Se trata además de minusvalorar mis preferencias frente a las de alguien con quien puedo no simpatizar en absoluto. Después de todo, sus preferencias son contrarias a las mías y este no es un buen principio para una buena amistad,

Brandt cree que podría proponerse el tener en cuenta solo los deseos que se tienen para el presente y/o en el momento de satisfacerlos (de hecho, en parte hacemos esto), adoptando planes flexibles. Es decir, tendríamos solo en cuenta las preferencias NN y TT, aunque considera que esta opción parece pobre y poco ambiciosa en comparación con la teoría de los deseos originaría, pues deja fuera de consideración algunos de nuestros deseos presentes (los NT). Aunque él plantea esta opción solo para el ámbito moral y político, creo que puede aplicarse también en el cálculo personal y que merece una atención mayor ya que puede proporcionar una respuesta afirmativa a nuestra pregunta en el momento N.

**Los planes flexibles** Los planes flexibles tienen sentido en los casos en los que nuestras preferencias NT son tales que podemos realizar alguna acción en el momento N. En efecto, a veces nos vemos en la necesidad de planificar la satisfacción de ciertos deseos. Quizá el deseo del niño de 7 años respecto a como celebrar su 15 cumpleaños requiera que el padre empiece a planearlo ya. Cuando llegue el momento, si el chico ya no lo desea no me empeñaré en satisfacerlo. Esto plantea un inconveniente, pues la planificación puede haber sido muy costosa y no

convertible (aunque el plan flexible puede incluir cosas del tipo “no te comprometas con planificaciones que no tienen vuelta atrás”). Pero si el chico a los 15 sigue deseándolo, y no lo he planificado, no podré satisfacer su deseo. Quizá lo invertido no permita que satisfaga sus deseos TT y no pueda celebrar su décimo quinto cumpleaños pagándole un curso de vela pues se ha quedado sin recursos después de costear el viaje a Disneyland. La realidad, a menudo, se encarga de que los planes flexibles no sean tan flexibles.

Creo sin embargo que los planes flexibles cuentan con dos ventajas. En primer lugar, reflejan un hecho indudable. Las medidas tomadas para la satisfacción futura pueden resultar en alguna medida gratificantes ahora: cuando se quiere algo para el futuro, se disfruta con los preparativos. Esto sucede tanto en los casos en los que la planificación es necesaria como en los que no lo es. Probablemente no es necesario que yo pase toda la semana probando ropa y zapatos para ir a bailar el sábado por la noche. Pero disfruto haciéndolo, y el placer tiene sentido porque pienso que voy a ir y ahora prefiero ir. Se que el sábado a las nueve tendré una pereza horrible, sobre todo si hace frío, y preferiré quedarme en casa. De acuerdo, cuando llegue el momento, si ya no quiero ir, no me empeñaré.

En segundo lugar, permite evitar un error que creo comete el requisito de prudencia. El requisito solo adquiere sentido si damos por sentado que sabemos ahora cuales serán nuestras preferencias TT. Y me temo que esto está tan alejado de la verdad que no puede ser utilizado ni como supuesto simplificador. Y no solo por el factor de incertidumbre que necesariamente rodea los acontecimientos futuros (¿hará tanto frío? ¿seguiré siendo yo tan friolera? El sábado seguramente sí. Pero ¿el año que viene?) sino, fundamentalmente, porque nuestras preferencias futuras dependen en parte de lo que hacemos ahora. No es posible ahora desarrollar este punto, pero baste considerar que planificar (innecesariamente) ahora e invertir (comprarme quizá unos zapatos de baile nuevos) contribuye a incrementar la probabilidad de que no surja la temida (por mi ahora) preferencia TT. Puede que así el sábado me parezca más a la que quiero ser.

Podemos concluir nuestro análisis del momento N diciendo que un individuo racional tendrá motivos para tomar en cuenta sus preferencias NT en los casos en que

- La planificación para la satisfacción de la preferencia en el momento T, sea o no estrictamente necesaria, le proporcione una satisfacción presente y/o ignorarla le suponga una insatisfacción
- Las preferencias previstas TT sean de tal modo contrarias a las preferencias NT que el individuo no se sienta identificado con aquellas e intente alterarlas mediante la toma de medidas pre-

sententes que garanticen o faciliten que sus preferencias TT sean lo más similares en contenido a sus preferencias NT<sup>11</sup>. O, dicho de otro modo, que llegado T haya una NN “correspondiente” (e.d., con el mismo objeto) a la actual NT.

Esto presenta a mi juicio una ventaja sobre la propuesta de Hare de ignorar nuestras NT en la momento N.

**El futuro** Momento T (en el que se ha de satisfacer la preferencia). La pregunta se transforma en ¿tengo alguna razón para satisfacer mis preferencias pasadas (o las de los demás)? ¿tengo motivos para preocuparme de personas que, en un sentido, ya no existen? Esas preferencias, esos deseos pasados ya no existen. ¿Tiene algún sentido preguntarse acerca de la posibilidad de satisfacer deseos que ya no existen? Una versión del utilitarismo centrada en la felicidad ignoraría las preferencias pasadas, al igual que lo hace el preferencialismo restringido que Hare defiende. El motivo de esta restricción es que las preferencias pasadas no tienen ninguna relación con mi felicidad (ni con la de nadie).

Sin embargo, este acuerdo mayoritario entre los utilitaristas en no tomar en cuenta las preferencias pasadas ha sido objeto de numerosas críticas que merecen consideración aparte, fundamentalmente porque iluminan la diferencia entre las versiones preferencialistas y de la felicidad.

#### LAS PREFERENCIAS PASADAS

La felicidad es un estado de un sujeto. La satisfacción de las preferencias no lo es: las preferencias son relativas a estados del mundo. Esos estados del mundo pueden o no provocar un estado de un sujeto (el que llamamos felicidad u otro). “Felicidad” en este sentido es una buena traducción, probablemente la más adecuada, de lo que los utilitaristas clásicos llamaron placer. Siguiendo con los clásicos, llamamos útiles a aquellas cosas que, debido a una propiedad suya que llamamos utilidad, producen placer o felicidad.

¿Tienen algo en común la felicidad y la satisfacción de los deseos? Tomemos “felicidad”, tal y como hacía Bentham y nosotros hemos estado haciendo, como un sinónimo de “placer”. Brandt<sup>12</sup> explica el significado de “placentero” diciendo que se trata de un término utilizado para calificar experiencias que contienen un elemento subjetivo (Brandt, TE, C.12,3). Este elemento subjetivo es aquello que llamamos “placer”. Podemos decir por tanto que el **placer** es un cierto **estado mental**, en el sentido de que llamamos placer a ese elemento subjetivo. Es el correlato subjetivo

11 En estos casos, la manipulación sería similar a la presentada por la estrategia de atarse a uno mismo (Elster, 1979)

12 Brandt (1959), cap. 12

de un tipo de experiencias de un sujeto, a las que, precisamente por contener ese elemento subjetivo, llamamos placenteras<sup>13</sup>.

Veamos qué sucede con la satisfacción de los deseos. La satisfacción de los deseos de un sujeto puede estar ligada a un estado mental en el sentido siguiente: si yo deseo algo, por ejemplo asistir a una representación de *Aida* en el Teatro del Liceo, la satisfacción de mi deseo hace referencia a un cierto estado mental: el de estar allí viviendo la experiencia (este es el estado mental que se desea), “contemplando” cómo mi deseo se satisface (cómo se produce el estado de cosas que satisface mi deseo). Digamos que se trata de un estado subjetivo de conciencia de que una determinada experiencia, que satisface un deseo mío, está teniendo lugar. Casi todos los deseos que tenemos son tales que su satisfacción incluye un estado mental.

Además, casi todos nuestros deseos son condicionales a su propia persistencia<sup>14</sup>: deseamos que se cumplan en la medida en que sigan deseándose. Casi todos los deseos ligados a un estado mental son de este tipo. Cuando se convierten en deseos pasados, en virtud de su carácter condicional, desaparecen en el sentido y por tanto ya no pueden ser satisfechos. Puede producirse un estado de cosas que corresponda al objeto de un deseo que ya no persiste en el presente, pero tal estado ya no supone la satisfacción de ningún deseo en el momento en que tal estado sucede.

Utilizando los términos de Brandt<sup>15</sup>, podemos decir que el hedonismo es una teoría acerca del valor de tipo personalista, entendiendo que un personalista “es alguien que considera que sólo los estados de bienestar poseen valor intrínseco”. Esto no puede afirmarse sin matices de la teoría de la satisfacción de los deseos. Esto sucede porque, mientras que no tiene sentido decir que alguien está teniendo una experiencia placentera pero no lo sabe (o que siente placer pero no lo sabe), sí lo tiene decir que los deseos de alguien se satisfacen pero que él/ella no lo sabe. El preferencialismo sería personalista si afirmamos que los únicos deseos cuya satisfacción tiene valor intrínseco son aquellos cuya satisfacción tiene una repercusión en el bienestar del sujeto (ahora usamos “bienestar”

13 Hay otro sentido en el que podemos decir que el placer es un estado mental: cuando alguien tiene una experiencia placentera y siente placer esto se corresponde con un cierto estado que se produce en nuestro cerebro. Sin embargo, esto es un uso impropio del término “estado mental”, y está mejor recogido por la expresión “estado cerebral”. Es habitual, y no sólo entre los filósofos, distinguir “cerebro” y “mente” (psicólogos, psiquiatras y neurólogos también utilizan esta diferencia. Una rama entera de la neurología, la neuropsicología, se encarga de estudiar las relaciones entre ambas). Que son distintos, al menos conceptualmente, no parece sujeto a discusión (incluso aunque algunos mantengan la tesis materialista extrema, habitualmente denominada tesis de la identidad, de que necesariamente un determinado estado cerebral y un determinado estado mental aparecen indisolublemente unidos).

14 Parfit los llama “deseos implícitamente condicionales a su propia persistencia” (1986, p.151).

15 Brandt (1956), Cap.2 Apar. 3

solo para referirnos a una experiencia que incluye un estado mental). Por ejemplo, la propuesta de Brandt (TRG) acerca del "plan flexible" de la satisfacción de los deseos, da lugar a una teoría en la que lo que se tiene en cuenta son los deseos cuya satisfacción tiene una clara repercusión en un cierto estado mental de felicidad del sujeto que tiene tales deseos. Un preferencialismo utilitarista sería personalista: queremos satisfacer los deseos de la gente, incluso los nuestros, porque habitualmente eso produce un estado mental al que llamamos felicidad. Por tanto, el UP sería necesariamente un UPR pues dejaría fuera de consideración al menos un tipo de preferencias, las pasadas.

¿Hay algún problema para dejar fuera de consideración las preferencias pasadas? Muchos autores piensan que sí lo hay, alegando que no todos los deseos que tenemos son condicionales a su persistencia ni están ligados a un estado mental.

Afirmar esto implica atribuir un valor intrínseco a la satisfacción de las preferencias, y, en el sentido anterior, suscribir una teoría no personalista del valor. El argumento esgrimido es que no todos nuestros deseos tienen tal conexión con la felicidad, y se acude a presentar como contraejemplos casos cuya validez me parece cuestionable:

#### LOS DESEOS DE LOS MUERTOS

En ocasiones, hablamos de satisfacer los deseos de alguien cuando tal satisfacción no puede traducirse en términos de estado mental. Hacemos cosas que satisfacen el deseo de un difunto y decimos "era su deseo". Estos casos pueden ser de dos tipos:

- Deseos que al ser formulados y tenidos incluyen una referencia necesaria a una satisfacción que se traduce en un estado mental, y que en su mayoría son además condicionales.

Tal es habitualmente el deseo anteriormente mencionado de asistir a la representación de *Aida*. Puede que, si pasara mi vida sin poder satisfacer ese deseo, y si hacerlo no estuviera penado por la ley, algún amado deudo me llevara embalsamada a un palco y dijera "era su deseo". Casi con toda seguridad podríamos decir que el deudo se equivoca. Yo no considero que esa situación satisfaga mi deseo. Esto también sucede si me lleva al palco viva pero en estado de coma, o inconsciente. No considero que eso sea una satisfacción a mi deseo. En estos casos, podríamos decir que los deseos, si se formularan expresa y detalladamente, incluirían una referencia a un estado mental (no formularlos así les produce muchos disgustos a los protagonistas de cuentos que buscan la satisfacción de sus deseos en un genio que toma sus palabras al pie de la letra y se niega, cosa que es habitual en ellos, a sobreentender nada. Les disculpamos porque son genios, pero no intentan satisfacer

los deseos de los humanos sino tomarles el pelo). No dudo que alguien puede tener razones para cumplir deseos de este tipo, pero se trata en todos los casos de sus propios deseos presentes, que sí se traducen en un estado mental del que los satisface.

- Deseos que al ser formulados y tenidos incluyen una referencia necesaria a una satisfacción que excluye la posibilidad de ser traducida en términos de estado mental.

Son del tipo que podríamos llamar “últimas voluntades”: deseo ser incinerada tras mi muerte y que mis cenizas se esparzan por un paisaje especialmente querido para mi. Por definición, no son tampoco condicionales. Podría pensarse que no todos los seres humanos tienen deseos de este tipo, y que es frecuente oír decir “que más me da a mi, si ya no me voy a enterar. Que hagan conmigo lo que quieran. Sin duda, pero sí es enormemente frecuente tener otros deseos de esta categoría: los relativos a la disposición de los bienes. Lo infrecuente es encontrar alguien que sea totalmente indiferente ante este asunto. En cualquier caso, existen y no son raros en absoluto: casi todos tenemos algún deseo de este estilo, aunque probablemente una parte mínima de nuestros deseos sean así<sup>16</sup>. Cuando son propiamente pasados ¿tenemos (los demás) motivos (utilitaristas) para satisfacerlas? Las habituales están relacionadas con su repercusión en estados mentales de otras personas o del propio ejecutor del deseo<sup>17</sup>.

**Mis viejos deseos.** Puede que yo tuviera en el pasado un deseo NT no condicional en su formulación, por ejemplo de contribuir en el futuro a la erradicación del SIDA. Puedo incluso tomar medidas para evitar que un posible yo egoísta futuro prefiera otra cosa llegado el momento. Pero, de nuevo, la cuestión es si hay alguna razón para satisfacerlo cuando pertenece al pasado. Los ejemplos como este, que son los habituales, pueden inducir a confusión pues, después de todo se trata de realizar algo que casi todos nosotros consideramos deseable y noble. Pero hay otra vertiente: el caso “morir joven y dejar un bonito cadáver”, piense lo que piense la maldita burguesa en que me voy a convertir. Si, tal y como creo, nuestra intuición cambia en este caso es porque no está tan claro que tales deseos NT tengan algún sentido vistos como pasados. Claramente, pertenecen a alguien que ya no existe: esa boba idealista,

16 Y estamos dispuestos a realizar considerables gastos, económicos y no económicos, para asegurar su satisfacción. Puede haber razones para iniciar acciones y poner los medios para su satisfacción vistas desde el punto de vista del N (y quedan sujetas a lo dicho anteriormente).

17 Estos casos hay que tratarlos con cuidado porque se mezclan, obviamente, con preferencias **externas**, altruistas o malevolentes que tienen su complicaciones propias. Un trabajo posterior debería tratar sobre la conexión preferencias pasadas y externas. Tiendo a pensar que no se puede dejar al margen unas sin dejar también al margen las otras. Otras, como las cenizas, no son externas. Podemos plantear la racionalidad de tales deseos, pero eso está fuera del propósito de este trabajo.

esa joven tremebunda, por la que yo puedo sentir cosas muy distintas, desde la nostalgia condescendiente (en fin, el que de joven no es de izquierdas...), al desafecto más o menos irónico (casi se sale con la suya: ya murió, aunque no dejo ningún cadáver. Qué razón tenía quien dijo no te fíes de nadie que tenga más de 30. Tendría que haberse suicidado a los 29.) Pero dudo que pueda darse ninguna razón para que la burguesa en cuestión se sienta ligada, salvo por la nostalgia, a ese deseo. Y desde luego no veo ninguna razón de carácter utilitarista para satisfacerlo.

Así pues, ninguno de estos tipos de deseos pasados presentan problemas para ser rechazados, en tanto que tales, desde el punto de vista utilitarista. Su aparente desafío, en tanto que casos de deseos que habitualmente creemos que tenemos motivos para satisfacer, se resuelve al considerar que tales motivos, cuando existen tienen relación con deseos presentes.

#### EN DEFENSA DE LA TEORÍA DE LA FELICIDAD

Un utilitarista no da valor intrínseco a la ocurrencia de eventos que satisfacen algún deseo. Creo que este es el principal motivo por el que el utilitarismo ha tendido a versiones de preferencialismo restringido. En este trabajo nos hemos centrado en el análisis de algunas de las preferencias que se dejan fuera de consideración. En el caso de las preferencias pasadas, no parece que tengamos ningún motivo para tomarlas en cuenta al margen de su repercusión presente en alguien.

Desde el momento en que renunciamos a tener en cuenta algunos tipos de preferencias, como las pasadas (o en otros casos las irracionales) debido a su falta de conexión con un estado de bienestar de un sujeto, estamos más cerca del hedonismo que del preferencialismo a nivel de teoría axiológica. Es decir, estamos más cerca de lo que Brandt llama teoría de la felicidad que de la teoría de la satisfacción de los deseos. Probablemente, la teoría de la felicidad es la más acorde con el utilitarismo, al menos si este es personalista en lo relativo a la identificación del valor intrínseco, como yo creo que lo es. El propio Hare mantiene su teoría como una cierta versión de la felicidad. Pero yo creo que, para mantener esta postura de una forma coherente, hay que dejar fuera de consideración no solo las preferencias pasadas sino también las futuras. En la teoría de la racionalidad, es decir, en la teoría normativa que indica como debe calcular su utilidad personal un individuo racional, parece poco discutible que solo las preferencias presentes (NN y NT) deben tener un peso directo. Esto es tan difícil de discutir que Hare, al intentar dar un peso a las preferencias futuras (TT) tiene que hacerlo introduciendo una preferencia presente dominante.

Si pensamos en los argumentos esgrimidos, veremos que se admite que lo que le da valor a la satisfacción de las preferencias es la felicidad

producida. El objetivo normativo es la felicidad, no la satisfacción de las preferencias. Cabría entonces preguntarse si tiene mucho sentido hablar de preferencialismo. Creo que ignorar algunas preferencias (las pasadas, las irracionales) nos aleja del preferencialismo y nos acerca a la teoría de la felicidad (o dicho de otro modo, nos aleja de la identificación felicidad = satisfacción de deseos): puesto que sólo la satisfacción de algunas preferencias produce felicidad, como utilitaristas sólo tenemos razones para satisfacer esas preferencias. Cabría entonces preguntarse si tiene mucho sentido seguir empleando el término preferencialismo o si por el contrario no habremos vuelto a una versión de la felicidad.

¿De dónde viene esta resistencia de los utilitaristas a abandonar la versión preferencialista? Una posible explicación está relacionada con el modo en que los utilitaristas clásicos defendían el hedonismo ético. Sin embargo, que tal defensa se basara en una teoría psicológica dudosa (HP) no es motivo para abandonar el hedonismo ético. Ayer decía que, si bien el rechazo/falsedad del hedonismo psicológico no condena al hedonismo como teoría normativa, es dudoso que nadie quiera mantenerlo como tal sin el apoyo de la teoría psicológica<sup>18</sup>. Pero se equivocaba. Brandt apunta la posibilidad de mantenerlo como una propuesta normativa<sup>19</sup>, y diseña las líneas por las que tal defensa podría transcurrir. (hasta donde yo sé es lo que hacían los epicureos). Otra línea de defensa, también iniciada por Brandt, es la seguida en este trabajo: mostrar los defectos de la versión preferencialista, de los que se ve libre la versión de la felicidad<sup>20</sup>.

Otro motivo puede ser el siguiente: parte de lo que nos importa no puede ser traducido en una experiencia del agente o al menos la experiencia del agente derivada de la satisfacción de esas preferencias es, en tanto que experiencia, menos placentera que otras alternativas. Son las preferencias externas. Pero las preferencias externas parecen problemáticas a nivel moral y político<sup>21</sup>. Si esto es así, resultaría que a ese nivel deberíamos abandonar el preferencialismo, aunque quizá no a nivel individual. Tendríamos que aceptar, aunque por otro motivo, la brecha que introduce Brandt entre las dos teorías normativas básicas (teoría moral y teoría de la conducta racional).

Quizá la brecha desaparecería si pudiéramos desestimar también las preferencias externas a nivel individual (por ejemplo, mostrando que son irracionales). Este camino no suena muy prometedor. Se puede plantear así: ¿podríamos recomendar que la gente abandonara sus preferencias

18 Ayer (1954)

19 Brandt (1959) Cap. 12.

20 Brandt (1979), Cap. 13.

21 Las preferencias externas, definidas por Dworkin (Dworkin 1977, cap.9) como aquellas relativas a la asignación de bienes y oportunidades a personas distintas de aquel que tiene la preferencia, plantean problemas que van desde desafíos a la idea de trato igualitario, cuando son benevolentes, a el reto que supone tener que tomar en cuenta preferencias malevolentes.

externas (y no solo que no se tengan en cuenta en el razonamiento moral y político)? Probablemente no. Debo reconocer que, por el momento, no se como salir de este problema.

Quizá, después de todo, no sea tan problemático que la ética y la política dejen fuera o incluso luchen contra cosas que no son irracionales (y puede que ni siquiera imprudentes). De hecho, hay quien dice que para eso queremos la ética y la política (para que nos saquen de los dilemas y las guerras de todos contra todos): lo que tu tienes razones para hacer, más lo que tengo razones para hacer yo, nos pueden llevar a un situación nada deseable.

No debemos tampoco olvidar que la versión de la felicidad no solo puede defenderse destacando los problemas de la versión preferencialista, sino que cuenta con méritos propios. Tal y como señalaba Brandt, cuenta con un apoyo intuitivo (cuando nos interesamos por alguien, como nuestros hijos, parece que lo que queremos es que sean felices, y nos preocupamos por que satisfagan sus deseos solo secundariamente, en tanto que suponemos que esto les hará felices) y la teoría psicológica que afirma que tenemos un sentimiento al que llamamos benevolencia dice que lo que nos motiva por simpatía es la felicidad ajena (entendida como estado mental, placer y ausencia de dolor).

Todas estas cuestiones deben ser planteadas y debatidas y debe procurarse dar una visión general de los problemas relativos a todas las categorías de preferencias desestimadas. Pero, al menos en lo que respecta a las preferencias pasadas, creo que podemos afirmar algo. Supongamos que podemos elegir que exista o continúe existiendo uno de estos dos mundo: en el mundo 1 existen unos seres que logran sólo la satisfacción de aquellas preferencias cuyo objeto es el placer (o, si queremos, que consisten en experiencias). En el mundo 2 no hay seres vivos: todos han muerto, pero se realizan los objetivos de las preferencias NT de los desaparecidos. ¿Cuál prefieres?<sup>22</sup> No hay duda de que un utilitarista se decantaría por el primero.

#### REFERENCIAS BIBLIOGRÁFICAS

- Ayer (1954) "El principio de Utilidad" en *Ensayos filosóficos*,  
 Brandt (1959) *Teoría ética*, Alianza Editorial, Madrid 1982.  
 Brandt (1979) *A Theory of the Good and the Right*. Oxford, Clarendon Press.  
 Brandt (1998) "The Rational Criticism of Preferences", en *Preferences*, Fehige y Wessels (Eds), de Gruyter, Berlin- New York.

22 El ejemplo está tomado de T. Ronnow-Rasmussen (1998)

- Dworkin (1977) *Taking Rights Seriously*, Duckworth, London.
- Elster (1979) *Ulysses and the Sirens*, Cambridge University Press, 1979.
- Hare (1981) *Moral Thinking: Its Levels, Method and Point*. Oxford, Clarendon Press.
- Hollis y Sudgen (1993) "Rationality in action", *Mind* 102.
- Parfit (1986) *Reasons and Persons*, Oxford University Press.
- Ronnow-Rasmussen (1998) "Preference-utilitarianism and Past Preferences", en *Proceedings of the Twentieth World Congress of Philosophy*, Boston, Massachusetts. Disponible en <http://www.bu.edu/wcp/Papers/Valu/ValuRonn.htm>