NOTA CRITICA

JOHN STUART MILL, LA FALACIA DE COMPOSICION Y LA NATURALEZA HUMANA *
(Observaciones sobre un artículo de Eduardo Rabossi)

THOMAS M. SIMPSON

I

En su defensa de los argumentos de John Stuart Mill en favor del principio de utilidad (*Análisis Filosófico*, Vol. VII, Nº 1), Eduardo Rabossi hace varias afirmaciones discutibles, destinadas a rescatar a Mill del reino de la falacia.

Para facilitar la comprensión de estas notas reproduciré *in extenso* el revulsivo fragmento de IV(3) que Rabossi incluye al comienzo de su artículo:

> La única prueba [*proof*] que puede darse de que un objeto es visible es que la gente lo ve. La única prueba de que un sonido es audible es que la gente lo oye; y lo mismo puede afirmarse de las otras fuentes de nuestra experiencia. De manera similar, creo yo, el único elemento de prueba [*evidence*] que es posible presentar en favor de la afirmación de que algo es deseable es que la gente lo desea. Si el fin que la doctrina utilitarista se propone a sí misma no fuera reconocido como un fin, en la teoría y en la práctica, nada podría convencer a nadie de que lo es. No puede darse razón alguna de por qué la felicidad general es deseable, excepto que cada persona, en la medida en que la crea alcanzable, desea su propia felicidad. Pero entonces, siendo esto un hecho, no sólo tenemos toda la prueba [*proof*] que el caso admite, sino toda la que es posible exigir, de que la felicidad es un bien: de que la felicidad de cada persona es un bien para esa persona [*that each person's happiness is a good for that person*] y de que, en consecuencia, la felicidad general es un bien para el conjunto [*the aggregate*] de todas las personas.[1]

---

* Me complace agradecer las provechosas conversaciones con Manuel Comesaña, siempre clarificadoras, durante la redacción de este comentario.

[1] Sólo he modificado ligeramente la traducción de Rabossi, distorsionada, sin duda, por reiterados *lapsus* tipográficos. La oración final se halla truncada doblemente, lo cual contribuye a oscurecer el ya conflictivo párrafo IV (3), que no necesita en castellano más tinieblas de las que posee en inglés. Se puede comprobar con desaliento que esta misma oración reaparece en la página 4 con errores distintos.

II

Con buena técnica expositiva, Rabossi divide los textos de los críticos de Mill en dos clases: los que se refieren al par *deseado/deseable* y los que se ocupan del par *felicidad individual/felicidad general*.[2] Es en relación con el segundo par que cita una oración final del fragmento de IV(3) y afirma que, si se atribuyera a la expresión "en consecuencia" un sentido lógico –lo cual, según él, no debe hacerse–, entonces Mill sería realmente culpable de "una burda falacia de composición", como K. Britton nos asegura con menos elocuencia desde la página 5.

Sin duda alguna, la oración de Mill que Rabossi *quiso* citar es la siguiente:

> (A) ...la felicidad de cada persona es un bien para esa persona, y la felicidad general, en consecuencia, es un bien para el conjunto de las personas.[3]

Pero ¿por qué afirma Rabossi que *si (A) pretendiera ser una inferencia deductiva entonces Mill incurriría en la falacia de composición?* El autor parece considerarlo evidente, pues no ofrece argumento alguno. Sin embargo, su imputación contrafáctica es digna por sí misma de ser testeada.

Si incluimos (A) en un fragmento más amplio de IV(3) (omitiendo las frases superfluas para nuestro propósito) obtenemos:

> (A') No puede darse razón alguna de por qué la felicidad general es deseable, excepto que *cada persona* [...] desea su propia felicidad. [...] Tenemos toda la prueba [...] de que la felicidad de *cada persona* es un bien para *esa persona* y de que, en consecuencia, la felicidad general es un bien para el conjunto de las personas (la bastardilla me pertenece; T.M.S.).

¿Cuál es el antecedente referencial de "esa persona" en (A')? Parece razonable pensar –a menos que se elimine esta interpretación mediante algún argumento– que "esa persona" remite a la segunda figuración de "cada persona". De acuerdo con esto, la oración subordinada: "la felicidad de cada per-

---

[2] Cometeré el pecado de señalar "una pequeña brizna en el ojo ajeno": la exposición de Rabossi nos deja sin saber si lo que Moore critica en Mill es la afirmación de que "deseable" *significa* lo mismo que "deseado", o la afirmación distinta de que "deseable" significa *susceptible de ser deseado*, que es lo que leemos en el texto de *Principia Ethica* citado por él. Podría ser que tampoco el obsesivo Moore hiciera la necesaria distinción; pero Rabossi nada dice al respecto.

[3] No es ésta, claro, la oración que figura en la página 4 de *Análisis Filosófico*. Cf. nota 1.

sona es un bien para esa persona", afirma que *cada persona es tal que su propia felicidad es un bien para ella (esa persona)*, lo que resultaría más evidente con una traducción menos literal y también menos equívoca: "la felicidad de *una* persona es un bien para esa persona", donde el artículo determinado funciona como un cuantificador universal. Esta paráfrasis elimina la tentación perversa de hacer decir a la oración que la felicidad de todos es un bien para *esa* persona, lectura que deja la expresión "esa persona" fuera de todo contexto referencial plausible, semánticamente a la intemperie.

Hay en (A'), y por lo tanto en IV(3), una secuencia inferencial de tres pasos, que simbolizaremos parcialmente como sigue, suponiendo que el dominio de "x" se halla constituido por personas:

(I) 1. (x) (x desea su propia felicidad);
$\therefore$ 2. (x) (su propia felicidad es un bien para x);
$\therefore$ 3. La felicidad general es un bien para el conjunto de las personas.

El paso de 1 a 2 se apoya simplemente en la tesis, defendida por Rabossi en la segunda parte de su artículo, según la cual *ser deseado* es un elemento de prueba válido (aunque no deductivo) para *ser deseable*, y no lo discutiremos por el momento.[4]

En cuanto al tránsito de 2 a 3, se abren dos posibilidades interpretativas, según la función referencial que se le asigne a la expresión "el conjunto de las personas": o bien designa una entidad colectiva distinguible de sus componentes, o bien es usada como sinónima de "todas y cada una de las personas", de acuerdo con la interpretación razonable que Rabossi adopta (sin ningún análisis) en la página 11. Pero en tal caso la conclusión (3) sería reformulable, más claramente, en la forma de

3'. (x) (la felicidad general es un bien para x).[5]

Falta aquí, sin embargo, un elemento típico de la falacia de composición:

---

[4] Téngase en cuenta, además, que según Mill si x *es deseable* entonces x *es un bien*.

[5] Si 'Ø es un bien para x' expresara una relación psicológica, la interpretación distributiva de "el conjunto de las personas" sería algo más compleja de lo que nuestro análisis exhibe. Suponiendo que 'felicidad general' = 'felicidad de todas y cada una de las personas', 3 podría interpretarse como

3'. (x) (x considera un bien que (y) (y sea feliz)), pero no, naturalmente, como

3". (x) (y) (x considera un bien que y sea feliz).

La primera interpretación es *opaca*; la segunda, *transparente*.

un sujeto constituido por un término singular referente a una "totalidad", cuya presencia hace posible el salto falaz ejemplificado en el cuasi esquema siguiente: "Las partes de la máquina *M* son livianas; por lo tanto, *M* es liviana". *Es difícil comprender cómo Rabossi adopta una interpretación distributiva de "el conjunto de las personas" y sostiene al mismo tiempo que si la relación inferencial fuera deductiva el paso de 2 a 3 sería una falacia de composición.* Ambas pretensiones son lógicamente incompatibles. Volvamos, pues, al razonamiento (I).

Es obvio, además, con prescindencia de cuál sea la naturaleza del paso inferencial, que 2 y 3 contienen predicados distintos, que expresan propiedades *prima facie* diferentes,[6] y que por lo tanto es *prima facie* increíble que haya una falacia de composición, a menos que se demuestre lo contrario.[7]

III

En la sección II de su artículo, Rabossi imagina a dos refinados contrincantes *A* y *B*; el primero sostiene que el fin último es *la felicidad* y el segundo *la santidad*. Con el objeto de hacer posible la discusión ambos convienen en el criterio de que, para merecer credibilidad, una doctrina de los fines últimos debe "estar de acuerdo con la naturaleza humana" (pág. 6). El supuesto pertenece al propio Mill (*U* IV(9)). Pero ¿qué implica la afirmación de que la deseabilidad de cierto estado Ø (la felicidad o la santidad) "está de acuerdo con la naturaleza humana"? Sorprendentemente, Rabossi responde con notable prontitud:

> En tal caso podríamos decir que *A* y *B* están de acuerdo en que lo que es deseable como fin (la felicidad, la santidad) tiene que ser efectivamente deseado como tal (p. 9).

Pero esta rápida conclusión es un simple *non sequitur*. Para verlo con claridad introduzcamos tentativamente algunas precisiones terminológicas mínimas.

(A) "La naturaleza humana" designa cierto conjunto *H* relativamente fijo de

---

[6] El predicado que figura en 2 es: "Su propia felicidad es un bien para...", mientras que en 3 tenemos: "la felicidad general es un bien para...".

[7] Existe una posibilidad extrema, aunque no atribuible a Mill, de crear una falacia de composición; pero Rabossi no puede apelar a ella, debido a su interpretación distributiva de "el conjunto de las personas".

propiedades peculiares de nuestra especie, algunas de las cuales son disposicionales.

(B) La deseabilidad de cierto estado Ø (la felicidad o la santidad) 'está de acuerdo' (o es 'compatible') con la naturaleza humana si y sólo si *desear* Ø no es nomológicamente incompatible con la posesión de las propiedades que la integran.[8]

De aquí surge sin esfuerzo que la doctrina utilitaria podría ser 'compatible' con la naturaleza humana sin que nadie deseara *de hecho* la felicidad: bien podría suceder que *desear la felicidad* fuese el consecuente de una propiedad disposicional más amplia, perteneciente a H y de la forma $C \rightarrow F$, donde "C" representa una condición no realizada y "F" el deseo de la felicidad. La compatibilidad de la doctrina utilitaria con el conjunto H no garantiza, pues, que lo que en ella es deseable como fin sea "efectivamente deseado".

Al menos si prestamos atención a la biología, es fácil advertir que las prescripciones genéticas no poseen en general una forma categórica sino condicional: dicen que si se da la circunstancia C se produce la respuesta R (o aparece la característica F). Los 'condicionales genéticos' pueden formar cadenas causales complejas:

$$C_1 \rightarrow R_1; \ldots C_n \rightarrow R_n,$$

lo cual explica que la variación de i en C_i produzca culturas muy diversas, todas ellas 'compatibles', no obstante, con la naturaleza humana. No es de extrañar que las ideologías más discrepantes (el igualitarismo, el anti-igualitarismo) hayan coincidido en apelar al "acuerdo con la naturaleza humana", pues las características manifiestas de los hombres en las distintas culturas son también incompatibles entre sí. La naturaleza humana determina a lo sumo el conjunto de culturas genéticamente posibles, pero ninguna en particular; y a esto hay que añadir las variaciones individuales, que dependen de la historia peculiar de cada *homo sapiens*.

Nadie negará que la construcción de nidos 'concuerda' con la naturaleza del hornero; pero también es compatible con ella que el hornero no construya su nido, si se dan condiciones algo extravagantes en las que no se producen, por ejemplo, algunas de las señales sucesivas que determinan el éxito de su empresa.

---

[8] Las convenciones semánticas (A) y (B) son perfectibles o reemplazables por otras; pero su sitio no puede quedar vacante, pues se necesitan para evitar un vacío inferencial. Su función aquí es fundamentalmente ilustrativa; si se las reemplaza por otras, las consecuencias serán también distintas. No parece factible, sin embargo, omitir el componente disposicional en cualquier definición alternativa de 'naturaleza humana'.

De modo análogo, bien podría ocurrir que aunque la deseabilidad de Ø fuera compatible, en el sentido indicado, con la naturaleza humana, de hecho el hombre no desee Ø, debido a circunstancias poco propicias para la ética de Mill. Suponiendo que Ø es el deseo de felicidad, las prescripciones genéticas podrían incluir al mismo tiempo los condicionales $C_1 \rightarrow Ø$ y $C_2 \rightarrow —(Ø)$; si para desgracia de John Stuart Mill se cumple C_2, entonces C_1 es falso y el hombre no desea la felicidad, todo en perfecto 'acuerdo' con la naturaleza humana.

IV

Al final de la página 9, Rabossi resume la argumentación de Mill de manera compacta y provocativa:

> El argumento de Mill es sencillo. Si alguien se propusiera hablar de entidades visibles o audibles que no son vistas ni oídas por nadie, concluiríamos, en base a esa sola circunstancia, que su propuesta carece de sentido. *El único test* que se puede proponer con seriedad de que una cosa es visible (audible) es que tal cosa es efectivamente vista (oída). Una teoría que chocara con esta exigencia pasaría el límite de lo razonable. Lo mismo ocurre con ciertas propuestas de los filósofos de la moral. Si el fin último que se propone no es de hecho deseado, se incurrirá en un sinsentido similar. La doctrina carecerá entonces de una base mínima de aceptabilidad; no podrá ser creída por nadie.

¿Es ésta una meditada defensa o una mera interpretación del texto de Stuart Mill? Como interpretación es dudosa; como defensa es problemática. ¿Cree realmente Rabossi que no es posible testear de manera indirecta la visibilidad de objetos nunca vistos y que hablar de tal cosa "carece de sentido"? Por fortuna para nuestra incertidumbre, el autor ha escrito un trabajo de temática similar que contiene el mismo párrafo que acabamos de reproducir y al término del cual dice textualmente: "Pienso que el requisito que sugiere Mill es correcto".[9] Parecería, pues, que, si Rabossi mantiene su palabra, la respuesta a nuestro interrogante es afirmativa.

Supongamos ahora que deseamos saber si un objeto *X* nunca visto (por ejemplo, el lado oculto de la Luna) es visible y que 'visible' = 'susceptible de ser visto en condiciones N por una personal normal'. ¿No sería pertinente para ello determinar si *X* refleja las ondas electromagnéticas de cierto rango de frecuencias?

---

[9] "Acerca de una prueba posible de los primeros principios éticos", incluido en *Estudios éticos* (Universidad de Carabobo, Valencia, Venezuela); la cita se halla en la página 115.

Es difícil creer que Rabossi sostenga una teoría tan extrema acerca del testeo de los enunciados que afirman la existencia de disposiciones y los criterios de significatividad aplicables a ellos; la extensión de esta teoría al lenguaje moral se presenta así como problemática *ab initio*. Pero sería pretencioso formular conjeturas sobre las motivaciones filosóficas del autor.

CONICET
INSTITUTO DI TELLA