

EL UTILITARISMO DE SATISFACCIÓN
DE PREFERENCIAS
Y EL «PRINCIPIO DE SIDGWICK»

JUAN RODRÍGUEZ LARRETA

Sociedad Argentina de Análisis Filosófico, Buenos Aires

ABSTRACT

This paper has two parts. In the first I shall try to challenge a claim made by M.D. Farrell in his «La satisfacción de preferencias y los límites del utilitarismo» (Τέλος, Volumen IV, N.º 1, 1995, España). Farrell claims that there is a plausible version of preference satisfaction utilitarianism which is compatible with an implicit utilitarian principle he called «Sidgwick's Principle».

In the second part I argue that if, as I think, preference satisfaction utilitarianism turns out to be incompatible with this principle then its intuitive plausibility diminishes considerably.

En este trabajo efectuaré una breve crítica de la variante actualmente más en boga del utilitarismo que es el *utilitarismo de satisfacción de preferencias*, (en adelante USP). Para ello me basaré en un análisis del mismo efectuado por Martin Farrell en su trabajo «La satisfacción de preferencias y los límites del utilitarismo», aparecido en el volumen IV, Número 1 (1995) de esta revista. El objetivo central del trabajo de Farrell es analizar la relación entre dicha postura y un principio implícito del utilitarismo que él denomina «principio de Sidgwick» (en adelante PS). Primero trataré de mostrar, contra la opinión de Farrell, que todas las caracterizaciones interesantes del USP violan dicho

principio, tal como él lo utiliza. Luego intentaré destacar que dicha doctrina, por el hecho de violarlo, choca con nuestras intuiciones morales, lo cual disminuye considerablemente su plausibilidad.

EL PRINCIPIO DE SIDGWICK

Farrell presenta el principio diciendo:

«Si examinamos cuidadosamente aquellos resultados que se consideran usualmente buenos, nada posee esa característica de bondad si no tiene relación con la existencia humana. Puede considerarse que algunos objetos inanimados, paisajes, etc., son buenos porque poseen belleza, mientras otros son malos porque resultan horribles; pero no es racional producir belleza sino para su posible contemplación por seres humanos.»

Esta formulación no es precisa y puede dar lugar a distintas interpretaciones. Pero si entendemos «bondad» como «utilidad» y si enfatizamos la última parte del texto citado en que se habla de la «contemplación por seres humanos», surge una interpretación que podríamos denominar «variante fuerte». Según la *variante fuerte*, los hechos que no son conocidos y por lo tanto no inciden en ningún estado mental de un sujeto no deben ser tomados en cuenta para calcular su utilidad (por lo cual, si ese sujeto hubiese vivido en un mundo diferente, en tanto sus estados mentales hubiesen sido los mismos que los que tiene en este mundo, su utilidad sería la misma). Esta variante no sólo concuerda con el utilitarismo tradicional «de estados mentales» que comparte Sidgwick, sino que es la que a mi juicio utiliza Farrell, y sobre todo, es la que él requiere para poder decir que una de las versiones del USP lo viola. Dicha *variante fuerte* será pues la que emplearé en adelante.

EL ANÁLISIS DE FARRELL

Del trabajo de Farrell surge que, bajo el rótulo de «utilitarismo de satisfacción de preferencias», se esconden tres teorías axiológicas que difieren sustantivamente entre sí.

La *versión subjetiva*, según la cual lo único relevante para determinar la utilidad (positiva o negativa)¹ de un sujeto son sus *creencias* acerca de si sus preferencias se satisfacen o no.

La *versión objetiva*, según la cual lo único relevante para determinar la utilidad de un sujeto es la satisfacción o no satisfacción *real* de sus preferencias, con independencia de que el sujeto lo crea o no lo crea.

La que podría denominarse *versión mixta*, según la cual lo único relevante para determinar la utilidad de un sujeto son sus *creencias* acerca de si sus preferencias se satisfacen o no pero sólo cuando éstas van unidas a la satisfacción o no satisfacción *real* de sus preferencias, en otras palabras, lo único relevante respecto de la utilidad de una persona son sus *creencias verdaderas*.

Farrell afirma, con razón, que la *versión subjetiva* del USP se reduce a una variante del *utilitarismo de estados mentales* (así se denomina al género que incluye a todas las versiones del utilitarismo de corte hedonista o eudemonista) y afirma también que éste sucumbe ante la conocida objeción de Nozick, basada en su «máquina de la experiencia»², por lo cual dicha versión debe ser abandonada. En cuanto a la versión objetiva, Farrell afirma que viola el PS, ya que una preferencia objetiva puede ser satisfecha o no satisfecha por un estado de cosas que *no es conocido* —«*no es contemplado*»— por el sujeto y que, por lo tanto, puede no incidir en absoluto en ninguno de sus estados mentales (nótese que aquí Farrell utiliza la versión fuerte del PS). El autor se inclina por la versión mixta, que exige lo que pide la versión objetiva (que la preferencia sea, de hecho, satisfecha o no satisfecha) pero también exige lo que pide la versión subjetiva (que el sujeto crea que la preferencia fue satisfecha o no satisfecha). Según Farrell, dicha

¹ En adelante «utilidad» debe entenderse como una abreviatura de «utilidad positiva o negativa».

² La objeción, contenida en su libro *Anarchy, State and Utopia* (Oxford, Basil Blackwell, 1974, pp. 42-43), se basa en imaginar una máquina que estimula a nuestro cerebro produciendo una corriente vivencial semejante a la de la vida real, y que puede ser programada de modo tal que, al ingresar en ella, nos asegure una vida ilusoria pero más placentera que la real.

Nozick nos pregunta ¿ingresaríamos en esa máquina? y contesta «sin duda no». En su opinión la justificación de nuestra negativa es que la calidad de nuestras experiencias no es lo único valioso: también es valioso «hacer ciertas cosas», «ser cierto tipo de persona». Pero entonces, el utilitarismo tradicional -que es el utilitarismo de estados mentales-, al sostener que sólo los estados mentales poseen valor intrínseco, estaría equivocado.

versión, por tener un componente objetivo, no colapsaría con el utilitarismo de estados mentales y en consecuencia no chocaría con la objeción de Nozick; y por otra parte, al exigir también que el sujeto crea que la preferencia fue satisfecha o no satisfecha, no asignaría utilidades cuando existe una desconexión total entre el estado de cosas que satisface la preferencia y los estados mentales del portador de la misma, con lo cual no violaría el PS.

CRÍTICA A FARRELL

A continuación trataré de mostrar que la versión mixta que defiende Farrell también viola el PS. Como paso previo presentaré un cuadro que permite visualizar la diferencia entre las tres versiones. La notación que utilizo en el cuadro es la siguiente: los signos «+» y «-», representan respectivamente el aumento y la disminución de la utilidad, y el signo «=» los casos en que la utilidad se mantiene invariable. Uso, además, «C(P)» y «P» como abreviaturas de «S cree que la preferencia p se satisface» y «la preferencia p se satisface».

| | | Versión subjetiva | Versión objetiva | Versión mixta |
|--------------|------------|----------------------|---------------------|------------------|
| Situación 1: | C(p) y p | + | + | + |
| Situación 2: | C(p) y ~p | + | - | = |
| Situación 3: | C(~p) y p | - | + | = |
| Situación 4: | C(~p) y ~p | - | - | - |

Podemos observar que en las situaciones 1 y 4, donde hay correspondencia entre la creencia y el hecho, las tres versiones concuerdan, mientras que en las situaciones 2 y 3, donde no se da dicha correspondencia, las teorías subjetiva y objetiva asignan a la utilidad variaciones opuestas y la teoría mixta no asigna variación alguna.

Según la versión mixta, para que una creencia acerca de una preferencia incida en el cómputo de utilidad de alguien, dicha creencia debe ser verdadera, con lo cual esta versión hace depender al cómputo utilitarista no sólo de ciertos estados mentales (ciertas actitudes proposicionales) sino de la *correpondencia* de esos estados con hechos que son independientes de esos estados y que pueden no tener incidencia alguna en los estados mentales del sujeto. Pero entonces, la utilidad de ese sujeto podría depender en parte de hechos que no inciden en absoluto en sus estados mentales (y si esos hechos fuesen diferentes aunque todos sus estados mentales fueran idénticos, la utilidad sería diferente). Vemos, pues, que la versión mixta viola también el PS. Consideremos el siguiente ejemplo:

1) El mayor deseo de Pedro es acertar la lotería y un día la acierta, pero la noticia jamás le llega, por lo cual muere sin saber que la ganó (situación 3).

Según la versión objetiva, dicho suceso implica, de hecho, la satisfacción de una intensa preferencia de Pedro, por lo cual debe ser computado como algo que incrementa considerablemente su utilidad. Ello contradice el PS ya que el hecho de que lo acierte, por ser ignorado por Pedro, no afecta a ninguno de sus estados mentales. En cambio, según la versión mixta, dicho suceso, al no ser creído por Pedro, no cumple con el requisito «subjetivo» que esa versión exige, y por lo tanto no incrementa la utilidad de Pedro. Vemos que, en este ejemplo, la versión mixta aventaja a la versión objetiva en el hecho de que no viola el PS. Pero ahora consideremos este otro ejemplo:

2) Pedro nunca acertó la lotería, pero un día antes de morir escucha por la radio el anuncio erróneo de que lo ganó, lo cual llena de satisfacción su último día de vida (situación 2).

Según la versión mixta, el hecho de que el anuncio sea erróneo hace que la segunda de las dos condiciones que deben cumplirse para que la situación determine alguna utilidad —la que afirma que la preferencia debe *de hecho* ser satisfecha— no se cumpla. Pero entonces, la falsedad del anuncio hace que la utilidad que hubiese adquirido Pedro de haber sido cierto el anuncio, no se produzca (de la situación 1 del gráfico, que asigna una utilidad positiva, pasamos a la situación 2, que no modifica la utilidad. Surge así que, según esta versión, también existen hechos, en este caso la falsedad de una noticia, que deben ser tomados en cuenta en el cómputo de la utilidad de una persona a pesar de no incidir en

ninguno de sus estados de conciencia. En este segundo ejemplo, entonces, la versión mixta, contra lo que afirma Farrell, también viola el PS³.

Como resultado del análisis de Farrell y de la crítica efectuada recién, podemos concluir que existen tres versiones diferentes del USP, y que las tres presentan dificultades. En efecto, la versión subjetiva se convierte en una variante muy poco plausible del utilitarismo de estados mentales. Resulta poco plausible porque de todos los estados mentales, la teoría considera valiosos sólo aquéllos que consisten en las experiencias supuestamente placenteras derivadas de la creencia de que una preferencia ha sido satisfecha, y esta restricción es antiintuitiva y arbitraria. Por otra parte, dicha versión no sería sostenida por ningún partidario del USP. En cuanto a las otras dos (la *objetiva* y la *mixta*) ellas violan el PS y, como veremos ahora, este hecho les resta credibilidad.

LA PLAUSIBILIDAD DEL USP

Como anuncié al principio, intentaré mostrar que el USP, en sus únicas dos variantes más o menos plausibles que son la *objetiva* y la *mixta* (descarto la *subjetiva*, por las razones mencionadas recién), por el hecho de violar el PS, se torna poco verosímil. En efecto, resulta poco verosímil sostener que toda la dicha que obtiene Pedro gracias al anuncio erróneo no le proporciona ninguna utilidad. Y, en general, resulta poco plausible sostener que toda la felicidad que una persona obtiene debido a una creencia errónea carece en absoluto de valor. Supongamos, por ejemplo que, como parece ser el caso, mucha gente desea intensamente que exista otra vida mejor (donde se reencuentra con sus seres

³ Merece destacarse que la versión mixta, al exigir que la preferencia se satisfaga *de hecho* y también que se dé la creencia de que se satisfizo, no pide más que la conjunción del hecho y de la creencia en el mismo. Ahora bien, esta conjunción puede ser accidental y la creencia no ser causada por el hecho. Pero entonces, un suceso *a* que satisface una preferencia de *x*, y tal que *x* cree por alguna razón fortuita que *a* acaeció, puede cumplir con las dos condiciones que exige la versión mixta; pero ésto resulta intuitivamente extraño. (Debo esta observación a Thomas Moro Simpson).

queridos, donde hay paz, etc.) y que la creencia en esa otra vida los hace más felices. Resulta poco plausible sostener que esa felicidad tiene valor sólo si la creencia es verdadera pero carece totalmente de valor en el caso de ser falsa. Sin embargo la versión mista, al igual que la objetiva, está obligada a sostener todas estas tesis.

Del mismo modo, resulta poco plausible sostener que toda desdicha producida por una creencia errónea no ocasiona ninguna disutilidad. Si alguien cree erróneamente que padece una enfermedad mortal (se supone que ello frustra sus preferencias más importantes), y vive atormentado por esa creencia hasta que un día muere en un accidente, resulta antiintuitivo sostener que toda la angustia y el sufrimiento que ocasionó esa creencia, por el hecho de ser errónea, no produjo ninguna disutilidad. La intuición nos dice —o me dice— que el hecho de que la creencia sea errónea, no sólo no anula por completo la disutilidad que hubiese generado de ser verídica, sino que no la disminuye en un ápice. El utilitarista de satisfacción de preferencias se defenderá diciendo que la apelación a estas intuiciones «sidgwickianas» conduce al utilitarismo de estados mentales y que esta postura está expuesta a la objeción planteada por la máquina de Nozick. Creo que la observación es válida, aunque pienso que la objeción de Nozick no es tan decisiva como parece; pero dicha cuestión excede los límites de este trabajo⁴.

⁴ Deseo agradecer a Martin Farrell, Eduardo Rivera López, Thomas Moro Simpson e Ignacio Zuberbühler por sus importantes observaciones que me permitieron eliminar una serie de errores y confusiones contenidas en una versión previa.