

# Sobre el tercer paralogismo: “Yo pienso” y el observador externo

## *On the Third Paralogism: “I think” and the Outside Observer*

Javier Vidal

### Resumen

Para J. Bennett el núcleo del tercer paralogismo está en decir algo sobre los juicios, basados en un aparente recuerdo, de la forma “Era yo quien era F en t”. La idea es que el juicio de una persona de que era ella quien era F en t podría ser erróneo en lo relativo a la *identidad* de la persona que realmente era F en t. Según A362-4, el rol del *observador externo* sería corregir un juicio de esa clase. Pero P. F. Strawson sostiene que lo que el tercer paralogismo trata de decir es algo derivado del hecho de que el “Yo pienso” expresa el pensamiento de la *identidad de la conciencia*. El error del psicólogo racional está en creer que el “Yo pienso” realmente expresa un *conocimiento de objetos* de la identidad del sujeto pensante como la identidad de un Ego cartesiano. Entonces, el rol del observador externo consiste en no estar sujeto a una tal ilusión de conocimiento del sujeto pensante. La tesis defendida en este artículo es que el error del psicólogo racional está en creer que el “Yo pienso” realmente expresa un *conocimiento a priori* de la identidad del sujeto de experiencias como la identidad de una sustancia pensante. El rol del observador externo consiste en no estar sujeto a una tal ilusión de conocimiento.

*Palabras clave:* Identidad de conciencia, sujeto de experiencias, “yo pienso”, cuasi-recuerdo, observador externo, sustancia pensante, conocimiento de objetos, conocimiento *a priori*.

## Abstract

For J. Bennett the kernel of the third paralogism is to say something about judgements, based on an apparent memory, of the form "It was I who was F at t". The idea is that a person's judgement that it was he who was F at t could be mistaken in respect of the *identity* of the person who actually was F at t. According to A362-4, the role of the *outside observer* would be to correct such a judgement. But P. F. Strawson supports that what the third paralogism tries to say is something from the fact that the "I think" expresses the thought of the *identity of the consciousness*. The mistake by the rational psychologist is to believe that the "I think" actually expresses a *knowledge of objects* of the identity of the thinking subject as the identity of a Cartesian Ego. Then the role of the outside observer is not to be liable to any such illusion of knowledge of the thinking subject. The thesis defended in this paper is that the mistake by the rational psychologist is to believe that the "I think" actually expresses an *a priori knowledge* of the identity of the subject of experiences as the identity of a thinking substance. The role of the outside observer is not to be liable to any such illusion of knowledge.

*Keywords:* Identity of Consciousness, Subject of Experiences, "I Think", Quasi-Memory, Outside Observer, Thinking Substance, Knowledge of Objects, *A Priori* Knowledge.

## 1. La interpretación de J. Bennett y la crítica de P. F. Strawson

Según la interpretación del tercer paralogismo por parte de J. Bennett<sup>1</sup>, el propósito de Kant es corregir cierto error filosófico sobre una clase de juicios de memoria, aquellos basados en los recuerdos aparentes de un sujeto respecto a un estado, presuntamente suyo, en el pasado. El error filosófico sería pensar que el juicio del sujeto así basado, aunque podría ser falso puesto que la memoria es falible, en cambio, no podría ser falso en lo relativo a la *identidad* del sujeto de tal estado<sup>2</sup>. De esta forma, el argumento de Kant está dirigido contra el lugar privilegiado concedido, supuestamente por la psicología racional, a la memoria entre los criterios empíricamente aplicables de la identidad personal. Como P. F. Strawson señala<sup>3</sup>, es un principio básico que una doctrina, sea o no errónea, sobre los crite-

<sup>1</sup> Bennett, J. : *Kant's Dialectic*, Cambridge, Cambridge University Press, 1974. En lo que sigue, haré uso de la traducción castellana; Bennett, J. : *La 'Crítica de la razón pura' de Kant. 2. La Dialéctica*, traducción de Julio César Armero, Madrid, Alianza Editorial, 1981, pp. 110-31.

<sup>2</sup> Cf. *Ibid.*, pp. 114-5.

<sup>3</sup> Strawson, P. F. : "Kant's Paralogisms: Self-Consciousness and the 'Outside Observer'", en

rios empíricamente aplicables de identidad tiene que ocuparse de un objeto de experiencia, un objeto de la intuición sensible, un tema de estudio para una psicología empírica del sentido externo o del sentido interno. En su caso, sería una doctrina sobre una persona individual o un sujeto de experiencias como fenómeno, es decir, tal y como se manifiesta para los sentidos externos y en el flujo de las experiencias internas, en la conciencia empírica, pero no tal y como es en sí mismo. Pero resulta que para Kant la psicología racional es una doctrina cuyo 'solo texto', el tema de sus argumentos y la base de sus pseudo-conclusiones, es un determinado 'pensamiento' o 'representación' que él introduce como el "Yo pienso", donde el "yo" de la representación no se refiere a un objeto de experiencia en absoluto: esto sería hacer de la psicología racional una psicología empírica<sup>4</sup>. Dejemos por el momento la cuestión de cuál es, además de esta tesis negativa, la tesis positiva de la psicología racional sobre el "Yo pienso".

Expongamos más detenidamente la interpretación de Bennett. Él dice que el problema del tercer paralogismo está en un intento de afirmar algo acerca de juicios de la forma "Era yo quien era F en t", juicios que expresan el recuerdo o el aparente recuerdo de un sujeto de ser F en t. De entrada, la raíz nutricia del tercer paralogismo sería: "Consiguientemente, refiero todas y cada una de mis determinaciones al yo numéricamente idéntico, es decir, en la forma de la intuición interna de mí mismo. Desde esta base, no deberíamos considerar la personalidad del alma<sup>5</sup> ni siquiera como inferida, sino como una proposición enteramente idéntica de la autoconciencia en el tiempo, lo que explica también el que posea validez *a priori*"<sup>6</sup>. Según una primera lectura, Kant quiere decir simplemente que la identidad de la conciencia en el tiempo no es algo inferido sino logrado de un modo más inmediato. De hecho, al principio de la *Crítica* ya afirma: "Todos los fenómenos posibles pertenecen, en cuanto representaciones, a toda la autoconciencia posible. Ahora bien, esta autoconciencia conlleva necesariamente, en cuanto representación trascendental, la identidad numérica, y es cierta *a priori*"<sup>7</sup>. Según la interpretación

---

Strawson, P. F. : *Entity and Identity and Other Essays*, Oxford, Clarendon Press, 1997, pp. 252-67. Cf., p. 255.

<sup>4</sup> Kant, I. : *Crítica de la razón pura*, traducción de Pedro Ribas, Madrid, Alfaguara, 1978/1993. Cf., A342-3/B400-1. Por ejemplo, en A342-3 Kant dice: "La doctrina *racional* del alma es, efectivamente, una tarea de esta índole, ya que, de mezclarse entre los fundamentos de esta ciencia el menor elemento empírico de mi pensar o alguna especial percepción de mi estado interno, no sería ya doctrina racional, sino doctrina *empírica*".

<sup>5</sup> Respecto a 'personalidad del alma' Kant parece referirse aquí trivialmente a la identidad de la conciencia empírica en el tiempo. Respecto a 'persona' Kant hace un uso ambiguo: como expresiones no cargadas parecen referirse trivialmente a la condición de ser un sujeto de experiencias, mientras que la lectura cargada del psicólogo racional, especialmente en la conclusión del paralogismo, apela a la condición cartesiana de ser una sustancia pensante.

<sup>6</sup> *Ibid.*, A362.

<sup>7</sup> *Ibid.*, A113.

de Bennett, lo que sería logrado de un modo más inmediato son juicios empíricos de autoidentidad de la forma "Era yo quien era F en t". Tal inmediatez es debida al principio cartesiano de que, al formar tales juicios, no puedo conocer que alguien era F y dudar de si era yo: recuerdo mis estados pasados como míos, y en consecuencia un juicio "Era yo quien era F en t" no se puede inferir, pues no puedo poseer el dato mnémico "Alguien era F" sin poseerlo en la forma "Yo era F" o en la forma "Alguien distinto era F"<sup>8</sup>.

Pero el psicólogo racional, además, trata de afirmar con este principio que no es posible el caso de que mi recuerdo indubitable de ser F en t sea erróneo en lo relativo a la identidad de la persona, supuestamente yo mismo, cuyo ser F en t se recuerda. El argumento de Kant a favor de tal posibilidad puede explicarse con una mayor precisión con la ayuda del concepto de Shoemaker<sup>9</sup> de *cuasi-recuerdo*. Esto significa que el aparente recuerdo de un sujeto de ser F en t podría ser realmente causado, y de una forma estándar, por ser alguien F en t pero dándose el hecho de que algún suceso, por ejemplo, de fisión personal, haya intervenido entre el estado de ser F en t y el aparente recuerdo del estado: la consecuencia es que esa intervención puede impedirnos afirmar que el sujeto del estado original es *idéntico* al sujeto que tiene el aparente recuerdo, es decir, el cuasi-recuerdo. En un caso de fisión, de ninguno de los productos (personas, sujetos) de la fisión puede decirse que es idéntico a la persona o sujeto de la fase pre-fisión así que cualquier juicio de identidad de la fase pos-fisión que esté basado en un cuasi-recuerdo de un estado real de la persona de la fase pre-fisión sería un juicio de identidad falso. Esto ayuda a explicar la posibilidad concebida por Kant: el juicio de un sujeto de que era él quien se encontraba en un estado, basado en un aparente recuerdo de tal estado, podría ser falso en lo relativo a la identidad de la persona, supuestamente él mismo, que realmente se encontraba en ese estado<sup>10</sup>.

Desde luego, no deja de ser sorprendente esta versión de lo que Kant ve en el tercer paralogismo. Bennett parece pensar que esta interpretación es obligada para poder explicar la referencia de Kant en A362-4 a un observador externo. Kant dice que todos nosotros debemos juzgar necesariamente determinada proposición pero que "no podemos dar todavía por válido este juicio desde el punto de vista de otro"<sup>11</sup>. De entrada, señalemos que la proposición en cuestión es realmente "somos los mismos en todo el tiempo del que poseemos conciencia"<sup>12</sup>. Bennett deja fuera de consideración esta proposición desde el principio porque entiende que, a lo

<sup>8</sup> Cf. Bennett, *op. cit.*, pp. 112-3.

<sup>9</sup> Cf. Shoemaker, S. : "Persons and their Pasts", *American Philosophical Quarterly*, 74, 1970, pp. 269-85.

<sup>10</sup> Cf. Bennett, *op. cit.*, pp. 114-7; Strawson, *art. cit.*, pp. 262-3.

<sup>11</sup> Kant, *op. cit.*, A364.

<sup>12</sup> *Ibid.*

sumo, expresa una trivialidad: necesariamente juzgamos que cualquier cosa es 'la misma en todo el tiempo' de su existencia, sea o no consciente. La evaluación del contenido de esta proposición será decisiva más tarde en mi discusión. Añadido a esto, Bennett se desentiende del hecho de que Kant utiliza gramaticalmente el singular para referirse al juicio, porque la proposición es una, y supone que Kant está refiriéndose a toda la clase de juicios basados en el recuerdo de la forma "Era yo quien era F en t". Pues bien, la conclusión de Bennett es que un observador externo es precisamente alguien que podría estar en una posición privilegiada para corregir un juicio de esa clase en lo relativo a la cuestión de la identidad, y tal sería el *quid* de Kant. La idea es que yo podría pensar que estoy recordando haber sido F pero un observador bien informado me podría decir: "Usted sólo está cuasi-recordando haber sido F. Ha habido una división y, aunque usted es continuo con la persona que fue F, no es idéntico a ella". La relevancia de una apelación al observador externo puede entenderse mejor desde el *dictum* analítico, explotado por el propio Strawson, de que mientras mis juicios de autoidentidad no hacen uso de criterios de identidad personal y, por tanto, yo mismo no estoy en una posición para corregirlos, en cambio, los juicios del observador sobre mi identidad están realmente basados en tales criterios, lo que avala su posición. Lo que está por ver es qué significa que yo *necesariamente* tenga que juzgar, si es un juicio basado en un aparente recuerdo, "Era yo quien era F en t". Parece que la explicación sería: yo tengo necesariamente que realizar ese juicio porque, no usando criterios de identidad personal al realizar el juicio, no estoy en una posición para corregirlo. Bennett, al introducir ese orden de explicación, rechaza la tesis kantiana sobre la necesidad del juicio: la razón es que cualquiera de los criterios que un observador externo podría usar para corregir un tal juicio de autoidentidad de una persona podrían estar igualmente a disposición de la persona misma. Por ejemplo, mi cuasi-recuerdo de ser F apoya, ciertamente no como criterio, el juicio de que yo era F, pero luego necesito criterios de identidad personal para determinar si realmente yo era F o alguien relacionado conmigo de determinada manera era F, dando entrada a elementos procedentes del dominio del observador. El resultado sería que el contraste entre la posición del observador externo y la posición del sujeto mismo en lo relativo a la identidad del sujeto es, en último caso, un contraste espurio<sup>13</sup>.

Pasemos ya a examinar la crítica de Strawson, aparte de la consideración general sobre la psicología racional hecha al principio. El objetivo es, claro está, refutar que Kant estuviera interesado en la posibilidad de que un sujeto cometa ciertos errores de identidad acerca de su propia historia pasada. Ténganse en cuenta dos textos muy relevantes. Primero, lo que sigue a la afirmación de que no podemos dar todavía por válido cierto juicio desde el punto de vista de otro : "...ya que, al no encontrar en el alma otro fenómeno permanente que la representación del yo, la cual

<sup>13</sup> Cf. Bennett, *op. cit.*, pp. 117-20.

las acompaña y enlaza todas, nunca podemos establecer si este yo (mero pensamiento) no fluye igual que los otros pensamientos que él liga entre sí"<sup>14</sup>. Y antes Kant ha dicho del observador externo lo siguiente: "En consecuencia, del yo que, con plena identidad, acompaña en todo tiempo las representaciones de *mi* conciencia el observador no inferirá todavía, aunque la admita, la permanencia objetiva del propio yo (de mí mismo)"<sup>15</sup>. La idea de Strawson es que los textos apuntan claramente en la dirección de que no está en juego cualquier error *particular* de identidad que una persona pudiera cometer acerca de su propia historia, sino que, de haber algún error, se trataría de un error *general* que cualquier hombre o persona, al ser consciente de la sucesión de sus propios estados de conciencia, está en condiciones de cometer sobre sí mismo como *ser pensante*, un error, sin embargo, que nadie, *observándole a él desde fuera*, está en condiciones de cometer sobre él<sup>16</sup>.

Strawson plantea entonces la cuestión de cuál sería ese error y la responde en términos 'strawsonianos'. Hay una práctica lingüística mediante la que uno ha aprendido a usar el pronombre personal "yo" en contextos donde uno es capaz de identificarse a sí mismo como un objeto de la intuición sensible, por ejemplo, siendo capaz de señalar su propio cuerpo. A partir de ahí, uno aprende a usar referencialmente el pronombre personal en contextos donde uno no es capaz de identificar el objeto de referencia, es decir, él mismo. En estos casos puede decirse que, aunque el uso de "yo" es referencial y mantiene la misma referencia, uno mismo, sin embargo, no es una referencia *personal* en el sentido de que no es una referencia a uno mismo en contextos donde uno puede identificarse como una persona. Tales contextos son aquellos en los que "yo" sólo se usa en conexión con estados de conciencia, como se usa, por ejemplo, en expresiones de la forma "Yo percibo...", donde los estados de conciencia en cuestión, las percepciones, tienen un contenido empírico particular. Respecto al dominio de estos usos, la idea de Kant es que, como él lo dice, aprendemos a unir el "yo" a una percepción *solamente* a causa de la *conciencia* de la percepción misma. El resultado es que nada cuenta como conciencia o experiencia de un estado presente de conciencia, una percepción en curso, excepto como una experiencia de encontrarse uno mismo en ese estado de conciencia, de tener esa percepción. Y del mismo modo, nada cuenta como un aparente recuerdo de un estado pasado de conciencia excepto como un aparente recuerdo de haberse encontrado uno mismo en ese estado de conciencia. Ahora bien, todos estos son usos empíricos de "yo". Pero, sabido es, sucede que hay un uso de "yo" en la representación "Yo pienso" cuya función es expresar la necesaria unidad de la conciencia o la pura forma de la conciencia en general, la razón de ser de que un estado de

---

<sup>14</sup> Kant, *op. cit.*, A364.

<sup>15</sup> *Ibid.*, A362-3. La expresión entre paréntesis es un añadido con fines aclarativos.

<sup>16</sup> Cf. Strawson, *art. cit.*, pp. 264-5.

conciencia sea un estado de *conciencia* por el solo hecho de estar enlazado con los otros estados de conciencia<sup>17</sup>.

Para Strawson que el uso de "yo" en la representación "Yo pienso" tenga esa función, obviamente no referencial en la medida en que la forma de la conciencia no es un objeto, no excluye que ese mismo uso de "yo" sea referencial: en concreto, el posible objeto de referencia sería el yo trascendental, el sujeto de experiencias tal y como es en sí mismo, es decir, como noúmeno<sup>18</sup>. Este doble aspecto de tal uso será discutido después. Lo relevante ahora es señalar que el uso de "yo" en la representación "Yo pienso" es indiscutiblemente introducido en la Deducción Trascendental sólo para expresar la necesaria unidad de la conciencia. Entonces surge la cuestión: ¿por qué tiene que haber un uso de "yo" para expresar la razón de ser de que un estado de conciencia sea un estado de conciencia por el *solo* hecho de estar enlazado con los demás? De entrada, la idea sería que tal uso se aprende mediante un proceso de *abstracción* por el que se dejan fuera los distintos contenidos empíricos particulares expresables como "Yo percibo...", expresables en un momento como "Yo percibo P", en otro momento como "Yo percibo Q" y así sucesivamente en el tiempo, para llegar a captar la forma de la conciencia en general, la forma de la conciencia común a los distintos estados de conciencia particulares en cuanto que están necesariamente unidos. Pues bien, parece que la unidad de la conciencia así abstraída, con independencia de las demandas de orden y conectividad que lleva aparejadas, demanda la *identidad de la conciencia* en el tiempo (lo que significa, de la 'corriente' de la conciencia empírica) en lo relativo a todos los estados de conciencia en que actualmente nos encontramos o aparentemente recordamos habernos encontrado. Y esta es la razón por la que "yo" es usado para expresar la necesaria unidad de la conciencia: porque hay una cuestión, aunque exclusivamente formal, sobre la identidad. Por tanto, lo que Kant consideraría trivial del tercer paralogismo es la afirmación de la identidad de la conciencia en el tiempo, la afirmación de que la corriente de conciencia es una y la misma y, por ello, hay una forma de la conciencia en general<sup>19</sup>.

¿Cuál sería, pues, para Strawson el error general que cualquier hombre está en condiciones de cometer y que el psicólogo racional realmente comete? Es evidente que un juicio de autoidentidad en el tiempo que uno haga que, al abstraer de todo contenido empírico, meramente exprese la identidad de la conciencia en el tiempo, no puede poner en juego la cuestión de la identidad personal. Abstraer del contenido empírico de un aparente recuerdo expresable como "Yo recuerdo..." es evidente-

<sup>17</sup> Cf. *Ibid.*, pp. 259-60.

<sup>18</sup> Cf. *Ibid.*, p. 257.

<sup>19</sup> Cf. *Ibid.*, pp. 255-6 y p. 265. Cf. McDowell, J. : *Mind and World*, Cambridge Mass., Harvard University Press, 1994, pp. 99-104, y, para el concepto de *abstracción*, especialmente Cassam, Q. : *Self and World*, Oxford, Oxford University Press, 1997, pp. 153-70.

mente abstraer del componente empírico de ese contenido cuyo referente podría ser una determinada persona. Pero dada la permanencia de la representación "Yo pienso" como acompañante de todas las representaciones, donde tal representación está desprovista de todo contenido empírico, el psicólogo racional puede interpretar la verdad de un juicio de autoidentidad en el tiempo que sólo es un juicio formal sobre la identidad de la conciencia ("somos los mismos en todo el tiempo del que poseemos conciencia" (A364)), que sólo expresa la representación "Yo pienso", como la *verdad* de un juicio sobre el sujeto de experiencias como *ser pensante*: en tanto que es una sustancia simple, inmaterial (no identificable empíricamente) e idéntica que permanece invariable a través de todos los cambios de estado, es decir, en tanto que es un *Ego cartesiano*. Ya hemos dicho que según Kant la proposición de A364 la debemos juzgar necesariamente, lo que significa que es algo de lo que estamos ciertos *a priori* y, por tanto, algo de lo que tenemos conocimiento. Luego, la proposición de A364, que corresponde al juicio de autoidentidad en el tiempo que Kant está considerando y que para él sólo expresa el trivial contenido de la identidad de la conciencia, es algo de lo que tenemos *conocimiento*. Este sería el *quid* del error del psicólogo racional. Es por ello que el *quid* no está en que el psicólogo racional tome un juicio de autoidentidad que sólo es un juicio de identidad de la conciencia como un juicio de identidad del sujeto de experiencias como ser pensante, sino que tome el conocimiento expresado por tal juicio de autoidentidad, la mera identidad de la conciencia, como un conocimiento de la identidad del sujeto. O, en los términos más propios de Strawson, el *quid* del error está en que el mero *pensamiento* expresado en tal juicio como un pensamiento de la identidad del sujeto de experiencias como un ser pensante, pensamiento cuya posibilidad no es cuestionada por Strawson, sea tomado como un *conocimiento* de la identidad del sujeto. Y por supuesto, dada la abstracción de todo componente empírico cuyo referente pudiera ser el sujeto de experiencias como ser pensante, resulta que el conocimiento de la identidad del sujeto no puede ser más que el conocimiento de la identidad de algo que no es identificable empíricamente, un Ego cartesiano (la simplicidad es una consecuencia del resto del vaciamiento empírico). El argumento de Kant es precisamente que, como ha probado en la *Crítica*, no puede haber *conocimiento de objetos* que no sean objetos empíricamente identificables, objetos de una experiencia posible, y así el psicólogo racional ha caído en la ilusión de tomar el mero conocimiento *a priori* de la identidad de la conciencia, que no es un conocimiento de objetos, como un auténtico conocimiento de objetos de la identidad del sujeto de experiencias como un ser pensante. En los términos de Strawson, el argumento es que, como no puede haber conocimiento de objetos que no sean empíricamente identificables, el psicólogo racional ha caído en la ilusión de tomar el mero pensamiento de la identidad del sujeto de experiencias como un ser pensante como un conocimiento de objetos de la identidad del sujeto así concebido. La ilusión es man-



ifiesta: al abstraer del componente empírico relativo al sujeto en la representación "Yo pienso" se deja fuera una de las condiciones de posibilidad de que haya conocimiento de objetos del sujeto en absoluto, a saber: la *intuición sensible* del mismo. No es que Kant niegue que haya intuición sensible de la persona o sujeto de experiencias tal y como aparece para los sentidos externos o en el flujo de las experiencias en el sentido interno: es que no cabe una intuición sensible del sujeto de experiencias como ser pensante, como es referido por abstracción en la representación "Yo pienso"<sup>20</sup>.

He aquí, en términos de unidad más que de identidad de la conciencia, cómo lo dice Kant: "Sin embargo, nada hay más natural y tentador que la ilusión consistente en tomar la unidad de la síntesis de los pensamientos por una unidad *percibida* en el sujeto de esos mismos pensamientos. Podríamos llamarla la subrepción de la conciencia hipostasiada"<sup>21</sup>; "De todo ello se desprende que la psicología racional debe su origen a un simple malentendido. Se toma la unidad de la conciencia, que sirve de base a las categorías, por intuición del sujeto como *objeto* y se le aplica la categoría de sustancia. Ahora bien, tal unidad es sólo la unidad del pensamiento. Ningún objeto se da en virtud de ella sola y, consiguientemente, no podemos aplicarle la categoría de sustancia, que presupone siempre una *intuición* dada; en consecuencia, no podemos conocer ese sujeto"<sup>22</sup>.

¿Cuál es, entonces, para Strawson la función asignada al observador externo en el desvanecimiento de la ilusión? La idea kantiana sería que el observador externo al que *yo*, el hombre o la persona, me presento como un objeto de la intuición externa, no está en condiciones como yo mismo sí lo estoy de caer en la ilusión de tener un conocimiento de *mí*, de mi identidad, como ser pensante. Como un objeto de la intuición externa yo soy ciertamente para él un objeto empíricamente identificable y reidentificable. Por otro lado, es claro que, aunque sea un mero observador que contempla un objeto desde fuera, el observador externo puede atribuirme a mí, precisamente como persona, percepciones, pensamientos y en general conciencia así como la identidad (y la unidad) de la conciencia: más aún, en tanto que toda persona tiene conocimiento *a priori* de la identidad de la conciencia, el observador, al tener conocimiento *a priori* de la identidad de la conciencia en su propio caso, puede atribuirme también a mí, como persona, tal conocimiento en mi caso. Del mismo modo, en los parámetros de Strawson, el observador externo, al tener en el propio caso el mero pensamiento de sí mismo, sujeto de experiencias, como ser pensante, puede atribuirme tal pensamiento en mi caso. Pues bien, el *quid* de la función del observador es que, a pesar de la riqueza de tales atribuciones, no podría haber un pensamiento suyo de *mí*, sujeto de experiencias, como ser pensante, ni hay por

<sup>20</sup> Cf. Strawson, *art. cit.*, pp. 265-6.

<sup>21</sup> Kant, *op. cit.*, A402.

<sup>22</sup> *Ibid.*, B421-2.

tanto el riesgo de tomar tal pensamiento como un conocimiento de *mi* como ser pensante. En efecto, la posición externa del observador hace que *él* no esté en condiciones de tener pensamiento alguno de *mi* identidad como ser pensante: la razón es que el pensamiento que yo tengo de mí mismo como ser pensante, en cuanto que soy consciente de mis estados de conciencia como *propios*, de la identidad de la conciencia como conciencia *propia*, hace que él no pueda tener un pensamiento de mí mismo como ser pensante, ya que él no es consciente de *mis* estados de conciencia como *propios*, ni de la identidad de *mi* conciencia como conciencia *propia*. Podríamos expresarlo diciendo que *su* representación "Yo pienso", que es aquella mediante la que podría pensarse en *algún* sujeto de experiencias como ser pensante al abstraer de todo contenido empírico representado en él, no acompaña las representaciones que el observador me adscribe a *mí* sino que acompaña *sus* representaciones: así, mediante *su* representación "Yo pienso" él sólo podría pensar en *su* identidad como ser pensante. En conclusión, el observador externo no está en condiciones de caer en la ilusión de tener *conocimiento* alguno de *mi* identidad como ser pensante, finalmente como una sustancia simple, inmaterial, permanente: un Ego cartesiano<sup>23</sup>.

## 2. Una nueva interpretación desde la crítica a P. F. Strawson

El diagnóstico del error del que el observador externo está exento me parece equivocado por parte de Strawson, y por la siguiente razón: por no cuestionarse la premisa del psicólogo racional de que es legítimo que uno mismo tenga un pensamiento (una autoconcepción) de la identidad del sujeto de experiencias como la identidad de una *sustancia pensante*. Más en concreto, la legitimidad de tal pensamiento es precisamente aquello en lo que consiste el error general. Se ha señalado que Strawson, junto a la función no referencial de expresar la identidad de la conciencia, añade al uso de "yo" en la representación "Yo pienso" la función de referirse al yo trascendental, es decir, al sujeto de experiencias tal y como es en sí mismo, y no tal y como se manifiesta en la intuición sensible. Podría preguntarse si necesariamente tal uso de "yo" se refiere al sujeto de experiencias como objeto y no, por el contrario, como sujeto. Podría responderse con consideraciones 'strawsonianas' sobre la distinción entre una referencia personal y una referencia no personal al sujeto, donde la distinción es entre los modos como se hace la referencia (referencia al sujeto como objeto y referencia al sujeto como sujeto) pero no entre referentes, ya que todo referente básico tiene que ser de suyo un objeto identificable y reidentificable. Pero esto significaría que la referencia al sujeto de experiencias

<sup>23</sup> Cf. Strawson, *art. cit.*, pp. 266-7. En los mismos términos se expresa Strawson en su versión de los paralogismos en Strawson, P. F. : *The Bounds of Sense*, London, Routledge, 1966. Cf., p. 166.

mediante el uso de "yo" en la representación "Yo pienso" es, aunque de un modo indirecto, la referencia a un objeto empíricamente identificable. Entonces, Kant ya no tendría una razón para decir que no puede haber conocimiento de la identidad del sujeto de experiencias.

Sin embargo, es evidente que para Kant la distinción entre una referencia al sujeto tal y como es en sí mismo y una referencia al sujeto tal y como se manifiesta en la intuición sensible, con una diferencia entre el sujeto como noúmeno y el sujeto como fenómeno, tiene un mayor alcance ontológico que una distinción entre modos de hacer referencia a una misma cosa (como sujeto y como objeto). Es que la noción de sujeto (como opuesta a objeto) es trascendental en el sentido de ser una condición *a priori* del conocimiento empírico mientras que la noción de noúmeno (como opuesta a fenómeno) es trascendental en el sentido de ser lo inaccesible (trascendente, diríamos) a la experiencia. De esta forma, hay una ambigüedad notoria al hablar de una referencia al 'yo trascendental'<sup>24</sup>. Pero en cuanto que referirse al sujeto de experiencias como yo trascendental es referirse al sujeto tal y como es en sí mismo, es decir, como noúmeno, es perfectamente legítimo referirse al sujeto de experiencias como un objeto, ya que no está en cuestión una referencia al sujeto como condición formal *a priori* de la experiencia sino como *algo* inaccesible a la experiencia, dado que no es algo que se manifieste en la intuición sensible. Esta referencia al sujeto de experiencias como objeto podría cuestionarse si sólo hubiera la posibilidad de referirse a algo como objeto *de experiencia*. La tesis kantiana es, claro está, que las condiciones en cuyos términos algo puede ser pensado como objeto se extienden más allá de la experiencia respecto al mero *pensamiento*<sup>25</sup>. Por tanto, referirse al sujeto como objeto es referirse a él como un objeto *en general*, sin pensarlo como algo dado en lo relativo a uno u otro modo de existencia, a una u otra determinación, en alguna intuición: "A través del pensamiento no me represento, pues, ni como soy ni como me manifiesto a mí mismo, sino que sólo me pienso como cualquier otro objeto en el que prescindo de cuál sea su modo de intuición"<sup>26</sup>.

La manera 'strawsoniana' de decir esto último es la tesis de que el uso de "yo" en la representación "Yo pienso" se refiere al sujeto de experiencias como objeto por abstracción de todos los contenidos empíricos particulares. Si la referencia al sujeto de experiencias es por abstracción de todo componente empírico cuya referencia sea un objeto empíricamente identificable, el sujeto no es referido como un objeto empíricamente identificable. Pero, primero, en tanto que algo es pensado como objeto en general, como cualquier otro objeto, el sujeto de experiencias tiene que ser pensado en la representación "Yo pienso" como un cierto 'sustrato' o *ser*<sup>27</sup>.

<sup>24</sup> Cf. Kant, *op. cit.*, A350.

<sup>25</sup> Cf. *Ibid.*, B148-50.

<sup>26</sup> *Ibid.*, B429.

<sup>27</sup> Cf. *Ibid.*, A349-50/A355/B428-9.

Segundo, según lo anterior y en tanto que la función referencial de "yo" en la representación "Yo pienso" depende de (la función no referencial de) haber llegado a captar la forma de la conciencia en general, el sujeto tiene que ser pensado en tal representación como una ser *pensante*. Finalmente, se sigue que, como en el tercer paralogismo haber llegado a captar la forma de la conciencia es haber llegado a captar la identidad de la conciencia, lo que es pensado en la representación "Yo pienso" es la identidad del sujeto como la *identidad* de un ser pensante. Así, según Strawson es legítimo el pensamiento (la autoconcepción) de la identidad del sujeto de experiencias como la identidad de un ser pensante. Ahora bien, Strawson no distingue explícitamente esta tesis del ulterior paso del psicólogo racional en el tercer paralogismo: que es legítima la autoconcepción de la identidad del sujeto de experiencias como la identidad de una *sustancia* pensante.

¿Cuál es, entonces, para Kant el error general consistente precisamente en considerar legítima tal autoconcepción? Lo primero es introducir el concepto kantiano de sustancia, atendiendo a sus dos polos. Por un lado, como un concepto puro del entendimiento, el concepto de sustancia es el concepto de un objeto como poseedor de ciertas determinaciones, cualidades o modos de existencia: "...algo capaz de ser pensado como sujeto, no como predicado de otra cosa"<sup>28</sup>. Por otro lado, como un concepto puro aplicable a la experiencia mediante el 'esquematismo' trascendental, el concepto de sustancia es el concepto de un objeto como algo permanente: "...la determinación sensible de la permanencia"<sup>29</sup>. El resultado es que en el conocimiento empírico "lo permanente es, pues, en todos los fenómenos, el objeto mismo, es decir, la sustancia, mientras que todo lo que cambia o puede cambiar pertenece únicamente al modo según el cual esa sustancia o esas sustancias existen y, consiguientemente, a sus determinaciones"<sup>30</sup>. Pero no está claro que el pensamiento de la identidad del sujeto de experiencias como la identidad de una sustancia pensante sea de suyo su autoconcepción como algo permanente: parece que en este caso emplearíamos el concepto puro de sustancia sin el esquematismo, en cuyo caso la identidad del sujeto sólo sería pensada como la de un objeto poseedor de ciertas determinaciones o modos de existencia, no como la de algo permanente, que dura en el tiempo. Ahora bien, si en el tercer paralogismo la cuestión de la identidad del sujeto de experiencia emerge al llegar a captar la identidad de la conciencia empírica que, como corriente o flujo de conciencia, tiene que consistir en cierta permanencia, entonces el pensamiento de la identidad del sujeto como la identidad de una sustancia pensante es su autoconcepción como algo permanente.

El error según Kant consistente en considerar como legítima la autoconcepción de la identidad del sujeto de experiencias como la identidad de una sustancia pen-

<sup>28</sup> *Ibid.*, A147.

<sup>29</sup> *Ibid.*, A147/B186.

<sup>30</sup> *Ibid.*, A183-4.

sante, emerge en el contexto del tercer paralogismo. En efecto, en tal contexto la proposición que Kant considera trivial y que el psicólogo racional sobrecarga es, como ya dijimos: "somos los mismos en todo el tiempo del que poseemos conciencia"<sup>31</sup>. Lo relevante es que en A364 Kant dice de esa proposición que debemos juzgarla *necesariamente*. La trivialidad que ahí ve Kant que debemos juzgar necesariamente no es obviamente, como lo es para Bennett, que cualquier cosa es la misma en todo el tiempo de su existencia, sea o no consciente. Precisamente la proposición que debemos juzgar necesariamente es que somos los mismos en todo el tiempo de existencia consciente, no de nuestra existencia, sea o no consciente. Pero si sustituimos 'somos los mismos' por 'cualquier cosa es la misma', resulta que no es verdad que debemos juzgar necesariamente que cualquier cosa es la misma en todo el tiempo de existencia consciente, en todo el tiempo en que hay una y la misma conciencia. Como ya hemos aprendido de los casos de fisión, la idea de todo el tiempo en que hay una y la misma conciencia no entraña la idea de todo el tiempo en que alguien es uno y el mismo: si el tiempo de existencia consciente es de una hora pero a la media hora, digamos en *t*, hay una fisión de personas, todo el tiempo en que hay una y la misma conciencia es, de un lado, la media hora hasta *t* en que hay una persona pre-fisión y es, de otro lado, la media hora desde *t* en que hay distintas personas pos-fisión. Desde luego, en circunstancias ordinarias juzgamos correctamente que alguien es la misma persona en todo el tiempo en que hay una y la misma conciencia: la cuestión sería que no tenemos que realizar ese juicio necesariamente. Pero es que no es esta la necesidad del juicio que Kant está evaluando. La necesidad del juicio que Kant evalúa está fundada en la validez *a priori* del juicio en cuestión, es decir, en el conocimiento *a priori* que el sujeto tiene de la proposición del juicio. Pero es evidente que Kant no puede haberse planteado que un sujeto tenga conocimiento *a priori* de que cualquier cosa *distinta de él mismo* es la misma en todo el tiempo en que hay una y la misma conciencia. Por tanto, en la proposición juzgada la expresión de *primera-persona* 'somos los mismos' es insustituible. Como ya dijimos, Kant puede ser interpretado como diciendo que el juicio en cuestión expresa el conocimiento *a priori* de la identidad de la conciencia en el tiempo. Esto significa que la proposición juzgada puede ser leída en términos de primera-persona como 'somos una y la misma conciencia en todo el tiempo del que poseemos conciencia', o como 'somos una y la misma conciencia en todo el tiempo de existencia consciente, o tautológicamente como: 'somos una y la misma conciencia en todo el tiempo en que hay una y la misma conciencia'. Según esta interpretación, todas las piezas encajan: la identidad de la conciencia en el tiempo es pensada sólo como una condición formal en la representación "Yo pienso" y una tautología es, claro está, válida *a priori*, algo que debemos juzgar necesariamente.

---

<sup>31</sup> *Ibid.*, A364.

Ahora bien, el psicólogo racional considera legítimo que la proposición en primera-persona de A364 (y otras semejantes pero más comprometidas como "la identidad de la conciencia de mí mismo en diferentes tiempos" o "la identidad numérica de nosotros mismos, que extraemos de la apercepción idéntica"<sup>32</sup>) exprese la autoconcepción de la identidad del sujeto de experiencias como la identidad de una sustancia pensante. Luego, la proposición de A364 puede ser leída como 'somos la misma sustancia pensante en todo el tiempo del que poseemos conciencia, 'somos la misma sustancia pensante en todo el tiempo de existencia consciente', o 'somos la misma sustancia pensante en todo el tiempo en que hay una y la misma conciencia'. Sin embargo, al sustituir 'somos los mismos' por 'somos la misma sustancia pensante', obtenemos una proposición no trivial que *no* podemos juzgar necesariamente. Hemos visto que pensar de la identidad del sujeto de experiencias como la identidad de una sustancia pensante es ciertamente pensar de él como algo permanente, que dura en el tiempo. Pero sucede lo siguiente. Por un lado, hemos llegado a pensar del sujeto como una *sustancia pensante* al pensarlo como un *objeto* con ciertas determinaciones o modos de existencia y al captar la forma de la conciencia en general, la razón de ser de que un estado de conciencia sea un estado de *conciencia*, y, por otro lado, llegamos a pensar la *permanencia* del sujeto al captar una demanda de la forma de la conciencia en general, la *identidad* de la conciencia en el tiempo. Esto significa que no pensamos del sujeto de experiencias como una *sustancia pensante* que permanece por haber captado la identidad de la conciencia en el tiempo, sino que pensamos del sujeto sólo como *permanente*, cualquiera que sea su condición ontológica, por haberla captado. La consecuencia es que pensar la permanencia del sujeto, su duración en el tiempo, en su condición de ser una sustancia pensante no es de suyo pensar su permanencia en términos de la mera identidad de la conciencia en el tiempo. Así, desde el puro pensamiento podríamos decir: cuál sea la permanencia del sujeto, lo que dure en el tiempo, como una sustancia pensante no depende de la identidad de la conciencia en el tiempo. Por eso, no es verdad que debamos juzgar necesariamente que somos la misma sustancia pensante en todo el tiempo en el que hay una y la misma conciencia. La idea de todo el tiempo en que hay una y la misma conciencia no entraña la idea de todo el tiempo en que somos una y la misma sustancia pensante. De nuevo, puede aplicarse aquí lo que hemos aprendido de los casos de fisión: si el tiempo de existencia consciente es de una hora pero a la media hora, digamos en *t*, hay una fisión de sustancias pensantes, todo el tiempo en que hay una y la misma conciencia es, de un lado, la media hora hasta *t* en que somos una sustancia pensante pre-fisión y es, de otro lado, la media hora desde *t* en que somos distintas sustancias pensantes post-fisión. Así, para Kant la cuestión es que realizar un juicio, tener una autoconcep-

---

<sup>32</sup> *Ibid.*, A363/A365.

ción, de la identidad del sujeto de experiencias como la identidad de una sustancia pensante, no es legítimo precisamente porque deberíamos realizar el juicio *necesariamente*: pero, dado que es concebible que seamos distintas sustancias pensantes, cada una con su duración en el tiempo, en todo el tiempo en que hay una y la misma conciencia, no cabe que sea necesario un juicio de la identidad del sujeto como la identidad de una sustancia pensante.

Kant expresa meridianamente el *quid* del error del psicólogo racional en los siguientes términos: "Supongamos ahora... unas sustancias de las que una comunica sus representaciones y la conciencia de éstas a la otra. Surgiría así una serie de sustancias, la primera de las cuales transmitiría su estado y la conciencia de éste a la segunda; ésta, a su vez, su estado propio y el de la anterior a la tercera; ésta, igualmente, los estados de todas las sustancias precedentes, juntamente con la conciencia propia y la de las demás. La última sustancia sería, pues, consciente de todos los estados de las sustancias previamente cambiadas como de estados propios, ya que tales estados, juntamente con la conciencia de los mismos, le habrían sido transmitidos. A pesar de ello, esta última sustancia no habría sido la misma persona en todos esos estados"<sup>33</sup>. Primero, Kant dice aquí que la identidad de la conciencia en el tiempo, garantizada por transmisión de los estados de conciencia como tales entre las sucesivas sustancias (pensantes), no entraña obviamente que haya una y la misma sustancia (pensante). Segundo, según una lectura justificada *ad hoc* donde 'la misma persona' sea leído como 'el mismo sujeto de experiencias'<sup>34</sup>, Kant dice que la identidad de una sustancia (pensante), en la que la identidad de la conciencia estuviera dada retrospectivamente, no entraña la identidad del sujeto de experiencias.

Estamos ya en condiciones de ver cuál es la diferencia de fondo con la versión 'strawsoniana' del error general. Para Strawson, que no se ocupa del problema de que tengamos que realizar cierto juicio necesariamente, la explicación del error está en la tesis de que el psicólogo racional sobrecarga tal juicio como expresión de un *conocimiento de objetos*. De entrada, Strawson no se cuestiona que para el psicólogo racional sea legítimo tomar un juicio de identidad de la conciencia en el tiempo como un juicio que expresa la mera autoconcepción de la identidad del sujeto de experiencias como la identidad de una sustancia pensante. Pero, como no es posi-

<sup>33</sup> *Ibid.*, A363-4n. Cf. B408. A este respecto, Kant parece estar diciendo lo mismo que Locke, J.: *Ensayo sobre el entendimiento humano*, Méjico, FCE, 1956. Locke dice: "La cuestión de si somos la misma...sustancia o no..., sea razonable o absurda, no se refiere en absoluto a la identidad *personal*. La cuestión es qué es lo que hace que se trate de la misma persona, y no si es la misma sustancia idéntica que siempre piensa la misma persona...Pues, dado que es la misma conciencia lo que le hace a un hombre ser él mismo para sí mismo, la identidad personal sólo de eso depende, ya esté unida solamente a una sustancia individual, o bien transcurra a través de una sucesión de varias sustancias", II, xxvii, 11.

<sup>34</sup> Cf., nota 5.

ble tener un conocimiento de objetos que no sean objetos de experiencia, y el sujeto pensado como una sustancia *pensante* no puede ser objeto de una experiencia posible, lo que es ilegítimo es tomar la mera autoconcepción de la identidad del sujeto como la identidad de una sustancia pensante como un conocimiento de objetos de tal identidad del sujeto. En los términos más estrictamente kantianos del paralogismo: como no es posible tener un conocimiento de objetos que no sean objetos de experiencia, y el sujeto pensado como una sustancia pensante no es un objeto de experiencia, lo que es ilegítimo es tomar el conocimiento *a priori* de la identidad de la conciencia como un conocimiento de objetos de la identidad del sujeto como la identidad de una sustancia pensante. Pero, según lo expuesto arriba, la explicación del error está en que el psicólogo racional sobrecarga el juicio en cuestión como expresión de un *conocimiento a priori*. Al valorar seriamente la condición kantiana de que tal juicio debe ser realizado necesariamente, es ilegítimo tomar un juicio de identidad de la conciencia en el tiempo como un juicio que expresa la autoconcepción de la identidad del sujeto de experiencias como la identidad de una sustancia pensante. La razón es que tomar el juicio de identidad de la conciencia como un juicio que expresa la autoconcepción de la identidad del sujeto como la identidad de una sustancia pensante entraña tomar el juicio como expresión de un conocimiento *a priori* de la identidad del sujeto como una sustancia pensante, lo que es ilegítimo. De esta forma, es la ilegitimidad de que haya un conocimiento *a priori* de la identidad del sujeto como la identidad de una sustancia pensante la que entraña la ilegitimidad de que haya la mera autoconcepción de tal identidad del sujeto. Pero, primero, ¿por qué en el tercer paralogismo el psicólogo racional tiene que comprometerse con un conocimiento *a priori* de tal identidad? A estas alturas debería resultar evidente. Dada la equivalencia entre validez *a priori* y necesidad, resulta que si la autoconcepción de la *identidad* del sujeto de experiencias es obtenida desde la autoconcepción de la *identidad* de la conciencia, y esta última es un pensamiento necesario, entonces la dependencia conceptual entre ambos hace que la autoconcepción de la identidad del sujeto de experiencias también sea *necesaria*: por su dependencia conceptual, que haya un conocimiento *a priori* de la identidad del sujeto de experiencias (para el psicólogo racional, además, como la identidad de una sustancia pensante) se sigue de que haya un conocimiento *a priori* de la identidad de la conciencia. Luego, segundo, ¿por qué en el tercer paralogismo es ilegítimo que el psicólogo racional se comprometa con un conocimiento *a priori* de tal identidad del sujeto? Como, según el psicólogo racional, la autoconcepción de la identidad del sujeto es precisamente la autoconcepción de su identidad como la identidad de una *sustancia pensante*, y por su dependencia conceptual la autoconcepción en cuestión tendría que ser necesaria, tal autoconcepción de la identidad del sujeto es *ilegítima*: hemos visto que es concebible que seamos distintas sustancias pensantes en todo el tiempo de la identidad de



la conciencia, en cuyo caso no cabe que sea necesaria la autoconcepción de la identidad del sujeto como la identidad de una sustancia pensante y, en consecuencia, no cabe un conocimiento *a priori* de la identidad del sujeto en los términos del psicólogo racional. Por tanto, lo ilegítimo no es, según la versión 'strawsoniana', que el psicólogo racional tome la mera autoconcepción de la identidad del sujeto como la identidad de una sustancia pensante como un conocimiento de objetos de tal identidad del sujeto, sino que lo ilegítimo es que la autoconcepción de la identidad del sujeto como la identidad de una sustancia pensante tenga que consistir en un conocimiento *a priori* de tal identidad del sujeto, lo que hace ilegítima la idea de la *mera* autoconcepción de tal identidad del sujeto. Lo ilegítimo no es, según la versión más kantiana, tomar el conocimiento *a priori* de la identidad de la conciencia como un conocimiento de objetos de la identidad del sujeto como la identidad de una sustancia pensante, sino que para Kant lo ilegítimo es tomar el conocimiento *a priori* de la identidad de la conciencia como un conocimiento *a priori* de tal identidad del sujeto.

Hay, sin duda, cierta tensión entre las tesis kantianas. Por un lado, parece que Kant acepta la posibilidad de autoconcebir la identidad del sujeto de experiencias como la identidad de un *ser* pensante, como lo expresa Strawson. En Kant mismo esto sería una consecuencia de la posibilidad de pensar la identidad del sujeto como la identidad de un objeto en general y, más específicamente, como la identidad del 'sustrato' en el que inhiere las representaciones: "sustrato que le sirve de base... a todos los pensamientos"<sup>35</sup>. Por otro lado, parece seguirse del tercer paralogismo que Kant no acepta la posibilidad de autoconcebir la identidad del sujeto de experiencias como la identidad de una *sustancia* pensante. Ahora bien, es que Kant habla de la posibilidad de pensar del sujeto de experiencias como un objeto en general y, por tanto, sin pensarlo como algo dado en lo relativo a su modo de existencia o determinación en alguna intuición sensible, lo que significa que no puede ser pensado en términos de los conceptos puros del entendimiento, como el de sustancia, aplicables precisamente a lo dado en alguna intuición sensible<sup>36</sup>, y así sólo cabe la posibilidad de autoconcebir la identidad del sujeto de experiencias como la identidad de un ser pensante *simpliciter*. De ahí el siguiente texto: "En la conciencia de mí mismo, en el caso del mero pensar, soy el *ser mismo*, pero, naturalmente, nada de él me es dado todavía al pensamiento"<sup>37</sup>.

La conclusión que podemos alcanzar respecto a la representación "Yo pienso" es la siguiente. Una versión 'strawsoniana' sostendría que en la representación "Yo pienso", representación desprovista de todos los contenidos empíricos particulares, de todas las percepciones genuinas y, por tanto, de toda intuición sensible, no es

<sup>35</sup> *Ibid.*, A350. Cf. A355.

<sup>36</sup> Cf. *Ibid.*, B148-50.

<sup>37</sup> *Ibid.*, B429.

posible autoreferirse al sujeto de experiencias, un objeto en general, como una sustancia pensante porque no es posible referirse a un objeto como sustancia en condiciones en las que, en ausencia de toda intuición sensible al representarnos la forma de la conciencia en general, no son aplicables los conceptos puros del entendimiento (las categorías). Luego, en la representación "Yo pienso" no es posible autoreferirse al sujeto, un objeto en general, como una sustancia pensante porque, en ausencia de toda intuición sensible, no es posible tener un conocimiento de objetos, en su modo de existencia como sustancia (pensante), del sujeto. Desde luego, esta concepción es correcta y tiene sus raíces en la Deducción Trascendental de las Categorías. Pero en el contexto del tercer paralogismo no es ésta la razón por la que en la representación "Yo pienso" no es posible autoreferirse al sujeto como una sustancia pensante. Aquí lo que cuenta, de entrada, es la identidad del sujeto como la identidad de una *sustancia* pensante: no están en cuestión los límites empíricos de aplicación de las categorías. La idea, entonces, es que en la representación "Yo pienso" no es posible autoreferirse al sujeto de experiencias, un objeto en general, como una sustancia pensante porque no es posible referirse a un objeto como una sustancia pensante en condiciones en las que, debido a la dependencia conceptual de tal referencia respecto a la representación necesaria de la identidad de la conciencia en la representación "Yo pienso", la referencia al objeto como una y la misma sustancia pensante tendría que ser *necesaria*. Por tanto, en la representación "Yo pienso" no es posible autoreferirse al sujeto de experiencias como una sustancia pensante porque, por la necesaria identidad de la conciencia, tendría que tratarse de una autoreferencia necesaria al sujeto en su modo de existencia como una y la misma sustancia pensante: pero resulta que es concebible que el sujeto exista como distintas sustancias pensantes.

¿Cuál es, pues, la función del observador externo en la corrección del error del psicólogo racional? La tesis kantiana sería que el observador externo al que yo, el hombre o la persona, me presento como un objeto de la intuición externa, no está en la situación como yo mismo sí lo estoy de caer en la ilusión de pensar *necesariamente* mi identidad como la identidad de una sustancia pensante. La razón no es, como lo expresaría Strawson, que el observador no puede tener un pensamiento de *mi* identidad como la identidad de una sustancia *pensante* porque no puede tener un pensamiento de primera-persona, una autoconcepción, de tal identidad, y *a fortiori* no puede tener ese pensamiento necesariamente. El argumento de Strawson era que el observador no puede tener un pensamiento de mí mismo como ser pensante en tanto que él no es consciente de mis estados de conciencia como estados propios ni de la identidad de mi conciencia como conciencia propia. Recuérdese que en el tercer paralogismo no cuenta la diferencia entre 'ser pensante' y 'sustancia pensante' porque no cuenta la cuestión del alcance empírico de las categorías. Ahora bien, Strawson concede que, a pesar de que para el observador soy un objeto contempla-

do desde fuera, él puede atribuirme a mí, precisamente como persona que soy para él, percepciones, pensamientos y en general conciencia así como la identidad (y la unidad) de la conciencia. La posibilidad de atribuirme desde fuera conciencia e identidad de la conciencia muestra que para tener un pensamiento de alguien como sustancia *pensante* no es necesario tener un pensamiento de primera-persona, y ello por la siguiente razón: en la representación "Yo pienso" el observador externo llega a captar la forma de la conciencia *en general*, de modo que, si bien él lo logra por ser consciente de *sus* estados de conciencia como estados de conciencia, lo cierto es que queda a su disposición la posibilidad de pensar la forma de la conciencia, que los estados de conciencia sean estados de *conciencia*, en lo relativo a cualquiera. De esta manera, no hay objeción a que el observador externo juzgue la proposición de A364 en los términos de la proposición de *tercera-persona*: 'alguien es la misma sustancia pensante en todo el tiempo del que posee conciencia', 'alguien es la misma sustancia pensante en todo el tiempo en que hay una y la misma conciencia'. La posición privilegiada del observador no consiste en que no pueda pensar de mi identidad como la identidad de una sustancia pensante al no poder juzgar la proposición de primera-persona, sino en que pudiendo pensar mi identidad como la identidad de una sustancia pensante al poder juzgar la proposición de tercera-persona, no tiene que pensarla necesariamente. Esta conclusión es una consecuencia de que él no tiene que pensar necesariamente la identidad de *mi* conciencia, es decir, no tiene un conocimiento *a priori* de tal identidad. En efecto, lo que en el análisis de Strawson hay de verdad es que la autoconcepción de mi identidad como la identidad de una sustancia pensante depende conceptualmente de ser consciente, mediante la representación "Yo pienso", de *mis* estados de conciencia como estados *propios* y de la identidad de *mi* conciencia como conciencia *propia*: es esto lo que hace que yo tenga que pensar necesariamente la identidad de *mi* conciencia y, por la dependencia conceptual, que tuviera que pensar necesariamente mi identidad como la identidad de una sustancia pensante, lo que es ilegítimo. Pero para el observador externo el pensamiento de mi identidad como la identidad de una sustancia pensante sólo depende conceptualmente de ser consciente, desde fuera, de mis estados de conciencia como estados de *conciencia* y de la identidad de mi conciencia como *conciencia*: esto significa que él no tiene que pensar necesariamente la identidad de *mi* conciencia en el sentido en que, obviamente, tiene que pensar necesariamente la identidad de *su* conciencia como conciencia *propia*, en cuyo caso no tiene que pensar necesariamente mi identidad como la identidad de una sustancia pensante. En otras palabras, la proposición de primera-persona 'soy una y la misma conciencia en todo el tiempo en que hay una y la misma conciencia' es, como vimos, una tautología que uno debe juzgar necesariamente: la segunda aparición de 'una y la misma conciencia' da ya expresión a la identidad de la conciencia como *mi* conciencia, en tanto que en mi propio caso no puedo dejar de pensar, mediante la repre-

sentación "Yo pienso", la identidad de la conciencia como conciencia *propia*. De ahí que tendría que juzgar necesariamente que 'soy la misma sustancia pensante en todo el tiempo en que hay una y la misma conciencia', lo que es ilegítimo. En cambio, la proposición de tercera-persona 'alguien (respecto a *mí*) es una y la misma conciencia en todo el tiempo en que hay una y la misma conciencia' no es una proposición trivial que el observador externo deba juzgar necesariamente: la segunda aparición no da expresión a la identidad de la conciencia como *mi* conciencia, en tanto que el observador externo sólo puede pensar, desde fuera, la identidad de la conciencia como *conciencia*. Luego, el observador externo no tendría que juzgar necesariamente que 'alguien (respecto a mí) es la misma sustancia pensante en todo el tiempo en que hay una y la misma conciencia'; es decir, "no podemos dar todavía por válido este juicio desde el punto de vista de otro"<sup>38</sup>.

Según esta interpretación, la apelación kantiana al observador externo resulta algo irónica. De entrada, parece que yo soy quien está en una posición privilegiada para pensar mi identidad como la identidad de una sustancia pensante, como Strawson concedería al psicólogo racional. Pero como la autoconcepción tendría que ser necesaria y es concebible que yo, el sujeto de experiencias, exista como distintas sustancias pensantes, resulta que no cabe tal autoconcepción. De esta manera para que mi posición privilegiada sea minada hay que alcanzar una perspectiva sobre mí mismo *desde fuera*: si yo tengo que pensar necesariamente mi identidad como la identidad de una sustancia pensante, la posibilidad de que yo exista como distintas sustancias pensantes sólo es concebible desde el *punto de vista de otro*. Y así es. Si el observador externo puede pensar mi identidad como la identidad de una sustancia pensante y no tiene que hacerlo necesariamente, entonces puede pensar mi identidad como la identidad de distintas sustancias pensantes. Pero la posibilidad de que para el observador externo yo exista como distintas sustancias pensantes descarga de peso ontológico su pensamiento de que soy una y la misma sustancia pensante: nada cambiaría para él, el hecho de que yo, el hombre o la persona, me presento como un objeto de la intuición externa y que, precisamente como persona, puede atribuirme conciencia e identidad de la conciencia, si yo existiera como distintas sustancias pensantes. Así, lo irónico es que, de un lado, es realmente el observador externo quien está en una posición privilegiada para pensar mi identidad como la identidad de una sustancia pensante ya que no tiene que pensarla necesariamente y, de otro lado, es él quien neutraliza la sobrecarga que el psicólogo racional

---

<sup>38</sup> *Ibid.*, A364. Y así sigue: "...ya que, al no encontrar en el alma otro fenómeno permanente que la representación del yo, la cual las acompaña y enlaza todas, nunca podemos establecer si este yo (mero pensamiento) no fluye igual que los otros pensamientos que él liga entre sí". Del mismo modo, con anterioridad en A362-3 ha dicho del observador: "En consecuencia, del yo que, con plena identidad, acompaña en todo tiempo las representaciones de *mi* conciencia el observador no inferirá, aunque la admita, la permanencia objetiva del propio yo (de mí mismo)".

pone en el pensamiento de mi identidad como la identidad de una sustancia pensante ya que para él es concebible que yo exista como distintas sustancias pensantes.

### **Bibliografía**

- BENNETT, J.: *La 'Crítica de la razón pura' de Kant. 2. La Dialéctica*, traducción de Julio César Armero, Madrid, Alianza Editorial, 1981. La versión inglesa es: *Kant's Dialectic*, Cambridge, Cambridge University Press, 1974.
- CASSAM, Q.: *Self and World*, Oxford, Oxford University Press, 1997.
- KANT, I.: *Crítica de la razón pura*, traducción de Pedro Ribas, Madrid, Alfaguara, 1978/1993.
- LOCKE, J.: *Ensayo sobre el entendimiento humano*, Méjico, FCE, 1956.
- MCDOWELL, J.: *Mind and World*, Cambridge Mass., Harvard University Press, 1994.
- SHOEMAKER, S.: "Persons and their Pasts", *American Philosophical Quarterly*, 74, 1970, pp. 269-85.
- STRAWSON, P. F.: *The Bounds of Sense*, London, Routledge, 1966.
- STRAWSON, P. F.: "Kant's Paralogisms: Self-Consciousness and the 'Outside Observer'", en Strawson, P. F.: *Entity and Identity and Other Essays*, Oxford, Clarendon Press, 1997, pp. 252-67.

Javier Vidal  
Paseo de Begoña, 24 11B  
33205 Gijón  
fjvidal@unav.es