



**FILO:UBA**  
Facultad de Filosofía y Letras  
Universidad de Buenos Aires

G

# El escepticismo y la condición de verdad del conocimiento

Autor:

Cresto, Eleonora

Tutor:

Olaso, Ezequiel de

1995

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título Licenciatura de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires en Filosofía.

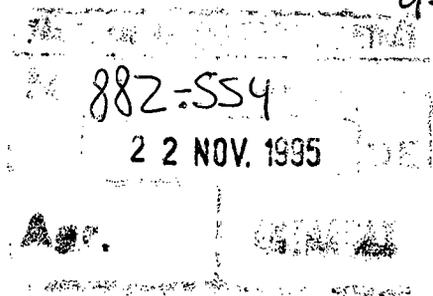
Grado



**FILO:UBA**  
Facultad de Filosofía y Letras

**FILODIGITAL**  
Repositorio Institucional de la Facultad  
de Filosofía y Letras, UBA

Tesis de licenciatura



*EL ESCEPTICISMO  
Y LA CONDICION DE VERDAD DEL CONOCIMIENTO*

**UNIVERSIDAD DE BUENOS AIRES  
FACULTAD DE FILOSOFIA Y LETRAS  
Dirección de Bibliotecas**

Director: Dr. Ezequiel de Olaso

Alumna: Eleonora Cresto ✓

Departamento de Filosofía  
Facultad de Filosofía y Letras  
Universidad de Buenos Aires

Noviembre de 1995

INDICE

	Página
INTRODUCCION	2
CAPITULO UNO:	
DOS CONCEPCIONES SOBRE LA VERDAD	11
1. Los compromisos de " <i>p</i> es verdadero sss <i>p</i> "	13
2. Compatibilidades del antirrealismo semántico	23
CAPITULO DOS:	
DIFERENTES ACTITUDES FRENTE AL PROBLEMA DEL ESCEPTICISMO	33
CAPITULO TRES:	
TEORIAS CAUSALES DEL CONOCIMIENTO Y CONFIABILISMO	40
1. Una teoría causal del conocimiento	40
2. Las tesis del confiabilismo	43
3. Algunas observaciones	50
CAPITULO CUATRO:	
EL CONOCIMIENTO EN EL MARCO DE LA TEORIA DE LA COMUNICACION	54
1. Una teoría informacional del conocimiento	54
2. La "Teoría de la información" contra el escéptico	65
CAPITULO CINCO:	
LA SOLUCION CONFIABILISTA Y EL PROBLEMA DE LA INDUCCION	73
1. Nozick y el escepticismo	75
2. La teoría informacional nuevamente	81
CAPITULO SEIS:	
LA SOLUCION CONFIABILISTA Y EL PROBLEMA DE LA VERDAD	85
1. Abandono del requisito de verdad	89
2. La teoría informacional del conocimiento y el antirrealismo semántico	95
CONCLUSIONES	105
BIBLIOGRAFIA	108

## INTRODUCCION

Hay, notoriamente, numerosos tipos de actitudes y argumentos escépticos, y respecto de muy diversos ámbitos del quehacer humano. Hay un sentido de "escepticismo" que remite al ámbito moral, o al religioso (¿refieren a algo nuestros términos morales, por ejemplo?; o ¿existe efectivamente alguna esfera de bienes trascendentes?). Suele hablarse también de escepticismo acerca de la posibilidad de diferentes clases de conocimiento; prácticamente cada disciplina científica ha desarrollado sus propias dudas escépticas. Encontramos, por último, escepticismo acerca de la posibilidad del conocimiento empírico en general: será éste el alcance del término que presupondré de aquí en adelante. Existen además distintos modos de formular el escepticismo en este último sentido: puede argumentarse que *no* sabemos si todo lo que creemos percibir no es más que una ilusión provocada por algún genio maligno, o por alguna supercomputadora a la cual se encuentra conectado nuestro cerebro (y es que, en realidad, podríamos no ser sino cerebros en una cubeta); puede argumentarse en cambio que, puesto que algunas veces creímos saber y nos hemos equivocado, nada garantiza que el error no vuelva a ocurrir, nada garantiza que tengamos conocimiento cuando así lo suponemos; puede esgrimirse también que no tenemos justificación para creer que sabemos algo no observado a partir de inducciones sobre casos observados<sup>1</sup>. Todos estos ejemplos tienen algo en común: describen situaciones en las que no podemos *asegurar* si una determinada creencia es o no es verdadera. El escéptico no nos dice que lo que creíamos verdadero era en realidad falso, sino que no tenemos forma de saber que no lo es.

Así como pueden construirse muchos y muy diferentes argumentos escépticos, así también son posibles muy diferentes tipos de respuesta a tales argumentos. La filosofía contemporánea de habla inglesa ha sido prolífica en ellos<sup>2</sup>. Algunos tipos de respuestas, en particular el tipo del que me ocuparé

---

1. Para esta diferenciación entre tipos de argumentos escépticos véase Dancy (1985), cap.1. Dancy dice, entre otras cosas, que todo escepticismo interesante depende de un argumento. He adoptado aquí esta posición.

2. En el Capítulo dos mencionaré, aunque sin desarrollar, algunas de las diferentes *actitudes* posibles ante el escéptico. En esta introducción me limitaré a caracterizar la respuesta *confiabilista*, sobre la cual se basa mi trabajo.

aquí, han surgido al calor de los intentos por "desarmar" los contraejemplos a la llamada definición tradicional del conocimiento, esto es, la definición de conocimiento como *creencia verdadera justificada*.<sup>3</sup> La idea de precisar o reformular la definición tradicional surge como un programa de investigación más o menos robusto a partir de un artículo de Gettier de 1963<sup>4</sup>. Gettier allí presenta algunas situaciones en las que no estaríamos dispuestos a afirmar que el sujeto conoce, pese a que en ellas se cumplen las tres condiciones (creencia, verdad y prueba) usualmente consideradas individualmente necesarias y conjuntamente suficientes para conocer. En tales ejemplos, las razones por las cuales el sujeto entiende estar justificado en su creencia de que *p*, *no son* las razones por las cuales *p* es verdadera<sup>5</sup>. Una respuesta estándar a este problema ha consistido en negar que en estos casos el sujeto tenga en realidad justificación para sostener su creencia. Ello en general ha ido acompañado de propuestas diversas para precisar la noción de justificación, o para reformular las condiciones de modo tal que las situaciones paradójicas mencionadas no puedan surgir. Paralelamente, este tipo de reformulación ha sido criticada por generar sus propios "contraejemplos Gettier", o tal vez por ser demasiado exigente y no permitir hablar de conocimiento cuando intuitivamente así lo deseáramos; a los nuevos contraejemplos se han sucedido luego nuevas reformulaciones y contraejemplos. A esta altura dicho programa de investigación parece más o menos agotado. Ya a mediados de los ochenta se

---

3. Generalmente se admite que la "noción tradicional de conocimiento" así entendida se remonta a Platón, aunque se han alzado voces negando que ésta sea, en efecto, una *buena* interpretación de los textos de Platón; cf. recientemente Schmitt (1992), por ejemplo, para una interpretación "externalista" de tales textos.

4. Véase Gettier "Is Justified True Belief Knowledge?", (1963).

5. Un ejemplo clásico: T. ha tenido toda su vida un Ford rojo, y S. tiene pruebas de ello -S. recuerda haberle prestado a T. dinero para adquirirlo, S. ha paseado en el Ford de T. en varias ocasiones, etc-. S. observa hoy a T. manejando un Ford rojo y, si le preguntan, cree que tiene derecho a afirmar que *sabe* que S. tiene hoy un Ford rojo. Sin embargo, T. ha vendido su viejo automóvil la semana pasada, y, extraña coincidencia, se ha ganado otro igual al anterior en una rifa. Puesto que S. desconoce estos últimos detalles, intuitivamente no diríamos que S. conoce que T. tiene un Ford, sino, más bien, que su creencia es verdadera por casualidad.

escriben algunos artículos en los que el objetivo es identificar... cuál es motivo por el cual el "problema Gettier" parece no poder resolverse<sup>6</sup>.

Entre tantos intentos por brindar condiciones del conocimiento más precisas han surgido, creo yo, algunas propuestas fructíferas. Un ejemplo de ello es la posición llamada "confiabilismo" o "fiabilismo" (en inglés, *reliabilism*). Según entiendo, ésta constituye una posición novedosa dentro de la gnoseología contemporánea, y la defensa que resulta frente al escéptico es, por lo menos, extraña, pero por ello mismo interesante. En este trabajo, entonces, me ocuparé solamente de cierto tipo de respuestas al escepticismo: aquellas surgidas en el marco de *teorías causales del conocimiento* que defiendan una concepción *confiabilista* de la condición de justificación.

Según el confiabilismo una creencia verdadera puede considerarse *conocimiento* si ha sido adquirida (causalmente) a través de un método "confiable" de formación de creencias, esto es, a través de un procedimiento que genere la mayoría de las veces creencias verdaderas<sup>7</sup>. Lo curioso consiste en que *la existencia de dicho procedimiento* es todo lo que se necesita para que una creencia esté justificada<sup>8</sup>. Es decir, no es necesario que el sujeto que tiene la creencia comprenda que existe tal proceso confiable; la justificación así entendida, y el conocimiento que resulta de ella, son procesos objetivos, naturales. Podría ocurrir que un individuo *conozca*, sin poder dar cuenta de aquello que justifica su creencia, inclusive sin poder dar cuenta de *si* su creencia en cuestión está justificada o no, o si es en efecto verdadera o no lo es. A esta característica peculiar del confiabilismo se le suele dar el nombre de

---

6. Véase por ejemplo Kirkham (1984). Una estrategia similar se adopta en libros más recientes de teoría del conocimiento. Véase también S.Haack (1993), p.7.

7. Cf. por ejemplo Armstrong (1973), Goldman (1979), (1986), Dretske (1981), (1989), Papineau (1992), Schmitt (1992).

8. También puede preferirse decir que en realidad, las opciones confiabilistas son "no justificacionistas". Esa es, por ejemplo, la idea de Dretske (1989), p.100 (nota 2), o de Robert Audi en (1988), pp.109 y ss. (aunque este último desde una perspectiva crítica al confiabilismo). Aquí consideraré, en un sentido amplio, que "confiabilismo" es un tipo de teoría en la cual se exige la existencia de un proceso confiable de adquisición de creencias como condición necesaria para conocer. Salvo referencia explícita, dejaré indeterminado si tal proceso puede utilizarse como una descripción adecuada de la condición de justificación, o si deberíamos mejor decir que la concepción del conocimiento así caracterizada no cuenta con *ninguna* condición de justificación. Personalmente, encuentro que este modo metódicamente indeterminado de hablar del confiabilismo es el más conveniente dados los intereses de este trabajo. Para una breve reseña (crítica) de los diferentes sentidos de "confiabilismo" véase Carl Ginet (1985).

"externalismo"; el externalismo se opone a "internalismo", o la posición más tradicional según la cual la justificación del conocimiento debe, para poder ser tal, ser accesible a la conciencia del sujeto cognoscente: *S* conoce que *p* sólo si *S* conoce que conoce que *p*<sup>9</sup>. Según algunos, la característica externalista del confiabilismo rompe bruscamente con siglos de una cierta tradición gnoseológica. Para sus detractores el externalismo no estaría sino "cambiando de tema" respecto de las preocupaciones básicas, y todavía legítimas, que deberían ocupar a una teoría del conocimiento<sup>10</sup>.

Es por cierto curioso que se hayan intentado respuestas al escepticismo en el marco de este proyecto con características externalistas<sup>11</sup>. Pues pareciera que es el escéptico, justamente, quien mejor puede capitalizar la sugerencia externalista de que tal vez no tengamos modo de *identificar* si en una situación dada hay justificación sobre la verdad de una creencia; para el escéptico, frente a esta catastrófica deficiencia humana poco importa que haya un proceso confiable objetivo desconocido, o que no la haya. La clave está, por supuesto, en que el confiabilista ha modificado la noción tradicional de conocimiento, y admite sin problemas que la aparente catástrofe no es tal.

Debajo de este giro conceptual se esconden varios problemas y detalles oscuros. Por ejemplo, ¿se trata tan sólo de un cambio terminológico? Si así fuera, a un trabajo de investigación sobre el confiabilismo deberíamos limitarnos a pedirle que aclarara por qué es más conveniente usar la palabra "conocimiento" en un sentido más bien que en otro. Pero tal vez haya otras cuestiones más urgentes para evaluar. Deseo llamar la atención sobre un punto que no siempre es tenido en cuenta a la hora de explicar este tipo de estrategia: el confiabilista *necesita* decir que HAY procesos confiables en la naturaleza. Es decir, el confiabilista asegura, frente al escéptico, que el conocimiento es

---

9. Véase Dancy (1985), pp. 64 y siguientes.

10. Esta última es por ejemplo la opinión de Laurence Bonjour. Véase "Externalism/internalism", en Dancy y Sosa, *A Companion to Epistemology*, (1994), pp. 132-136.

11. No todos los confiabilistas han hecho este intento. Goldman, por ejemplo, -quien es algo así como uno de los padres fundadores de las teorías causales contemporáneas y del confiabilismo- es consciente de que su posición puede engendrar dudas escépticas, pero ello no parece preocuparle en absoluto, y expresamente se ha desentendido de la relación eventual entre su teoría y los problemas clásicos del escepticismo. Véase un pasaje clave al respecto en Goldman (1967), p.372.

posible; el escéptico no se habría dado cuenta de ello porque está limitado por una definición de conocimiento inadecuada. Pero ¿qué es lo que le permite al confiabilista afirmar que es posible conocer? Podríamos aventurar dos respuestas. El conocimiento es posible,

1) porque nada *impide* que haya conocimiento; en particular, no lo impide la circunstancia de que tal vez no haya justificaciones accesibles a nuestra conciencia. El conocimiento es así una mera posibilidad de hecho.

2) porque *efectivamente* conocemos, aunque, una vez más, no es necesario que siempre podamos reconocer que lo hacemos.

No es previsible que algún escéptico quede convencido con 1), con la afirmación de la posibilidad del conocimiento. 1) es compatible con la posibilidad de que *de hecho, no conozcamos nada*, y ello no es aceptable tampoco para un confiabilista. Admitamos, entonces, que una refutación del escepticismo nos compromete con 2). Pero si 2) no es una mera profesión de fe, debería ir acompañado con la *identificación concreta* de algunos (por lo menos uno) procesos confiables de adquisición de creencias<sup>12</sup>. Por otra parte, la noción de proceso confiable está íntimamente ligada con el concepto de verdad, puesto que "verdad" es el término por medio del cual se define un "método confiable" de adquisición de creencias. Por lo tanto, el confiabilista está obligado a decir primero que estamos, de hecho, en posesión de algunas creencias *verdaderas*, y debería ser posible, una vez más, identificar alguna. Este es, según creo, el nudo conceptual de la respuesta confiabilista al escéptico: mostrar, de hecho, ejemplos de creencias verdaderas, y ejemplos de creencias verdaderas adquiridas a través del procedimiento adecuado. Pero mostrar esto no es tan sencillo, inclusive según los propios parámetros confiabilistas; es mucho menos sencillo, por ejemplo, que cambiar la definición de conocimiento.

Para evaluar si ésta es una respuesta viable, creo especialmente importante analizar con cierto cuidado la vinculación ya mencionada entre confiabilismo y verdad. Ningún confiabilista, hasta donde sé, ha estudiado en profundidad esta relación. Sin embargo, considero que ésta es una cuestión absolutamente central para el problema que nos ocupa. Pues, intuitivamente, tengo la convicción de que una respuesta plausible al escéptico tiene que ver

---

12. Más adelante (Capítulo seis) argumentaré que esta exigencia no se contradice con el espíritu externalista del confiabilismo.

con adoptar una concepción adecuada de la verdad. Podría justificar (parcialmente) esta intuición como sigue:

En primer lugar, debería observarse que, en general, los argumentos escépticos atacan lo que podríamos llamar la posibilidad de tener "conocimiento ordinario", esto es, conocimiento sobre hechos de esta clase: que somos seres humanos (y no cerebros en cubetas), que estamos rodeados por otros seres humanos en un mundo externo a nuestra conciencia, que normalmente percibimos árboles en las plazas cuando *hay, de hecho*, árboles, que necesitamos alimentarnos para vivir y que necesitamos respirar aire con oxígeno, que hay agua en los mares, que nuestro vecino X vive en el piso tercero. Todos los intentos por solucionar los "contraejemplos Gettier" presuponen que se está hablando de conocimiento ordinario; correlativamente, los casos preferidos que analizan los confiabilistas, por ejemplo, remiten en su gran mayoría a problemas de psicología de la percepción (ordinaria). Los argumentos escépticos típicos no son argumentos que ataquen específicamente cuestiones de conocimiento *científico*, tales como que las leyes de la mecánica cuántica puedan ser falsas. La razón supongo que es, entre otras cosas, que la ciencia actual ha incorporado hace tiempo el requisito de falibilidad de todos sus enunciados; no se exige seriamente la necesidad de tener enunciados *verdaderos* (a diferencia de, por ejemplo, *probablemente verdaderos*) para que algo cuente como "conocimiento científico". (Por supuesto, hay puntos de contacto entre los dos ámbitos; algunos enunciados científicos dependerán de cuestiones que llamé de conocimiento ordinario, y viceversa.) Pero la mayoría de las formulaciones del conocimiento ordinario han persistido en pedir la *verdad* de una cierta creencia como requisito para conocer; la posibilidad del error, que en ciencia pertenece a la misma dinámica del progreso científico, en algunos sectores del ámbito cotidiano resulta intolerable.

¿Pedir que la verdad sea condición necesaria del conocimiento no es, entonces, pedir demasiado? En efecto, podríamos plantearnos la conveniencia de retirar el requisito de verdad para conocer. Sin embargo, lo cierto es que *usamos* el término "verdad", y por lo menos en el ámbito ordinario, no solamente como idea regulativa. Si no deseamos resignarnos a dejar los términos "verdad" o "conocimiento" sin aplicación, tal vez haya que formular el problema de otro modo. El requisito de verdad parece exigir demasiado *si* entendemos la verdad de un modo que llamaré "semánticamente realista", esto es, "verdad" como cierta propiedad que pueden tener o no algunas entidades

(proposiciones, creencias, oraciones)<sup>13</sup> con independencia de cualquier capacidad epistémica para reconocer si una entidad determinada la posee o no. Es pedir demasiado porque de esta definición de verdad se infiere que, por principio, siempre podemos estar equivocados y creer verdadero algo que en realidad no lo es<sup>14</sup>. *Esto* es todo lo que el escéptico necesita para plantear su desafío.

Pero la anterior no es la única noción de verdad posible. Podemos adoptar alguna versión no-realista de la verdad (antirrealismo semántico) en la cual la situación escéptica no surja inevitablemente. Para saber si semejante procedimiento es conveniente, además de meramente posible, sería deseable analizar qué compromisos trae aparejados tal noción de verdad, en particular, si es imprescindible asociarla con una postura metafísica idealista<sup>15</sup>, o si el antirrealismo semántico es compatible con alguna forma de realismo más amplio, metafísico (realismo acerca de la existencia del mundo externo, de las entidades que existen en él, etc., como diferente de realismo acerca de la verdad de nuestros enunciados).

La intuición pre-teórica de la que intenté dar cuenta en los párrafos anteriores, entonces, afirma que muy posiblemente detrás de cierto tipo de estrategias contra el escéptico, si son efectivas, encontraremos alguna reformulación de la noción realista de la verdad. *En este trabajo me ocuparé de analizar si la estrategia confiabilista es compatible, o no, con esta intuición.* En especial, trataré de responder preguntas como:

---

13. Usaré indistintamente como portadores de verdad a las oraciones, enunciados, proposiciones o creencias, sin especificar tampoco si se trata en cada caso de tipos (*types*) o ejemplares (*tokens*). Creo, con Kirkham, que es preferible un uso laxo de la idea de "portador de verdad". Véase Kirkham (1992), capítulo 2.

14. Esta observación puede encontrarse en Putnam (1981). No deseo adoptar aquí su solución al problema (el "realismo interno") ni hablaré especialmente de ella en este trabajo. Mi impresión, como volveré a manifestar en el Capítulo 1, es que deben discriminarse claramente las cuestiones metafísicas de las cuestiones semánticas, y el "realismo interno" de Putnam parece, al menos en parte, no discriminar tesis de ambas categorías (incluye un cierto modo de concebir la verdad y supone a la vez una toma de posición acerca de las entidades que existen en el mundo).

15. En cuyo caso, como dice Dancy, nos toparíamos con una solución efectiva frente al escepticismo... aunque casi tan implausible como él. (Dancy (1985), pp.35-36).

- ¿Qué noción de verdad adopta implícitamente el confiabilismo? O, mejor, ¿con qué nociones de verdad es compatible?

-Si el confiabilismo fuera únicamente compatible con el realismo semántico -y hay razones para pensar que esto es así- ¿puede la solución confiabilista neutralizar las consecuencias escépticas de una concepción realista de la verdad?

Desarrollaré los temas comentados en el siguiente orden:

En el Capítulo uno me ocuparé de algunas cuestiones preliminares al examen puntual de la relación entre confiabilismo, escepticismo y verdad. Específicamente, estudiaré en qué difieren entre sí concepciones semánticas realistas y antirrealistas. Prestaré especial atención a los compromisos metafísicos del antirrealismo semántico, lo cual será de interés a la hora de evaluar si el antirrealismo semántico es una posición plausible. Esto último, a su vez, importa para averiguar si una respuesta al escepticismo puede provechosamente hacer uso de nociones semánticas no realistas. Consideraré las observaciones de Putnam (1978) sobre Field (1972); Dummett (1978), y otras.

En el segundo capítulo esbozaré un esquema de las diferentes actitudes que pueden adoptarse frente al problema del escepticismo. El objetivo será mostrar cuál es la ubicación del confiabilismo, *tal como yo lo veo*, en relación con el espectro de las soluciones posibles. Será, por lo tanto, una suerte de capítulo introductorio al problema; solamente *mencionaré* diferentes opciones frente al escepticismo, pero no las explicaré con detalle.

En el Capítulo tres haré una exposición breve de las notas distintivas de las teorías causales del conocimiento y del confiabilismo. Me basaré sobre todo en Goldman (1967) y Goldman (1979). Expondré también los que a mi juicio constituyen los principales supuestos y méritos de este tipo de concepción. En el capítulo siguiente presentaré un caso particular de posición confiabilista, a saber, la "teoría informacional del conocimiento", tal como la desarrolla Dretske en *Knowledge and The Flow of Information* (1981) a partir de algunas herramientas de la teoría matemática de la comunicación. Según entiendo, ella constituye una de las opciones confiabilistas más originales e interesantes. En lo que sigue me ocuparé sobre todo de la solución confiabilista al escepticismo que es posible construir en el marco de dicha teoría informacional.

En los capítulos cinco y seis trataré de dar cuenta, respectivamente, de lo que sostengo son los dos grandes supuestos de una respuesta confiabilista frente al escéptico (según habrá surgido de algunas observaciones de los

capítulos anteriores). El confiabilista asume que podemos hacer un uso no problemático de la inducción, y que es posible identificar ciertas creencias como verdaderas. Así pues, en el Capítulo cinco abordaré la relación entre confiabilismo y el problema de la inducción. Recordaré en este punto la solución al escepticismo propuesta por Nozick (1981), y, en la segunda parte del capítulo cinco, intentaré un modo de *evitar* el problema de la inducción con base en algunas consideraciones de Dretske.

En el Capítulo seis, por último, encararé el problema de la verdad, esto es, intentaré afrontar el supuesto de que es posible identificar creencias verdaderas. En la primera parte analizaré la posible compatibilidad del confiabilismo, y en particular de la teoría informacional del conocimiento, con la idea de que no es necesaria la verdad para conocer; para ello haré algunas observaciones sobre modos de interpretar el externalismo, y apelaré a Dretske (1989) para rastrear un criterio de aceptabilidad de creencias como verdaderas compatible con el confiabilismo. En la segunda parte examinaré la relación entre confiabilismo y antirrealismo semántico. Sugeriré que puede concebirse una respuesta confiabilista satisfactoria frente al escéptico que haga uso para ello de una noción no realista de la verdad. Las afirmaciones de esta última parte del Capítulo seis serán en gran medida tentativas y programáticas.

Como ha quedado implícitamente establecido, es un objetivo general de este trabajo evaluar la efectividad de un cierto tipo de respuesta frente al problema del escepticismo. Los planteos escépticos, con sus relatos de ciencia ficción y otros menos fantásticos, son capaces de ejercer una extraña fascinación. Sugiero, como punto de partida provisorio, que una solución *buen*a debe habilitarnos para comprender el por qué del innegable atractivo de los argumentos escépticos.

CAPITULO 1:  
DOS CONCEPCIONES SOBRE LA VERDAD.

Este capítulo está especialmente destinado a aclarar en qué difieren, de un modo general, concepciones realistas o antirrealistas de la verdad. Mencionaré para ello qué se entiende bajo los conceptos de realismo semántico, realismo metafísico, antirrealismo semántico y antirrealismo metafísico; indagaré luego qué compatibilidades o incompatibilidades pueden encontrarse entre estas posiciones. Los resultados a los que arribemos serán de vital importancia para abordar, en los capítulos siguientes de la tesis, cuál es la concepción de verdad que se presupone en algunas respuestas al escepticismo, o tal vez cuál es la concepción que mejor nos habilita para una respuesta plausible.

Los conceptos de realismo/antirrealismo semántico/metafísico suelen ser confusos, y no siempre aparecen usados de la misma manera en la bibliografía, por lo que optaré por aclarar cómo voy a usarlos en esta ocasión. Por "realismo semántico" entenderé la idea de que todo enunciado tiene un valor de verdad determinado, verdadero o falso, que es independiente de que se lo conozca o no; la verdad de un enunciado no depende de nuestros medios para conocerla. El realismo semántico defiende así básicamente la bivalencia. El "realismo metafísico" entiendo que afirma algo muy diferente: que existe un mundo exterior al sujeto de conocimiento, y que las entidades que pueblan el mundo lo hacen con independencia de nosotros, de nuestro conocimiento de ellas, e inclusive de las capacidades expresivas de nuestro lenguaje; esto último lleva a no identificar los límites del lenguaje con los límites del mundo (suele decirse que el mundo puede ser mucho más rico que lo que los términos de nuestro lenguaje permitan describir)<sup>16</sup>. Ninguna de estas dos

---

<sup>16</sup>. La caracterización que he dado de realismo metafísico es básicamente la que puede encontrarse en Devitt, (1984), p.11 y ss., donde se afirma que el realismo supone una dimensión de existencia (de un mundo externo) y otra de independencia (de dicho mundo respecto del sujeto). Devitt no desea aplicar el término "realismo" a cuestiones semánticas, pues opina que en este uso (equivocado) se encuentra la raíz de muchas confusiones. Estoy básicamente de acuerdo en que no deben mezclarse cuestiones semánticas con cuestiones metafísicas (todo mi esfuerzo por distinguir tipos de realismos tiene como fundamento mi intuición de que deben evitarse tales mezclas) pero a los fines

caracterizaciones define unívocamente una determinada posición semántica o metafísica; es decir, bajo estos rótulos podrían razonablemente ubicarse autores diferentes con sus respectivas teorías, lo que implica cierto margen de flexibilidad en ambas nociones. Entenderé además por "antirrealismo semántico" la negación de las tesis del realismo semántico, y por "antirrealismo metafísico" la negación de las tesis del realismo metafísico. Usaré indistintamente, por otra parte, los términos "antirrealismo" y "no realismo" (semántico o metafísico).

El objetivo fundamental en lo que sigue será analizar si es posible adoptar una posición que hable de antirrealismo semántico y realismo metafísico simultáneamente. Es decir: si es posible una posición que afirme simultáneamente que a) las entidades que pueblan el mundo tienen existencia con independencia del sujeto cognoscente y b) la *verdad* de los enunciados no es independiente de nuestros medios para conocerla. Pero antes diré algo sobre los compromisos semánticos o metafísicos del esquema " 'p' es verdadero sss p ".

---

prácticos de este trabajo encuentro *útil* hablar de "realismo/ antirrealismo semántico". Creo que no hay razones para no usar tales nombres, una vez que se hacen las aclaraciones del caso: no pretendo dar connotaciones metafísicas con estos términos. En cuanto a la relación entre lenguaje y mundo en el realismo, Armstrong por ejemplo es uno de los que admite explícitamente que los límites del lenguaje no pueden ser la medida de la realidad. Lo que aquí llamo realismo metafísico él lo denomina "realismo científico", y lo conecta con una teoría realista sobre el problema de los universales Cf. Armstrong (1978), especialmente Vol.2.

Para una distinción diferente entre realismos, cf. por ejemplo Horwich, (1982). Allí Horwich comienza por diferenciar tres formas de realismo: epistemológico, semántico y metafísico. A decir verdad, no me seduce demasiado el modo en que las caracteriza. Por ejemplo, no me parece que por realismo metafísico (=R.M.) haya que entender básicamente una determinada teoría de la verdad (i.e., la verdad como correspondencia). Pues, en caso de hacerlo, se confunden los límites del R.M. con el R.semántico. Si se afirma que en el R.M. "*the concept of truth involves a primitive non-epistemic idea*", p.182, ¿no es esto otro modo de decir que la verdad de un enunciado es independiente de nuestros medios para conocerla?. Pero esta última era una de las tesis que creo apropiadas para definir al R.semántico. La culpa de esta caracterización a mi juicio errónea tal vez haya que buscarla en Putnam, en su descripción de lo que es el R.M. como opuesto al realismo "interno", en Putnam (1978 c) y (1981). "Realismo metafísico" no intenta tener aquí ni remotamente la connotación peyorativa que le da Putnam en (1981), ni implica forzosamente todo lo que allí Putnam opina que implica.

### 1. Los compromisos de " 'p' es verdadero sss p "

Tal vez uno de los pocos hechos incontrovertidos sobre la verdad es que " 'la nieve es blanca' es verdadera si y sólo si la nieve es blanca"; de manera general y esquemática, " 'p' es V. sss p " <sup>17</sup>. Esto es lo que Tarski llamó la "convención T" o criterio de adecuación material; según Tarski, de cualquier definición *adecuada* de verdad, esto es, de una definición que respete nuestras intuiciones pre-teóricas sobre la verdad, deberían poder deducirse todas las instancias de sustitución del esquema T <sup>18</sup>. Tarski, como es sabido, ofrece su propia definición de verdad para lenguajes formalizados, pero ha habido intentos de trasplantar su idea para lenguajes naturales (por ejemplo, el intento de Davidson <sup>19</sup>).

El punto es si la concepción en cuestión sobre la verdad nos compromete con una posición semánticamente realista, y tal vez con lo que he llamado realismo metafísico, o si no es forzosa dicha vinculación. De hecho, en la formulación de Tarski, del lado derecho de la equivalencia de la Convención T se dice expresamente que no es asunto del "lenguaje", sino del "mundo".

El asunto nos importa porque la opción confiabilista sobre teoría del conocimiento, que analizaremos en la segunda parte de la tesis, previsiblemente lo mínimo que requerirá para admitir una concepción sobre la verdad es que deben valer las equivalencias de la forma T <sup>20</sup>. Abordaré aquí, entonces, el siguiente problema: si el uso del formalismo de Tarski para la definición de verdad, o al menos el uso de algunas ideas derivadas de las de Tarski, debe comprometernos forzosamente con una determinada perspectiva semántica (por ejemplo, con el realismo semántico) y/o metafísica.

Por supuesto, se puede estar de acuerdo en la conveniencia de admitir el esquema T y *no* adherir a la *definición* de Tarski. Podemos pensar, por ejemplo, que los bicondicionales T son todo lo que una definición de verdad

---

17. Cf. Horwich, en Dancy y Sosa (1994), p.511.

18. Cf. Tarski (1933) y (1944).

19. Para una compilación de los artículos de Davidson hasta 1982 véase Davidson (1984). Véase una explicación detallada del "programa Davidson" en Kirkham (1992), cap.8.

20. Véase por ejemplo Nozick (1981), p.680, nota 6 al capítulo tres.

necesita<sup>21</sup>. En ese caso, no por ello las conclusiones de esta parte serán inútiles, sino que aún valdrán, restringidamente, para clarificar los compromisos del esquema T, esto es, del criterio tarskiano de adecuación material.

Una definición de verdad, según Tarski, debe cumplir con una serie de requisitos, bastante conocidos, en los que aquí no profundizaré<sup>22</sup>. Por ejemplo, debe cumplirse con el requisito de corrección formal: la definición de verdad debe hacerse para un lenguaje formal L completamente especificado, y que no sea semánticamente cerrado (esto es, se evitará que L posea términos semánticos y nombres para sus propias expresiones), desde un metalenguaje esencialmente más rico que L; ello pretende evitar el surgimiento de paradojas semánticas. Mencionaré con un poco más de detalle *en qué consiste* la definición de verdad que ofrece Tarski, pues resultará útil para los comentarios posteriores de este capítulo. Tarski define la verdad de una oración en términos de "satisfacción". Para explicar esta idea habría que recordar la distinción entre "oración", o fórmula cerrada, y "función proposicional", o fórmula abierta:

- (1)  $x_1$  es rojo.
- (2)  $x_1$  es un tomate.
- (3)  $x_1$  está entre  $x_7$  y  $x_8$ .
- (4)  $x_3$  es alto o  $x_3$  es verde.

(1), (2), (3) y (4) son ejemplos de fórmulas abiertas, pues contienen variables libres. El modo standard de "cerrar" una fórmula abierta es agregándole un cuantificador (existencial o universal) o reemplazando las variables por constantes de individuo:

- (5)  $a_1$  es rojo.
- (6)  $(\exists x_1) (x_1 \text{ es un tomate})$ .
- (7)  $(x_1) (\exists x_7) (\exists x_8) (x_1 \text{ está entre } x_7 \text{ y } x_8)$
- (8)  $(x_3) (x_3 \text{ es alto o } x_3 \text{ es verde})$ .

---

21. Cf. la teoría de la verdad como "desentrecomilladora", o la teoría minimal de Horwich, en Horwich (1990).

22. Cf. Tarski, *op.cit.*

Una fórmula cerrada puede verse como el caso límite de una fórmula abierta, esto es, una fórmula con 0 variables libres.

De las funciones proposicionales no se dice que son verdaderas o falsas, sino satisfechas o no satisfechas por secuencias infinitas de objetos. Una secuencia es un conjunto de elementos en el cual importa sobre todo el orden en que dichos elementos aparecen. Así, (1) es satisfecha por una secuencia  $S$  si el primer objeto de  $S$  es rojo; (3) es satisfecha por  $S'$  si el primer objeto de la secuencia  $S'$  está entre el séptimo y el octavo; (4) es satisfecho por  $S''$  si el tercer objeto de  $S''$  es alto o es verde.

Con base en esto, podemos construir las siguientes definiciones, por ejemplo, para un lenguaje que posea solamente los predicados "ser rojo" y "estar entre", y las conectivas "no" (negación) y "o" (disyunción):

Una fórmula  $\alpha$  es satisfecha por una secuencia infinita  $S$  ssi:

1. ( $\alpha$  = "xi es rojo" y el  $i$ ésimo objeto de  $S$  es rojo); o
2. ( $\alpha$  = "xi está entre  $x_j$  y  $x_k$ " y el  $i$ ésimo objeto de  $S$  está entre los objetos ubicados en los lugares  $j$  y  $k$ ); o
3. ( $\alpha$  = "no  $\psi$ ", y  $S$  no satisface  $\psi$ ); o
4. ( $\alpha$  = " $\lambda$  o  $\mu$ ", y  $S$  satisface  $\lambda$  o  $S$  satisface  $\mu$ ); o
5. ( $\alpha$  = " $(\forall x_i) \psi$ ", y *toda* secuencia que difiera de  $S$  a lo sumo en el  $i$ ésimo lugar satisface  $\psi$ ); o
6. ( $\alpha$  = " $(\exists x_i) \psi$ ", y *alguna* secuencia que difiera de  $S$  a lo sumo en el  $i$ ésimo lugar satisface  $\psi$ ).

Pero dado como han sido construidas las definiciones de "satisface" en los pasos 5. y 6., de allí se sigue que si una oración (cerrada) es satisfecha por una secuencia cualquiera  $S$ , entonces es satisfecha por toda secuencia. Con lo cual Tarski puede definir la verdad del siguiente modo:

Una oración es verdadera si y sólo si es satisfecha por toda secuencia .

"Existe algo que es rojo" es una oración satisfecha por una secuencia cualquiera  $S$  (esto es, por toda secuencia) en el caso en que haya alguna secuencia  $S''$  cuyo primer elemento sea rojo, y, por ende, satisfaga la fórmula abierta que resulta de eliminar el cuantificador existencial. "Todo objeto es rojo" es satisfecha por alguna secuencia cualquiera (por toda secuencia) en

caso de que el primer elemento de toda secuencia sea rojo. Esta última oración normalmente es falsa; una oración falsa, como puede verse, no puede satisfacerse con *ninguna* secuencia.

Recursos similares pueden utilizarse si se desea incorporar nombres al lenguaje, lo cual Tarski no considera. Aquí hemos construido la definición de verdad *à la Tarski* para un lenguaje con solamente dos predicados: el predicado monádico "ser rojo" y el predicado de tres lugares "estar entre". Es de notar que *debe* haber un número finito de predicados. a partir de allí, las conectivas lógicas nos permiten, por supuesto, construir oraciones tan complejas como queramos.

De la definición de verdad que hemos explicado Tarski muestra que pueden derivarse todas las instancias del esquema T, de adecuación material.

Creo que puede intentarse un abordaje a nuestro problema sobre los compromisos de una definición tarskiana de la verdad a partir de algunas críticas de Putnam a Hartry Field en las Conferencias John Locke de 1976<sup>23</sup>. Para llevar a cabo esta crítica, como veremos, Putnam propone un modo no estándar de interpretar el formalismo de Tarski, en lo que llama una posición "cuasi-intuicionista", inspirado en ideas de M. Dummett.

Field escribe en 1972 un artículo crítico sobre la concepción de la verdad desarrollada por Tarski<sup>24</sup>. En él opina que Tarski, fiel a ciertos principios fisicalistas y con la idea de construir las bases de una semántica científica, habría querido reducir el concepto de verdad a nociones *no* semánticas, pero fracasó en el intento: consiguió tan sólo reducir "verdad" a *otras* nociones semánticas, por ejemplo, a la noción de "aplicarse" un predicado a ciertos objetos. La teoría de Tarski podría así, según Field, descomponerse en la conjunción de dos partes muy desparejas. En una de ellas, el Tarski "valioso", se efectúa la reducción de la noción de verdad a otras nociones semánticas (por ejemplo, la noción de satisfacción). En la otra se efectúa la pretensión fallida de reducir también estas otras nociones, (reducirlas, presuntamente, a términos exclusivamente físicos, lógicos y/o matemáticos) y lo único que se consigue es proporcionar "listas". Por ejemplo, listas de predicados en relación con los objetos a los que decidimos primitivamente que dichos predicados se aplican: estas listas definirían la

---

<sup>23</sup>. Putnam (1978 a)

<sup>24</sup>. Field, H., "Tarski's Theory of Truth", (1972)

“aplicación primitiva” de los predicados. (Field considera también que podría brindarse una estrategia equivalente para tratar a) nombres: mediante listas que definan la “denotación, o referencia, primitiva”, y b) letras de función que seguidas de un término singular dan como valor otro término singular. Claro está, en los trabajos de Tarski en realidad no se alude a nombres de ninguna especie.) Tales listas son tan poco interesantes como reducción como lo sería una definición de valencia de la forma

“(E) (n) (E tiene valencia n = E es potasio y n es +1, o..... o E es azufre y n es -2)”<sup>25</sup>.

Field sugiere que se debe recurrir a una teoría fisicalista de la referencia para completar la tarea que Tarski dejó inconclusa. Dicha teoría consistiría en especificar que “el hablante sólo se refiere a una cosa al utilizar un término si el uso que le da se encuentra en una relación causal definida (por la ciencia empírica, en el mismo espíritu en que ésta ha “definido” *agua* como H<sub>2</sub>O y *temperatura* como energía cinética molecular) con ese algo, o con cosas de la misma clase que ese algo (...)”<sup>26</sup>.

La estrategia de Putnam, para intentar una defensa de Tarski, consiste entre otras cosas en esbozar una *justificación* del uso de listas cuando se da la denotación primitiva. Para llevar a cabo esta defensa apelará a ciertas ideas intuicionistas<sup>27</sup>. El intuicionismo explica el significado de las conectivas lógicas en términos de *demostrabilidad constructiva y absurdo* y no en términos de verdad y falsedad. Así, afirmar “p” es afirmar “p es demostrable”, y afirmar “¬p” (= “no p”, “¬” consideraremos que es el símbolo intuicionista para la negación) es afirmar el absurdo de la demostrabilidad de “p”, esto es, es demostrable que una prueba de “p” implica la demostrabilidad de un absurdo evidente. “p.q” significa “p es demostrable y q es demostrable”, “p v q” significa “existe una prueba de p o una prueba de q y se puede saber cuál es”, y “p ⊃ q” significa “existe un método que una vez aplicado a una cualquier prueba de p lleva a una prueba de q, y una prueba del método realiza esto”.

---

25. Field, H., Op. cit., p.362.

26. Putnam, op.cit., p.13.

27. En realidad, el primer paso de Putnam consiste en reivindicar el formalismo de Tarski apelando a algunas ideas de Stephen Leeds, que aquí no comentaré. De todas maneras, véase S. Leeds, (1978).

Entre otras cosas, todo lo anterior impide que algo como " $p$  o no  $p$ " sea teorema del cálculo proposicional intuicionista.

Sería conveniente recordar aquí ciertas ideas vertidas por Dummett en un artículo de 1959 sobre el problema de la verdad y el antirrealismo<sup>28</sup>. Dummett se pronuncia allí en contra de que el significado de una oración esté dado por sus condiciones de verdad. Ello no puede sostenerse simultáneamente con la idea de que, para entender lo que la verdad es, basta considerar aquello de que " $p$  es verdadero si y sólo si  $p$ ", pues, de sostener ambas tesis simultáneamente, nos veríamos envueltos en un círculo vicioso. A saber, para comprender  $P$  a partir de saber en qué circunstancias  $P$  sería verdadero, debe comprenderse primero lo que significa decir que  $P$  es verdadero. Si la única explicación de esto último es que es lo mismo que afirmar que  $P$ , llegamos a que tenemos que presuponer aquello que estábamos tratando de explicar: esto es, tenemos que presuponer en qué consiste la comprensión de  $P$ . (Dummett (1990), p.72). La posición de Dummett será la de no explicar más "el sentido de un enunciado estipulando sus condiciones de verdad en términos de los valores de verdad de sus constituyentes, sino estableciendo cuándo puede afirmarse en términos de las condiciones bajo las cuales pueden afirmarse sus constituyentes." ((1990), p.83). Es decir, podemos afirmar que un enunciado es verdadero o falso "sólo cuando  $P$  sea un enunciado de tal clase que podamos en un tiempo finito encontrarnos en posición de estar justificados o bien en afirmar o bien en negar  $P$ ; es decir, cuando  $P$  es efectivamente un enunciado decidible." ((1990), p.82); habrá casos, entonces, en los que un enunciado *con sentido* no será ni verdadero ni falso. Los conceptos de verdad y falsedad no tienen un particular papel central en la teoría del significado, y a efectos de usar dichos conceptos, para la mayoría de los contextos comunes nos bastan las explicaciones del tipo " $P$  es verdadero si y sólo si  $P$ ", y " $P$  es falso si y sólo si no  $P$ ". ((1990), p.84)<sup>29</sup>.

Dummett, entonces, propone una teoría intuicionista del significado; es el intuicionismo matemático, justamente, aquel que dice que "verdad" y "falsedad" deben entenderse como "demostrabilidad constructiva" y "absurdo"

---

28. cf. "Truth", en Dummett (1990), pp.65-90. Dummett es de algún modo el inventor del término "antirrealismo semántico"; no entraré en el problema de si su posición es además metafísicamente idealista.

29. Véanse algunas críticas muy claras a esta posición en Haack (1980), pp.112-115. Véase también Kirkham (1992), capítulo 8.

respectivamente. *Pero además* el intuicionismo "está disponible para otras regiones de la realidad [fuera de la matemática] como una alternativa a la concepción realista del mundo" ((1990), p.84); ideas muy similares a estas últimas son las que inspiran la concepción de Putnam. Putnam ensaya una *aplicación* del intuicionismo matemático al conocimiento empírico, lo que resulta en una posición que llama "cuasi-intuicionismo"<sup>30</sup>.

Putnam alude en primer lugar a un breve artículo de Gödel donde se establece que, si reinterpretemos las conectivas clásicas de modo que

1) la negación y la conjunción clásicas se identifiquen con la negación y conjunción intuicionistas, pero

2) la disyunción clásica se identifique con " $\neg (\neg p \wedge \neg q)$ ", y

3) la implicación material clásica se identifique con " $\neg (p \wedge \neg q)$ ",

entonces todo teorema del cálculo proposicional clásico sale también como teorema en el cálculo proposicional intuicionista. En este caso, traduciendo el cálculo clásico al cálculo intuicionista, "los significados [de las conectivas] siguen hallándose fuera de lo clásico, pues se explican en términos de *demostrabilidad* y no de *verdad* y *falsedad*." (Putnam (1978 a), p.39). Y, en este espíritu, Putnam sugiere interpretar las conectivas *del cálculo clásico no traducido* en un sentido intuicionista, esto es, en términos de demostrabilidad constructiva; el formalismo de Tarski puede interpretarse también en este sentido.

Así, Putnam propone *interpretar* de modo intuicionista el significado de las conectivas lógicas en el *cálculo clásico*, y, en segundo lugar, propone "aplicar" este proyecto a las ciencias empíricas: reemplazar la idea de "demostrabilidad constructiva" del intuicionismo matemático por la de "demostrabilidad constructiva derivada de (una reconstrucción coherente y adecuada de) los postulados pertenecientes a la ciencia empírica aceptada en ese momento" (p.39). Entonces, "si B1 es la ciencia empírica aceptada en *una* época y B2 es la ciencia empírica aceptada en una época *diferente*, entonces, de acuerdo con esta interpretación 'cuasi-intuicionista', las mismas conectivas lógicas se referirían a la 'demostrabilidad en B1' cuando se utilizaran en B1 y a la 'demostrabilidad en B2' cuando se utilizaran en B2. Las conectivas lógicas

---

<sup>30</sup>. Putnam ofrece esta respuesta en el marco de un análisis sobre la convergencia de las teorías científicas (la idea de que las teorías científicas tienden en el largo plazo hacia la verdad), lo cual no nos interesa en este momento: el cuasi-intuicionismo sería un medio de dar un sentido coherente a la idea de que en realidad *no hay* convergencia en la ciencia.

modificarían su significado de manera sistemática al cambiar el conocimiento empírico". (p.40).

Ahora bien, apelar al cuasi-intuicionismo negaría, según la óptica de Putnam, uno de los postulados realistas fundamentales: para un realista, "resulta *por lo menos lógicamente posible* que un juicio se siga de B1 (nuestro mejor cuerpo teórico actual) y que no sea *verdadero*." (p.48). Y, evidentemente, esta es una posibilidad que no puede darse dentro de la perspectiva del cuasi-intuicionismo, una vez que reinterpretemos convenientemente el significado de "verdadero". Por otra parte, otro de los efectos de apelar al cuasi-intuicionismo sería que "la existencia se vuelve intrateórica": "En la definición de verdad y referencia de Tarski, (a) "'Electrón' tiene referencia" equivale a (b) "Existen electrones". Pero si 'existen' se interpreta de manera intuicionista, (b) sólo afirma que (c) "Existe una descripción D tal que 'D es un electrón' es demostrable en B1". Y esto puede ser verdadero (para un B1 adecuado) aunque no existieran electrones. En suma, el efecto de reinterpretar de manera intuicionista las conectivas lógicas es que la 'existencia' se vuelve *intrateórica*." (p.41). En fin, el cuasi-intuicionismo así entendido encuadra, por lo menos, dentro de la definición que dimos al comienzo sobre "antirrealismo semántico", y está por verse si no tiene además implicaciones *metafísicamente* antirrealistas (la observación de que "la existencia se vuelve intrateórica" parece apuntar en ese sentido). No nos preocupará por ahora cuál es exactamente el alcance antirrealista de esta posición, sino su compatibilidad con una definición de verdad *à la Tarski*.

Lo que se ha mostrado hasta ahora es que es posible conservar la lógica clásica y *también la semántica de Tarski* y sin embargo transformar la noción de verdad en algo semejante a la "aseverabilidad justificada", dando una interpretación cuasi-intuicionista de las conectivas. Esto es, la lógica clásica no es incompatible con una interpretación "no clásica" y la teoría de Tarski valdría también en este último caso. Debemos para ello reinterpretar el término "verdadero" como "demostrabilidad constructiva", aunque manteniendo los aspectos formales de la definición. Pero si el cuasi-intuicionismo nos compromete, por lo menos, con una posición semánticamente antirrealista, ello muestra que la semántica tarskiana es valiosa más allá de que decidamos optar por el realismo o el antirrealismo. El formalismo ideado por Tarski es viable aún en caso de dar una interpretación no clásica de la lógica y, consecuentemente, no realista. Considerar en toda su magnitud el valor del formalismo de Tarski incluye una explicación, o mejor

dicho, *justificación*, del uso de listas cuando se da la denotación primitiva, justificación que Field ha fallado en reconocer. A saber:

En primer lugar debemos notar, según Putnam, que las listas tienen una estructura muy especial. "'Electrón' se refiere a los electrones" puede reescribirse como " 'electrón(x)' es satisfecho por  $y_1$  si y sólo si  $y_1$  es un electrón", y ello se parece bastante al criterio T - lo que no es casual, pues "verdadero", como es sabido, es el caso 0-ádico de satisfacción. Con lo cual Putnam sugiere generalizar el criterio T en un "Criterio S": "una definición adecuada de satisfacción para L debe reconocer como teoremas a todas las instancias del siguiente esquema:

" '[P(x<sub>1</sub>, ..., x<sub>n</sub>)]' se *satisface* con la secuencia  $y_1$ , ...,  $y_n$  si y sólo si P( $y_1$ , ...,  $y_n$ )." (p.43-44)

En segundo lugar, Putnam propone una suerte de argumento trascendental (según sus propias palabras). *Necesitamos* tener una noción metalingüística que satisfaga al criterio T y a su generalización en S, puesto que necesitamos tener predicados que funcionen formalmente del modo descrito por T y S. Y esta noción es, justamente, la noción de verdad/satisfacción, o, eventualmente, su reinterpretación en términos no clásicos. Es decir, "conservaríamos al criterio S aunque eligiéramos una acepción intuicionista o cuasi-intuicionista para las conectivas lógicas." (p.44). En fin, no solamente no está mal en absoluto haber proporcionado listas para la referencia primitiva, sino que era necesario darlas, para así poner de manifiesto las propiedades *formales* que queremos que tengan nuestras nociones de verdad y referencia.

En conclusión, la estructura formal de la definición de la verdad de Tarski rescata ciertos rasgos básicos inherentes a *cualquier* interpretación de la verdad, y por lo tanto, intrínsecamente valiosos. Es decir que podemos conservar la definición tarskiana, o al menos el criterio de adecuación material, e inclusive una generalización de dicho criterio para la noción de satisfacción, sin estar obligados a realizar una elección realista (semántica y/o metafísica).

No es Putnam el único autor que afirma esto. He optado aquí por elegir un texto de Putnam, simplemente, porque creo que su estrategia tiene algunos méritos, como ser:

- 1- *mostrar* en qué consistiría una interpretación no realista de Tarski, y
- 2- explicar cómo es que la posibilidad de tal interpretación revela ciertas propiedades formales importantes que deseáramos previsiblemente que respete cualquier concepción de la verdad que adoptemos, lo que le permite a su vez sacar conclusiones para hacer frente al polémico artículo de Field.

Podríamos ahora volver sobre la cuestión de cuáles tesis realistas niega exactamente el cuasi-intuicionismo (aunque sin entrar en demasiados detalles, puesto que no es la idea del capítulo hacer una exégesis sobre Putnam).

¿Por qué el cuasi-intuicionismo supone, según Putnam, una opción antirrealista? El realismo semántico dice que la verdad es independiente de que la conozcamos o no, y el cuasi-intuicionismo, basado en las correspondientes ideas de Dummett, nos dice que la verdad de un enunciado está dada por la capacidad concreta de construir una prueba de él. Indudablemente, el cuasi-intuicionismo es entonces la antítesis del realismo semántico. ¿ES además incompatible con el realismo metafísico, tal como lo he definido más arriba? Eso pareciera si, como ya mencionamos, aparentemente en el cuasi-intuicionismo la "existencia" pasa a ser intrateórica, según las propias palabras de Putnam. Si se desdibuja la línea entre lo que es afirmable y lo que es verdadero, pareciera decir Putnam, lo que en realidad ocurre es que se desdibuja la línea entre *lo que existe intrateóricamente* y *lo que existe*, a secas. Es decir, *se desdibuja la tesis central del realismo metafísico*.

Esto último pareciera apoyar la idea de que el antirrealismo semántico (el cuasi-intuicionismo lo es) implica el antirrealismo metafísico. Esto es, el cuasi-intuicionismo aparece como una tesis semántica según la cual la verdad de un enunciado se define como la capacidad de mostrar una prueba de él, pero una conclusión de ello parece ser la tesis metafísica según la cual lo que haya en el mundo depende en algún sentido de nuestras teorías, es decir, del sujeto cognoscente. O, al menos, así pareciera entenderlo implícitamente Putnam.

Ahora bien, ¿es esto realmente así? Creo que hay aquí varias cosas juntas para evaluar. En primer lugar, creo que habría que tomar con cierta cautela la noción de "existencia intrateórica" que maneja el cuasi-intuicionismo. "Existencia intrateórica" no tiene por qué, en principio, ser equivalente a "afirmación de que la existencia de las entidades de las que se habla depende del sujeto". "Existencia intrateórica" es más bien una expresión metafórica para aludir a "descripción de ciertas entidades/ situaciones, en términos de las expresiones disponibles en una determinada teoría". Y ello, *per*

se, nada dice sobre cuestiones más amplias, metafísicas. No intenta decir *nada* sobre lo que sucede más allá del lenguaje, ni para negar que exista tal cosa como un mundo de entidades independientes, ni para afirmarlo.

En segundo lugar, pareciera que el único medio de mostrar que el antirrealismo semántico implica el antirrealismo metafísico sería probar que *no puede darse* una posición que conjugue antirrealismo semántico y realismo metafísico a la vez. Pero ello está por verse.

## 2. Compatibilidades del antirrealismo semántico.

Tal vez deberíamos encarar la pregunta de esta manera: ¿por qué se cree o se sospecha usualmente que NO es posible la compatibilidad entre antirrealismo semántico y realismo metafísico? La razón fundamental reside, posiblemente, en la noción de *correspondencia* : es habitual suponer que el antirrealismo semántico (en el sentido en que yo lo estoy usando aquí) rechaza la idea de que la verdad sea explicada por una relación de correspondencia entre entidades lingüísticas (o mentales) y entidades extralingüísticas (o extramentales), mientras que sería esa misma relación de correspondencia la que aseguraría que hay *algo más* que meramente nuestros enunciados y nuestras creencias. Es decir, la relación de correspondencia estaría garantizando nuestro compromiso con el realismo metafísico.

Trataré de mostrar aquí que el antirrealismo semántico no tiene por qué negar que exista algún tipo de relación de correspondencia. Sólo niega un cierto modo de entender en qué consiste esta relación; niega, digamos, un sentido "fuerte" de entender la correspondencia.

Creo que un buen comienzo para ello es retomar la crítica de Putnam a Field en las Conferencias. Field, podemos sospechar, ES un buen representante de realismo semántico y realismo metafísico a la vez. Esta vez será útil relacionar la crítica de Putnam a Field con ciertas afirmaciones de un artículo posterior de Putnam, "Referencia y comprensión", en el que comenta un modelo de la comprensión lingüística basado en ideas de Carnap y Reichenbach.

De acuerdo con lo que vimos en la primera parte del capítulo, para Putnam es altamente criticable la idea de Field de "completar" a Tarski

brindando la caracterización de nexos causales únicos entre un término y su referencia. (Sí es legítimo, en cambio, pensar en una aclaración de Tarski que tenga que ver con reflexiones sobre el modo de entender las conectivas, o algún otro comentario metateórico de índole similar.) Pero podríamos muy bien pretender alguna justificación un poco más sustantiva de porqué Field está equivocado; hasta ahora Putnam se había limitado a expresar que no era *necesario* completar a Tarski de este modo. El último paso de Putnam, a partir de la Conferencia Tercera, consiste en decir que tampoco es posible.

Antes de explicar la argumentación de Putnam, convendría recordar cuál es la relación de correspondencia en la que Field está pensando. En una perspectiva fisicalista como la de Field, una determinada referencia física causa el uso de ciertos términos y, subsidiariamente, la verdad de ciertas oraciones. HAY una relación privilegiada que conecta un término con su referencia, y una teoría de la verdad satisfactoria debería dar cuenta de esta relación. En el fondo todo hecho semántico debe poder ser reducido a hechos físicos, más lógicos y matemáticos (por ello lo de "fisicalismo", según expresa Field<sup>31</sup>), o, si ello es imposible, las nociones semánticas irreducibles deben ser abandonadas. En palabras de Field "si podemos argumentar desde nuestra teoría del mundo que a la noción de enunciar algo verdadero, o referir a una cosa particular, no se le puede dar sentido en términos fisicalistas (por ejemplo, argumentando que cualquier noción semántica que tenga sentido fisicalista puede ser explicada en términos de Skinner, y que las nociones de verdad y referencia no pueden ser explicadas en tales términos) entonces, si tal argumento es convincente debemos ser llevados a concluir que, si hemos de permanecer fisicalistas, las nociones de verdad y referencia deben ser abandonadas." (Field (1972), p.373).

¿Qué consecuencias trae afirmar la posibilidad de identificar dicha relación de correspondencia?. El efecto más obvio parece ser la posibilidad de identificar de manera unívoca una cierta referencia en el mundo. Es decir, el realismo metafísico con el que nos comprometemos nos dice que cada término del lenguaje se corresponde con alguna referencia definida, y que es dicha referencia la que, en cada caso, ha determinado causalmente el uso del

---

<sup>31</sup>. Véase sin embargo el artículo de Mc.Dowell (1978), según el cual el "semanticalismo" es en realidad compatible con el fisicalismo. Véase también Kirkham (1992), pp.202-208, para una interesante crítica a Field según la cual i) Tarski no entiende el fisicalismo en el modo en que Field lo hace; y ii) hay serias dudas sobre si el fisicalismo de Field es un programa filosóficamente valioso.

término. Pero entonces detrás de las nociones de referencia y verdad hay una porción identificable de ontología que puede explicar por qué un término refiere o una oración es verdadera. ¿Qué significa que "el pasto es verde" es una oración verdadera? Bueno, que existe cierta entidad, que es el pasto, a la que se le aplica el predicado "ser verde"; los términos de nuestra oración refieren a algo en el mundo, ... y allí afuera hay una porción definida de mundo que nos muestra de qué se trata todo el asunto. Que pueda explicar los fenómenos semánticos es *parte* de la concepción de una tal noción de correspondencia. Si esta idea resultare imposible, en realidad lo que cae es todo *un modo de entender la realidad*, esto es, cae la idea de entenderla *como la contrapartida exacta de los términos de nuestro lenguaje* y, en última instancia, la responsable exclusiva -con prescindencia de consideraciones subjetivas (alusiones a "niveles de interés", "marcos conceptuales" o cosas por el estilo)- de que los términos de nuestro lenguaje sean tales. Tal noción de correspondencia, evidentemente, es correlativa a *un cierto modo* de entender las tesis del realismo metafísico (esto es, las tesis sobre la existencia e independencia del mundo externo), un modo que inevitablemente implica el realismo semántico.

Según Putnam, no pueden caracterizarse nexos causales únicos entre un término y su referencia sencillamente porque *no existe* tal relación única. Para ello se basa en la tesis de Quine acerca de la indeterminación de la traducción y la referencia<sup>32</sup>; no hay una relación única porque los datos con que contamos subdeterminan las posibles interpretaciones que podemos dar sobre la referencia de los términos. Putnam propone que se debe considerar como traducción -de un idioma a otro- o como definición de verdad correcta a la que *explique* mejor la conducta del hablante ((1978 a), p.54). Pero una explicación es una "noción de interés relativo"; si dos culturas que poseen *intereses diferentes* son enfrentadas a la tarea de traducir el mismo lenguaje, puede ocurrir que dispongan de expresiones cortas para referir a entidades por completo diferentes (y que expliquen igualmente bien la conducta del hablante, claro está). Comprender que hay niveles de interés en toda explicación es lo que permite, según Putnam, que la doctrina de Quine de la indeterminación de la traducción resulte verosímil.

Se podría pensar en remediar el problema de la relatividad del interés, dice Putnam, mediante la siguiente estrategia: ampliar el intento fiscalista de

---

<sup>32</sup>. Véase para ello sobre todo el capítulo dos de *Word and Object*, Quine (1960).

Field de modo que la descripción de una teoría causal de la referencia comprenda también los intereses del hablante y otros detalles relevantes. Pero ello equivaldría a un intento por simular "la capacidad humana total", lo cual requeriría un esfuerzo científico para el cual no estamos capacitados. "No existen motivos para creer que se pueden definir explícitamente algunas de estas nociones ["traducción", "explicación"] de acuerdo con lo deseado por los filósofos fisicalistas, ni explicitar todas las 'restricciones' que imponemos de manera inconsciente al uso de esas nociones, sin contar con *un detallado modelo explicativo sobre la organización funcional del ser humano total*. Y afirmo que esperar esto es utópico." (p.78). La conclusión final de Putnam es que el programa de Field es, sencillamente, "utopismo científico" (p.76).

Quisiera llamar la atención sobre la peculiar posición en la que quedamos ubicados si Putnam tiene razón en criticar a Field de este modo, haciendo uso de ideas de Quine. Putnam efectúa esta crítica pero *no pretende por ello deshacerse de la noción de correspondencia*. El nudo de nuestro problema sobre la compatibilidad entre realismos y antirrealismos creo que consiste en decir cómo puede ser posible una crítica del estilo mencionado sin abandonar la idea de correspondencia. Putnam, lamentablemente, no lo aclara en lo más mínimo, pero no por ello la situación es desesperante. Según creo, hemos dado con una posición correspondentista sobre la verdad que, al parecer, todo lo que viene a querer decir es que hay un mundo exterior que es en última instancia el responsable de que lo que afirmamos sea verdadero o falso. Sin embargo, todo intento por aclarar la referencia se transforma en una misión imposible, y, en tanto tal, el intento es condenable (es, como ya vimos, "utopismo científico"). Pero entonces la posición en cuestión se resuelve en una suerte de realismo (metafísico) *de mínima*, o una noción "débil" de correspondencia, pero *no por ello menos "realista"*, según el cual nuestras afirmaciones versan sin dudas sobre el mundo externo, aunque en un sentido bastante indeterminado. Por si quedan algunas dudas de que esto es efectivamente lo que hace Putnam, creo que este pasaje del final de la Segunda Conferencia es más que elocuente:

" 'Electrón' se refiere a los electrones", ¿de qué otro modo podríamos decir a qué se refiere 'electrón' desde el interior de un sistema conceptual en el cual 'electrón' es un término primitivo? Tan pronto como analizamos los electrones -a saber, que 'los electrones son partículas de tal y cual masa y unidad de carga negativa'- podemos afirmar que "'electrón' se refiere a partículas de tal y cual masa de carga negativa", pero entonces 'carga' (o cualesquiera que sean las

nociones primitivas de nuestra nueva teoría) tendrá que ser explicada 'trivialmente', esto es, de acuerdo con el criterio S. Dado el predicamento de Quine (¿de Kant?), según el cual existe un mundo real pero sólo podemos describirlo en términos de nuestro propio sistema conceptual (¿deberíamos acaso utilizar el sistema conceptual de otro?), ¿resulta sorprendente que la referencia primitiva posea este carácter en apariencia trivial?" ((1978a), p. 44-45)<sup>33</sup>.

Hasta ahora tan sólo hemos mostrado con qué herramientas es posible cuestionar una noción de correspondencia como la que pareciera requerir el realismo semántico. Y sugerimos que no parece imposible la idea de rescatar una noción más débil o indeterminada de correspondencia que baste para asegurar las tesis del realismo metafísico aunque no tal vez las del realismo semántico.

Para afianzar esta última idea mostraré un ejemplo en el cual, según creo yo, podemos interpretar que un autor *explícitamente* advierte sobre el antirrealismo (semántico) de una posición, y sin embargo trata de separarlo de una noción más amplia de idealismo. En "Referencia y comprensión" Putnam propone una "explicación verificacionista de la comprensión lingüística", basada en ideas de Carnap y Reichenbach<sup>34</sup>. La idea es la siguiente:

Supongamos un modelo del oyente/hablante -puede pensarse en un autómeta- tal que está dotado de una lógica inductiva (calcula grados de confirmación), una lógica deductiva, un ordenamiento preferencial (una forma idealizada de modelar las estructuras preferenciales del ser humano) y una regla de acción (por ejemplo, aumentar al máximo la utilidad estimada). Tal hablante procede concediendo probabilidades a las oraciones expresadas por otros, y está autorizado a formular él mismo oraciones cuando estas tienen una probabilidad que supera cierto valor. El sistema, además, es holístico, en el

---

<sup>33</sup>. Hasta donde puedo ver, ciertas ideas del párrafo anterior parecen en el fondo un antecedente directo de lo que más adelante Putnam llamará "realismo interno" o "realismo pragmático". Por supuesto, hay diferencias - no por nada Putnam mismo considera que se ha operado un cambio en su pensamiento-, pero las bases para ese cambio ya estaban dadas. Tal vez el punto nodal de dicho cambio consista en lo siguiente, según se afirma en "Realismo y Razón" (1978 c): la anterior opinión de Putnam era que podían darse descripciones equivalentes del mundo sin que por ello el realista se sintiera molesto, pues aquello era equivalente a la posibilidad que hay de hacer mapas alternativos de la Tierra a partir de proyecciones diferentes (Mercator, Polar, etc.); la opinión más reciente consiste en afirmar que *es una propiedad fundamental del mundo mismo* el admitir tales proyecciones distintas.

<sup>34</sup>. Putnam (1978 b). Véase también Reichenbach (1938).

sentido de que "las condiciones en las cuales se emite cualquier oración en particular y la conducta que resulta cuando se expresa una oración dada no dependen de una única cosa que pueda ser denominado el 'sentido' de la oración, sino del sistema completo. Cambiar la lógica inductiva, o la deductiva, o la función de utilidad afectará hasta cierto punto las condiciones de emisión y las respuestas de la conducta para cada oración del lenguaje (aunque, desde luego, los enunciados comunes sobre macroobservables estarán asociados con las experiencias canónicas de 'verificación' y 'falsificación'.)" (Putnam (1978 b), p.116-117).

En primer lugar habría que analizar si efectivamente el modelo idealizado de hablante que propone Putnam se compromete con el antirrealismo semántico según nuestros propios cánones. Esto último equivaldría a sostener -recordamos- que la verdad de los enunciados no es independiente del sujeto. Pues bien, en dicho modelo, según nos dice Putnam, es el sujeto el que va otorgando valores de verdad (en realidad, grados de probabilidad) de acuerdo con la lógica inductiva que tiene incorporada. Las oraciones sobre entidades inobservables, por ejemplo, son evaluadas de acuerdo con el "peso" que dicha oración tiene *a efectos de deducir o inducir consecuencias (observacionales) a partir de ella*. El significado de una oración no está dado por las condiciones en las cuales sería verdadera (entendiendo "verdadera" al modo realista, esto es, como condiciones de las cuales podría tal vez no tenerse nunca experiencia), sino por el peso que la oración tiene, en el marco de nuestra lógica inductiva<sup>35</sup>. En realidad la noción de "verdadero" *entendida al modo realista* no desempeña en esta teoría ningún papel relevante, puesto que más bien lo que encontramos son, simplemente, "instrucciones para asignar un gran peso a determinadas oraciones cuando se tienen ciertas experiencias, instrucciones para expresarlas, para llevar a cabo transformaciones sintácticas, para producir una conducta no verbal (por ejemplo, 'levanta un brazo cuando estimes que 'levantar un brazo posee la máxima utilidad prevista'), etc. (p.118). Según creo, deberíamos decir -Putnam no lo hace, pero no veo que pudiera considerarlo incorrecto- que, en todo caso,

---

35. Al igual que Dummett, Putnam desea rechazar la idea de que la verdad de una oración sea *anterior* al significado de la misma (cf. Putnam (1978 b) p.130; véase también (1978 c), p.148)). Putnam también aclara, sin embargo, que esta "teoría probabilística" no constituye en realidad una explicación correcta del *significado*, sino de la *comprensión lingüística*, pues el significado es en realidad un compuesto de varias nociones, entre ellas la de 'referencia' y 'estereotipo' (nociones de las que no da cuenta la teoría verificacionista que está considerando). Cf. Putnam (1975).

"oración verdadera" ha sido convenientemente traducida *dentro de dicha teoría* a "oración que puede ser legítimamente usada como premisa en ulteriores inferencias que cumplan con ciertos requisitos", y, en esta reinterpretación, la noción de verdad está irremediabilmente ligada a *nuestra capacidad* para inferir situaciones que, a su vez, seamos *capaces* de evaluar experimentalmente.

Todo lo anterior hace comprensible el rótulo de teoría *verificacionista* ; la idea subyacente tiene sin dudas puntos de contacto, o al menos un cierto aire de familia, con la propuesta de Dummett o, mejor, con el uso que hace Putnam de dicha propuesta.

El próximo paso es notar cómo una teoría verificacionista puede evitar implicaciones antirrealistas, en el sentido de lo que hemos llamado antirrealismo metafísico. Bajo el apartado "Una objeción realista", Putnam expresa: "Imagino que algunos realistas bien pueden sentirse incómodos con la explicación delineada en estas páginas acerca de la comprensión, incluso como modelo idealizado y excesivamente simplificado. ¿No es este modelo verificacionista? ¿Y no es el verificacionismo, en el fondo, una forma de idealismo?" (p.131). Sugiero que lo que se está preguntando Putnam aquí es: ¿no es el antirrealismo semántico un pasaporte seguro para el antirrealismo metafísico?

Para responder esto, por empezar, Putnam intenta distinguir su teoría de un fenomenalismo tradicional. Este último intentaría reducir lo no observado a objetos observados (y lo "observado", dicho sea de paso, puede traducirse en términos de "impresiones sensoriales" o algo por el estilo). Pero ello no es lo que la teoría que Putnam propone, sino que la idea es, como ya se dijo, asignar un peso a una oración sobre inobservados tal que de ella puedan inferirse consecuencias, sin hablar en ningún momento de "reducción" (p.132). Y "¿Por qué habrá de ser imposible desde el punto de vista realista que (1) toda oración significante posea uno u otro peso dentro de una situación observable, y (2) toda *diferencia* de significado se refleje en cierta diferencia de peso dentro de una situación observable? (...) Al menos, si ello es así o no, debería ser una pregunta abierta para el realista en cuanto realista, y no algo descartado por el realismo." (p.133). Después de plantear estas cuestiones Putnam no parece dar mayores argumentos para que esta "pregunta abierta" se responda en el sentido que Putnam desea.

Es razonable sostener que efectivamente lo que Putnam quiere hacer es rescatar las tesis de lo que aquí hemos llamado realismo metafísico. Habría que

volver con algo de cuidado entonces sobre cómo entiende Putnam la noción de "correspondencia". Putnam recuerda que "la definición de verdad para un lenguaje natural se halla *subdeterminada* por el tipo de consideraciones causales que discutimos aquí, además de que bien puede haber cierta *relatividad del interés* en nociones como 'verdad' y 'referencia' dentro del contexto de la traducción radical." (p.126; subrayado mío). Y agrega de inmediato, "Pero el hecho de que las explicaciones sean relativas con respecto al interés no implica que no puedan ser *correctas*, lo cual constituye lo único que importa para la argumentación que estoy desarrollando a favor del realismo." (p. 126). Uno puede muy bien preguntarse, ¿qué quiere decir aquí "correctas"? No puede significar que la relación establecida entre cada hecho o situación del mundo y una oración del lenguaje es la *correcta*, pues ya ha aclarado (lo hace con detalle en la Conferencia Tercera) que no existe una relación única que sea la "correcta". Lo que parece ser "correcta" es la estrategia de aludir a un mundo independiente como forma de explicar el éxito en la comunicación de los seres humanos. Pero entonces, aludir a una "correspondencia" entre mundo y lenguaje no equivale más que a la afirmación un tanto vaga de que Putnam desea comprometerse con dicha tesis del realismo metafísico, esto es, con la tesis de que existen entidades que pertenecen a un mundo externo e independiente del sujeto.

Llegamos aquí a un punto central: ¿qué es lo que hace *verdadera* a una oración, el *peso* que le asigna el sujeto o el hecho de que se "corresponda" con algo en el mundo?. La respuesta del antirrealista semántico podría ser: las dos cosas, pero sobre todo, en mayor medida, la *explicación* de la verdad recae sobre el *peso* asignado por el hablante. No niega (necesariamente) que exista, de hecho, una correspondencia con el mundo; pero no es esa correspondencia la que determina que llamemos verdadera a una oración, sino que ello queda determinado por la *evaluación* que podamos hacer de tal correspondencia. ¿No tenemos acaso la fuerte intuición pre-teórica de que "el gato está sobre la mesa" es verdadera *porque* el gato, efectivamente, está sobre la mesa? Pero eso no es sino decir que tenemos la fuerte intuición de que el criterio T es correcto, lo cual, como vimos, puede aceptar tanto un realista como un antirrealista. Podríamos inclusive sostener que dicho criterio *presupone* la noción de correspondencia<sup>36</sup>: esto tan sólo presupondría que ciertas tesis muy básicas del

---

<sup>36</sup>. Esta es la idea de Davidson en "True to the facts", en Davidson (1984), pp.37-54. (En artículos posteriores se retracta de ello).

realismo metafísico son ciertas. Y la correspondencia entendida de este modo, de un modo, digamos, "quineano", en el fondo no "explica" nada sobre la verdad, pues cualquier pregunta ulterior sobre tal "correspondencia" está condenada a quedar sin respuesta<sup>37</sup>. Un realista semántico podría contraargumentar que si admitimos la noción de correspondencia, aún debilitada, el concepto de verdad lleva implícito inevitablemente la idea de un mundo externo. El antirrealista semántico, por su parte, podría perfectamente conceder esto último, pues ello no muestra que sea necesario pensar la noción de verdad como independiente del uso del lenguaje por una comunidad de sujetos; *lo que muestra, en todo caso, es tan sólo que no podemos pensar ciertas nociones semánticas sin una correspondiente toma de posición metafísica.*<sup>38</sup>

En fin, postular una relación de correspondencia que padezca de las restricciones inevitables derivadas de la defensa de ciertas tesis quineanas, entonces, no tiene la fuerza suficiente para *explicar* el por qué de la verdad de una oración determinada, *pero sí se compromete con la existencia del mundo externo*. La respuesta más natural a por qué una oración es verdadera es, simplemente, porque hacemos cierto *uso* de esa oración en el lenguaje, y aquí sí podríamos explayarnos en aclarar en qué consiste exactamente ese uso. Pero esto ya es antirrealismo semántico. Con lo cual mostramos lo que queríamos mostrar: que es posible *conjugar realismo metafísico con antirrealismo semántico*, y que por lo tanto el antirrealismo semántico no necesariamente implica un compromiso metafísico idealista. La teoría del uso que Putnam adopta en "Referencia y comprensión", junto con el tipo de consideraciones metafísicas que allí hace, creo que es un buen ejemplo de tal conjunción, si bien Putnam no lo expresa de este modo -y no estoy demasiado segura de si aceptaría que ello es efectivamente lo que está haciendo.

---

37. Es evidentemente un problema para esta parte de mi argumentación el hecho de que no haya criterios demasiado compartidos acerca de qué constituye una buena explicación de un concepto. Si apelamos a la idea intuitiva de que el *explanans* debe ser más comprensible que el concepto a explicar, la noción indeterminada de correspondencia no parece un buen candidato para explicar nada. Igualmente, si preferimos la idea un poco más elaborada de que una explicación no debe aludir a propiedades subjetivas como la inteligibilidad, y sí en cambio debe procurar que el *explanans* forme parte de una teoría, entonces recordemos que en el caso que estamos considerando *tenemos* una teoría sobre el uso del lenguaje pero ninguna sobre una noción tan extremadamente vaga de correspondencia.

38. Esto último es abiertamente cuestionado, por ejemplo, por Horwich (1990).

¿Por qué alguien puede querer rechazar el antirrealismo semántico? Mi impresión es que en general se teme que implique el antirrealismo metafísico, y pocos filósofos actualmente estarían dispuestos a reivindicar una posición idealista más o menos extrema. Pero sostener esta implicación no es correcto, según he tratado de mostrar, y el antirrealismo semántico es compatible con más de una posición de carácter metafísico. Puede haber otros motivos para rechazar el antirrealismo semántico, desde luego. Uno de ellos tal vez sea el deseo de defender la bivalencia; el realismo semántico está inmediatamente comprometido con ella, y si bien tal vez no esté claro si todo antirrealismo tendría que ser "antibivalente", muchas posiciones dentro del antirrealismo semántico lo son; el intuicionismo, por ejemplo, desde luego que lo es. En este caso no estarían en juego cuestiones metafísicas, entonces, sino tal vez de simplicidad teórica u otros motivos semejantes<sup>39</sup>.

---

<sup>39</sup>. Véase por ejemplo Quine, (1986), p.45-51.

## CAPITULO DOS:

### DIFERENTES ACTITUDES FRENTE AL PROBLEMA DEL ESCEPTICISMO

En la Introducción de esta Tesis presenté la idea de que hay más de un estilo de argumentación escéptica, y, previsiblemente, más de una forma de encarar dichos argumentos. Es decir que podemos encontrar muy diferentes actitudes frente al problema del escepticismo. Podemos:

1) Desconocer el desafío.

2) Aceptar que el escéptico tiene razón. Conocemos muy pocas cosas o tal vez ninguna. Pero no es importante. Para sobrevivir nos basta con tener creencias medianamente justificadas (que sin embargo no constituyen conocimiento).

3) Intentar una estrategia para mostrar que el escéptico está equivocado. Tal estrategia podría consistir en:

3.1.) Mostrar que hay algo intrínsecamente errado en la argumentación / actitud del escéptico. Deberíamos, entonces,

3.1.1.) Mostrar que los argumentos escépticos ya presuponen que el conocimiento es posible. (Por ejemplo, argumentos trascendentales).

3.1.2.) Recurrir a la noción de mundo de certezas no cognoscitivas (aplicar a ellas la idea de "conocimiento" equivale a cometer una suerte de error categorial), verdades "de marco", del sentido común, etc.

3.2.) Intentar bajar las pretensiones para que algo pueda acceder al nombre de "conocimiento". (Las dudas escépticas serían legítimas, pero sólo se aplican a estándares demasiado elevados de conocimiento)

3.2.1) La verdad no es un requisito para conocer.

3.2.2) La verdad es una noción epistémica, al menos en parte.

3.2.3.) El conocimiento no requiere certeza; basta con una *justificación* más modesta de nuestras creencias<sup>40</sup>.

Por supuesto, este esquema no pretende dar un panorama exhaustivo; tampoco lo explicaré con demasiado detalle. El objetivo de la clasificación, simplemente, es proporcionar un modo de ubicar el planteo confiabilista (claramente, en 3.2.3.) tal como yo lo concibo.

¿Con qué criterio podemos decidir cuál es una opción *buena* frente al escepticismo? Parece razonable elegir la estrategia que más se adecue a nuestras intuiciones pre-analíticas. Por ejemplo, encuentro que la actitud señalada bajo el punto 2) decididamente no da cuenta de ellas. "Aceptar que el escéptico tiene razón" implica dejar sin aplicación la noción de conocimiento. Sin embargo, notoriamente usamos este concepto, constantemente, tanto en el lenguaje ordinario como en el plano científico. El partidario de 2) debería exigir entonces, correlativamente, una modificación de los usos lingüísticos de manera de hacerlos consistentes con su propuesta, y debería también, por supuesto, explicar por qué esto último es deseable. ¿Cuál es la razón de querer disponer de una noción tan estricta de conocimiento que resulte *en principio*

---

40. Me limitaré a citar un autor paradigmático de cada actitud, sólo con el objeto de ilustrar mejor a qué me refiero con cada punto. 2): R. Kirkham (1984); véase la nota siguiente. 3.1.1): D. Davidson, por ejemplo en "Thought and Talk", en Davidson (1984), pp. 155-170: si comprendemos el significado de "creencia" no podemos sino rechazar la posibilidad de "error masivo" en ellas. 3.1.2.): L. Wittgenstein (1969): dado cierto juego de lenguaje, *ni siquiera tiene sentido* dudar de proposiciones como "esta es tu mano". 3.2.1.): Villoro (1982): una concepción adecuada del conocimiento cuenta con las siguientes condiciones (individualmente necesarias y conjuntamente suficientes): (i) S cree que p - (ii) S tiene razones objetivamente suficientes para creer que p; "verdad" es un concepto epistemológicamente inaplicable. 3.2.2.) Putnam (1981): la verdad es aceptabilidad racional idealizada (p. 65); Putnam argumenta que una noción realista de la verdad lleva inevitablemente al argumento escéptico según el cual podríamos ser cerebros en una cubeta. 3.2.3.): Dretske (1989); véase capítulos siguientes. (Téngase en cuenta también que, como se verá más adelante, Dretske no piensa que esté reformulando la condición de justificación, sino que la está *eliminando*). Para un panorama alternativo sobre respuestas frente al escepticismo véase Goldman (1986), cap.2, pp.28-41.

inutilizable?<sup>41</sup>. Hay, por cierto, una respuesta para estas críticas: siempre puede decirse que es útil una noción de este tipo como idea regulativa, como meta que guía nuestro proceder. Sin embargo, no creo que el caso de 2) pueda considerarse bajo esta perspectiva. Pues la idea misma de un ideal regulativo tiene sentido si nos sirve para observar que algunas cosas (en este caso, algunas creencias) se acercan a él más que otras. Pero el escéptico, por lo menos el escéptico radical con que estamos tratando aquí, creo que no aceptaría esto tampoco. Es decir: si existe algún conjunto de enunciados más cercanos que otros al ideal de conocimiento, sin duda entre ellos deberían estar ejemplos como "yo estoy escribiendo en este momento", "el mundo existe desde hace más de cinco minutos", "no somos cerebros en una cubeta"<sup>42</sup>. Pero el escéptico no tiene ninguna consideración especial para con ejemplos de esta clase; más bien todo lo contrario. Entonces, en un sentido importante el escéptico subvierte el uso habitual del término "conocimiento". ¿Por qué sería *lícito*, como quiere 2), llevar a cabo tal maniobra?

La argumentación anterior no muestra que la movida del escéptico no sea lícita, sino sólo que no va de suyo que ello es así. Ello brinda apoyo a la idea de que es en principio correcto intentar mostrar que el escéptico está equivocado. No me ocuparé en esta Tesis de las opciones que he agrupado bajo

---

41. El siguiente pasaje de Kirkham (1984) ilustra bastante bien lo que creo es una concepción *equivocada* sobre el conocimiento y el escepticismo: " Podemos concluir que hay dos sentidos de conocimiento. Uno de ellos se utiliza en la vida ordinaria cuando decimos que conocemos toda clase de proposiciones incluyendo muchas que no están justificadas del modo en que mi análisis lo requiere. El otro es revelado cuando los contraejemplos Gettier nos fuerzan a aguzar nuestras intuiciones conceptuales sobre el conocimiento. Por otro lado, podemos simplemente concluir que el concepto de conocimiento es desesperadamente confuso y contradictorio; por lo tanto ningún análisis es posible. La opción que prefiero consiste en concluir que los contraejemplos Gettier revelan el análisis correcto de conocimiento y, de esta manera, si mis argumentos son válidos, una forma muy radical de escepticismo es correcta. Esta conclusión implica que la mayoría de las afirmaciones de conocimiento que hacemos en la vida ordinaria son simplemente incorrectas. La opción que he preferido se hace más plausible si recordamos que una creencia o proposición no es menos valiosa meramente porque no podamos aplicarle el feliz término de 'conocimiento' a ella." pp. 512-513. Es importante observar que las conclusiones de Kirkham están motivadas por los problemas que generan los contraejemplos Gettier. Con lo cual lo que Kirkham señala es que no es posible cumplir con un requisito de *justificación* lo suficientemente exigente como para que una creencia sea "conocimiento".

42. A menos que sostengamos que este tipo de enunciados constituyen "certezas no cognoscitivas", o algo similar, al estilo de Wittgenstein. Esto no es lo que dice 2), evidentemente (en mi cuadro intenté recoger tal idea en la opción 3.1.2.). No haré comentarios sobre esta estrategia.

3.1). Consideraré, en cambio, 3.2.): la estrategia en este caso consiste en ofrecer una definición de conocimiento diferente en algún punto de la tradicional, con la esperanza de que los argumentos del escéptico no se le apliquen. En este caso, el diagnóstico sobre cuál es el error que comete el escéptico será *indirecto* (a diferencia de lo que ocurriría en 3.1.), pues debe partir de la consideración de una nueva definición de conocimiento y *desde allí* acusar al escepticismo de disponer de una concepción gnoseológica equivocada.

En especial, en los próximos capítulos me concentraré en el punto 3.2.3.). (El confiabilismo, como veremos en detalle desde el Capítulo tres, se ocupa básicamente de recusar la concepción tradicional de justificación<sup>43</sup>). Pero antes de ello quisiera plantear algunas observaciones respecto de la condición de verdad.

Respecto de 3.2.1.) creo que pueden efectuarse observaciones similares a las expresadas a propósito de 2). Este análisis condena al ostracismo el uso del concepto "verdadero", y ello no se condice con la práctica lingüística real. Sin embargo, aquí tal vez sí tenga sentido sugerir que la verdad funciona como una idea regulativa. Lo único que puedo hacer entonces frente al planteo de 3.2.1.) es oponer *intuiciones alternativas* sobre el conocimiento o sobre la verdad: por ejemplo, las intuiciones sobre las que descansa 3.2.2.). (Y no sabría, en este punto, ofrecer un argumento diferente para optar entre una u otra estrategia, entre 3.2.1) y 3.2.2.)).

En 3.2.2) creo que vale la pena detenernos un momento. En *Theories of Truth* <sup>44</sup>, por ejemplo, Kirkham sostiene que uno de los principales móviles para construir teorías no-realistas de la verdad es la suposición de que este tipo de teorías proporciona una respuesta al escéptico, puesto que en ellas se define "verdad" de modo tal que resulta fácil tener creencias justificadas como verdaderas. Esta estrategia según Kirkham no es sino un engaño intelectual: ningún problema puede resolverse simulando que nuestros intereses son diferentes de lo que realmente son, y, evidentemente, para Kirkham nuestro interés acerca de la posibilidad del conocimiento incluye gran interés en que

---

<sup>43</sup>. Es bastante habitual (y no solamente entre los confiabilistas) encontrar la idea de que el escepticismo refiere sobre todo al problema de la *justificación*, no del *conocimiento*. Véase por ejemplo L. Bonjour: las versiones más profundas y problemáticas del escepticismo parecen primordialmente concernir a la justificación más que al conocimiento, en Dancy y Sosa (1994), pp.135-136.

<sup>44</sup>. Kirkham (1992). Este, dicho sea de paso, es el único libro sistemático sobre teorías de la verdad que se ha publicado recientemente.

*verdadero* se entienda en un sentido inequívocamente no epistémico. Sobre esto quisiera aclarar dos cosas:

1) El rechazo de Kirkham hacia este tipo de solución semánticamente antirrealista, me parece, está en parte motivado por el temor de que el antirrealismo semántico implique una posición idealista en terreno metafísico. Si las conclusiones del Capítulo uno son correctas, no hay ninguna necesidad de que ello suceda.

2) Habría que tener cierto cuidado en afirmar rotundamente que una teoría semántica antirrealista (*cualquier* teoría de este tipo) soluciona (o tal vez "solucionaría, si nuestros intereses fueran diferentes de lo que realmente son") el problema del escepticismo. Intentaré mostrar aquí que no *toda* teoría no realista de la verdad tiene éxito. Daré un ejemplo de ello:

Hay todo un rubro de teorías no realistas de la verdad llamadas usualmente "teorías consensuales", o teorías de la verdad como justificación racional idealizada. La idea, expresada muy brevemente, es que la verdad de una proposición (oración, creencia, o el portador de verdad que se considere más adecuado) consiste en sus condiciones de justificación tal como éstas sean acordadas por el consenso de una comunidad ideal en condiciones óptimas de investigación. Tal concepción tiene sus orígenes en Charles S. Peirce, y ha sido adoptada, con desigual éxito, por tradiciones filosóficas muy disímiles. Putnam acude a ella cuando necesita una concepción sobre la verdad apropiada para su "realismo interno"<sup>45</sup>: "...la verdad es una *idealización* de la aceptabilidad racional. Hablamos como si hubiera tales cosas como condiciones epistemológicamente ideales, y llamamos "verdadero" a un enunciado que estaría justificado bajo tales condiciones. Las condiciones epistemológicamente ideales son como las superficies sin rozamiento: en realidad no podemos obtener condiciones epistemológicamente ideales, ni siquiera tener la certeza de que nos hemos aproximado a ellas. Pero tampoco podemos conseguir superficies sin rozamiento, y aún así decimos que las superficies sin rozamiento tienen "valor efectivo" gracias a que podemos acercarnos a ellos con un grado de aproximación bastante alto."<sup>46</sup>. Y, desde una tradición por completo diferente, Karl-Otto Apel, por ejemplo, se basa explícitamente en una idea semejante para ofrecer su propia formulación de verdad para las ciencias

---

45. En Putnam (1981).

46. Putnam (1981), p.65.

de la naturaleza: la verdad consiste en el consenso de una comunidad científica ideal e ilimitada sobre todos los criterios de verdad disponibles <sup>47</sup>(dicho sea de paso, la formulación de Apel es significativamente más rica que la noción más indeterminada de "aceptabilidad racional idealizada", pues hay en aquella una alusión expresa a (i) el papel del consenso, (ii) el papel de los criterios de verdad, como instancias claramente diferentes de lo que la verdad es).

Evidentemente, la anterior ES una concepción epistémica de la verdad, pues se está diciendo que lo que la verdad de una proposición sea depende al menos en parte<sup>48</sup> del sujeto que considera/ evalúa su verdad. Pero nótese también que, así formulada, la verdad es una noción límite, como se observa muy claramente en la cita de Putnam. No es previsible que las condiciones de investigación de alguna comunidad futura real sean reconocidas efectivamente como condiciones *óptimas*. Menos aún que tal reconocimiento alcance a nuestra situación actual.

Ahora bien, una perspectiva correspondentista tradicional, realista, sobre la verdad nos exige pensar que todo lo que hoy tomamos por verdadero podría ser, de hecho, falso; de este modo queda garantizado que la verdad es independiente de nuestro conocimiento de ella. Resulta, entonces, que podríamos estar completamente equivocados acerca del mundo, podríamos haber fallado sistemáticamente a la hora de identificar los hechos que se corresponden con nuestras estructuras semánticas. En fin, todo podría ser un sueño; podríamos ser "cerebros en una cubeta". El antirrealismo semántico en cambio pareciera que nos permite eliminar tales fantasías, puesto que en este caso la verdad es por definición *dependiente* de condiciones epistémicas. Sin embargo, la perspectiva antirrealista que dimos más arriba nos habilita para imaginar una situación bastante extraña. ¿Qué pasa si en condiciones óptimas de investigación se concluye -consenso mediante, por supuesto- que somos cerebros en una cubeta?. No tenemos razones para excluir tal posibilidad (sin cambiar, claro está, el tipo de definición de verdad que hemos dado<sup>49</sup>). La duda

---

47. En Apel (1991).

48. Por supuesto, habría que aclarar este punto. De las definiciones dadas no se sigue, sin más, que la verdad dependa *totalmente* del sujeto; por ejemplo, ello no se sigue si se sostiene simultáneamente que el consenso resultante es función de que el mundo sea de cierto modo más bien que de otro.

49. Por ejemplo, podríamos decir que es una exigencia de la teoría considerar que la mayoría de nuestras afirmaciones actuales será mantenida en condiciones ideales de justificación. Si incorporamos esta restricción, independientemente del modo en que

escéptica no puede surgir en el marco de la comunidad ideal (a lo sumo lo que podemos decir que ocurre es que el individuo *sabe* que es un cerebro en una cubeta), pero reaparece bajo otro ropaje: todo lo que *hoy* consideramos verdadero podría ser, otra vez, *falso* en condiciones ideales de investigación; no hay garantías de que la mayoría de nuestras afirmaciones de hoy, ni siquiera una pequeña porción de ellas, pueda seguir manteniéndose bajo condiciones ideales. Con lo cual este tipo de noción antirrealista de verdad es tan ineficiente para combatir dudas escépticas *actuales* como lo es la mejor teoría realista de que podamos disponer.

En lo que sigue me concentraré en el confiabilismo, esto es, en una posición que pretende inscribirse en el marco de una reformulación de la condición de justificación. Sin embargo, intentaré mostrar a lo largo de la Tesis que una solución de este tipo, enmarcada en 3.2.3.), no puede desentenderse de consideraciones sobre el problema de la verdad: no es en realidad independiente de la adopción de 3.2.1.) o 3.2.2.). Pero primero haré una exposición general acerca de las tesis confiabilistas.

---

decidamos justificarla, es indudable que, en un sentido importante, estamos *cambiando* de noción de verdad respecto de la que habíamos dado más arriba.

## CAPITULO TRES:

### TEORIAS CAUSALES DEL CONOCIMIENTO Y CONFIABILISMO.

Muchas veces se ha supuesto que, puesto que la duda escéptica surge como consecuencia de falencias en el modo de justificación de nuestras creencias/proposiciones (cae en el escepticismo quien tiene estándares de justificación muy elevados), una respuesta al escepticismo debe contemplar alguna reformulación en el modo de entender la *condición de justificación*. Ese es típicamente el caso de las respuestas que suelen ofrecerse en el marco de las llamadas "teorías causales del conocimiento". En este capítulo no me ocuparé aún de respuestas confiabilistas al escepticismo, sino de presentar en qué consisten exactamente este tipo de teorías del conocimiento. Ya he ofrecido una caracterización muy general en la Introducción; aquí abordaré el tema con un poco más de profundidad.

La concepción confiabilista del conocimiento ha sido introducida en la filosofía contemporánea de un modo paradigmático por Alvin Goldman<sup>50</sup>. En 1967 Goldman publica "A Causal Theory of Knowing", que, como veremos, en algún sentido es un antecedente directo del confiabilismo. En primer lugar, entonces, explicaré muy brevemente cuáles son los principales rasgos de un análisis del conocimiento que apele a cadenas causales. Luego trataré de reconstruir algunos aspectos de la idea un poco más sofisticada de "confiabilismo".

#### 1. Una teoría causal del conocimiento

La "teoría causal del conocimiento" en su versión contemporánea surge como un intento por dar una respuesta a los contraejemplos del tipo Gettier<sup>51</sup>. En los casos Gettier, según el diagnóstico de Goldman, la razón por la cual *S no conoce que p*, consiste en que no hay conexión causal entre la verdad de *p* y la creencia que *S* tiene de *p*. En este análisis "causal", a diferencia del análisis

---

50. Para comentarios sobre otros exponentes del confiabilismo véase Armstrong (1973), p.159, o Goldman, en Dancy y Sosa (1994), pp. 433-436.

51. Véase la Introducción de esta tesis.

tradicional, todo lo que requiere para que haya conocimiento es que exista alguna cadena causal apropiada entre cierto hecho y la creencia de ese hecho por un sujeto:

\* *S* conoce que *p* si y sólo si  
el hecho de que *p* está conectado causalmente de un modo "apropiado"  
con la creencia de *S* en *p*.

\* Los "Procesos causales apropiados" incluyen (por lo menos) los siguientes:

- (1) Percepción
- (2) Memoria
- (3) una cadena causal reconstruida correctamente mediante inferencias,  
cada una de las cuales debe estar garantizada.
- (4) combinaciones de (1), (2) y (3).

(Goldman (1967), p.369-370).

Veamos algún ejemplo:

A cuenta a B, en *t*, que en cierta región tuvo lugar una erupción volcánica hace algunos cientos de años. (He modificado ligeramente el ejemplo para poder ilustrar varios tipos de procesos causales a la vez). ¿Puede decirse que B *conoce* lo que A acaba de contarle? La respuesta es que para que ello ocurra debe existir una cadena causal que conecte el hecho en cuestión (esto es, la erupción del volcán, que resumiremos por "*p*") con la creencia de B. Podemos intentar una reconstrucción de dicha cadena causal en estos términos. *p* ha causado *q*, esto es, ha causado la presencia de lava solidificada en las inmediaciones de la región en la cual ocurrió la erupción. *q* causa, a través de la percepción visual, que A *crea* en *q*. A cree a su vez en la verdad de *r*, una proposición auxiliar que describe el modo en que la lava se produce y solidifica. La creencia de A en *q* y en *r* causan que infiera (no necesariamente de un modo consciente) que *p*, y, por lo tanto, causan su creencia en *p*. B, por su parte, a través de la percepción auditiva cree que A ha dicho que *p*; esta última creencia ha sido causada por la emisión de voz de A. B cuenta además con creencias previas que refieren a la sinceridad de A y otras cuestiones relacionadas (por ejemplo, que se encuentran ambos en ese momento en un contexto lingüístico en el cual las emisiones de oraciones deben entenderse con su significado literal, etc.). Estas últimas creencias, junto con la creencia de B de que A ha dicho que *p*, causan la creencia de B en *p*. Si

toda esta reconstrucción es correcta, entonces la creencia de B en  $p$  constituye, en efecto, conocimiento. Supongamos en cambio que

(i) A ha inventado  $p$  sin tener bases para ello (por ejemplo, tan sólo quiso aparentar conocimientos arqueológicos delante de B), pero casualmente, un volcán *realmente* entró en erupción en la región un par de siglos atrás; A, sin embargo, desconoce esta circunstancia.

(ii) La "lava" encontrada por A consiste en realidad en réplicas exactas en papel maché, dejadas por un equipo de filmación. Sin embargo, una vez más, *hubo* realmente una erupción volcánica en la región, aunque sus verdaderos rastros no han sido hallados por A.

(i) y (ii) constituyen típicos ejemplos del estilo Gettier. La razón por la cual ni en (i) ni en (ii) hablaríamos de conocimiento consiste en que en ninguno de estos dos casos existe una cadena causal entre  $p$  y la creencia de Pedro.

Esta construcción se diferencia sobre todo de una concepción tradicional en que, en general, según el análisis tripartito del conocimiento el sujeto cognoscente debe ser capaz de *justificar* o mostrar que está en posesión de cierta *evidencia* sobre la verdad de una proposición para poder afirmar que la *conoce*. Este requisito, en cambio, no se pide en el marco de la teoría causal: "Por lo menos en una interpretación popular del análisis tradicional, el que conoce debe poder justificar o dar evidencia de cualquier proposición que conozca. Para que  $S$  conozca que  $p$  en  $t$ ,  $S$  debe ser capaz, en  $t$ , de enunciar su justificación para creer en  $p$ , o sus bases para  $p$ . Mi análisis no pide tal requerimiento, y la ausencia del mismo me permite dar cuenta de casos de conocimiento que serían erróneamente excluidos por el análisis tradicional" ((1967), p.370).

La ausencia del requisito de justificación según Goldman supone una ventaja frente al análisis tradicional, puesto que nos permite dar cuenta del hecho de que conocemos muchas cosas que han sido aprendidas de un modo que ya no recordamos. Sé que San Martín nació en Yapeyú; no recuerdo a esta altura de mi vida el modo en que adquirí esta creencia, pero ello no tiene por qué quitarle el status de conocimiento. Puedo suponer cómo sería la cadena causal apropiada que conecta el nacimiento de San Martín en Yapeyú con mi creencia; por ejemplo, puedo suponer que lo leí en un (buen) libro de historia de escuela primaria, y que mi conocimiento original fue preservado hasta ahora por el proceso causal de la memoria, etcétera. El conjunto de proposiciones conocidas se reduciría drásticamente si a casos como éste le

negáramos el *status* de "conocimiento" ((1967), p.370). Este tipo de análisis creo que es especialmente útil, si bien Goldman no lo dice, para reivindicar la intuición de que muchas de nuestras creencias de "sentido común" *constituyen* conocimiento, y no son, por ejemplo, certezas no cognoscitivas. (Puedo imaginar cómo sería la cadena causal que conecta el hecho de haber nacido yo en Buenos Aires con mi respectiva creencia; podría inclusive intentar una reconstrucción de la hipotética cadena causal que conecta el hecho de que mi nombre es "Eleonora" -a través del planteo de una situación de bautismo inicial, por ejemplo- con mi creencia de ello).<sup>52</sup>.

## 2. Las tesis del confiabilismo

En "What is Justified Belief?"<sup>53</sup> Goldman recupera la idea de que la justificación es necesaria para el conocimiento, pero diferencia su concepción de las que llama descripciones "cartesianas" de la justificación; estas últimas serían las únicas que Goldman habría tenido en cuenta en sus artículos anteriores, y ello explicaría su anterior rechazo, ya comentado, de la definición tradicional de conocimiento como creencia verdadera *justificada*. La idea de un abordaje "causal" del problema del conocimiento, entonces, ha decantado aquí en lo que se conocerá como "confiabilismo". Es decir que el confiabilismo tal como lo presenta Goldman aquí puede entenderse como una *especificación* de la teoría causal del conocimiento<sup>54</sup>; el confiabilismo consiste en la explicación del *tipo de condición de justificación* que *constituye*, después de todo, un requisito necesario para conocer. No cualquier proceso causal confiere *justificación* a una creencia, sino sólo aquellos procesos que son *confiables* ((1979), p.9). Señalaré los puntos más importantes de esta concepción (no seguiré en todo momento el orden de exposición del artículo citado):

---

<sup>52</sup>. La teoría causal del conocimiento ha recibido toda clase de críticas, que aquí no comentaré. Véanse algunas de ellas por ejemplo en Armstrong (1973), pp. 158-159; o Dancy (1985), pp. 50-51. En Dretske (1981) se presentan algunas soluciones a las críticas más usuales al trabajo de Goldman (1967).

<sup>53</sup>. Goldman (1979).

<sup>54</sup>. Creo que este abordaje del confiabilismo (confiabilismo como especificación de una teoría causal del conocimiento) es especialmente útil para entender lo que hace Dretske en (1981). Véase el Capítulo cuatro.

1) Goldman presenta su trabajo como un análisis del concepto de *justificación*. Pero, como era previsible dados los comentarios anteriores, Goldman defiende una actitud *externalista* respecto de la justificación: para que una creencia esté justificada no es necesario que el sujeto que tiene la creencia *sepa* que está justificada, o pueda ser capaz de explicar *en qué consiste* su justificación. El punto es importante, pues podemos encontrar diferentes sentidos de "externalismo": hay concepciones externalistas respecto de diferentes ámbitos, y no siempre es claro cuáles son las notas distintivas que se intentan defender. Optaré por transcribir el siguiente párrafo:

"Unas pocas palabras introductorias sobre mi *explicandum* son apropiadas en este punto. A menudo se asume que toda vez que un individuo tiene una creencia justificada, conoce que está justificada y conoce cuál es la justificación. Se asume además que la persona puede indicar o explicar cuál es su justificación. En esta perspectiva, una justificación es un argumento, defensa, o conjunto de razones que pueden ser ofrecidos en apoyo de una creencia. Por lo tanto, se estudia la naturaleza de la creencia justificada a partir de considerar lo que la persona diría si le pidieran que defienda, o justifique, su creencia. Aquí no haré ninguna de esta clase de supuestos. Dejaré como una cuestión abierta si, cuando una creencia está justificada, el sujeto que cree *conoce* que está justificada. También dejaré como una cuestión abierta si, cuando una creencia está justificada, el que cree puede señalar o dar alguna justificación para ella. Ni siquiera supongo que cuando una creencia está justificada existe algo que "posee" el que cree y que pueda ser llamado una "justificación". Supongo en cambio que una creencia justificada adquiere su status de justificada a partir de algún proceso o propiedad que la hacen justificada. En suma, debe haber algunos procesos o propiedades que confieran justificación (*some justification-conferring processes or properties*). Pero ello no implica que deba haber un argumento, o razón, o ninguna otra cosa, "poseído" conjuntamente con la creencia por el que cree" ((1979), p. 2).

En fin, el externalismo parece suponer un cambio radical en el modo de entender la gnoseología. Pero, como veremos más adelante, algunas concepciones externalistas resultan muy difíciles de sostener consistentemente con el resto de las tesis de una teoría gnoseológica plausible. En el último capítulo trataré de aclarar cuál es el sentido de "externalismo" que yo considero interesante, y por qué valdría la pena defender una concepción tal de la gnoseología.

2) Goldman cree necesario dar su definición de lo que es una "creencia justificada" en términos no epistémicos, y a través de alguna/s cláusula/s básica/s, cláusulas recursivas, y una cláusula de cierre. Este formato de definición se adecua especialmente bien a una teoría fundacionista; efectivamente, lo que Goldman admite estar construyendo aquí es un tipo *especial* de teoría fundacionista<sup>55</sup>. Es importante advertir que propone como ejemplos de expresiones no epistémicas a las siguientes: "cree que", "es verdadero", "causa", "es necesario que", "implica que", "es deducible de", y "es probable" (entendiendo en este último caso la probabilidad en un sentido no subjetivista). Por lo tanto, queda claro que "verdad" se entenderá de un modo inequívocamente realista.

3) Una explicación correcta de lo que es una creencia justificada debe aludir a *cadena causal apropiada*. Sería justamente la ausencia de requerimientos causales lo que, según Goldman, convierte en inaceptables a la mayor parte de las reconstrucciones alternativas a la suya propia.

4) Ahora bien, ¿qué clase de procesos causales podemos decir que confieren justificación? Para responder a esta pregunta Goldman sugiere pensar ejemplos de procesos de formación de creencias de los cuales obtengamos como resultado creencias que intuitivamente clasificaríamos como justificadas o injustificadas. En el segundo caso mencionaríamos seguramente procesos como razonar confusamente, o intentar adivinar al azar; ellos comparten la característica de *no ser confiables*, pues en el largo plazo tienden a producir errores la mayor parte de las veces. En cambio, la formación de creencias a través de la percepción normal, el recuerdo, el uso de formas lógicamente correctas de razonamiento, y otros, parecen tener en común el rasgo de ser *confiables*, esto es, las creencias que producen son generalmente verdaderas. El carácter de "justificada" de una creencia, entonces, es función de la confiabilidad del proceso que la ha causado, y la confiabilidad de un proceso consiste en la tendencia a producir en su mayoría creencias verdaderas antes que falsas. Así como la confiabilidad es muchas veces una cuestión de grados, así también "justificación" es un concepto

---

<sup>55</sup>. En textos posteriores, sin embargo, Goldman mismo admite que el confiabilismo es compatible tanto con posiciones fundacionistas como coherentistas: es un complemento de ambas antes que un competidor. Véase Goldman (1986); véase también "Reliabilism", en Dancy y Sosa (1994), p. 435.

gradual y comparativo: tiene sentido, por ejemplo, decir de una creencia que está mejor justificada que otras.

Goldman sugiere que permanezca indeterminado cuán confiable deba ser un proceso para brindar justificación, pues nuestra concepción pre-teórica de justificación también es vaga en este mismo sentido. Es importante señalar que, tal como ha sido definido "proceso confiable", habrá oportunidades en que dichos procesos producirán creencias erróneas: es decir, puede haber creencias justificadas... y falsas.

Habría que recordar también una distinción entre "procesos tipo" y "procesos caso". Aquí se está hablando de "procesos tipo": son ellos los que tienen propiedades estadísticas, tales como producir verdades el 80% de las veces. Goldman también deja indeterminado, aunque reconoce que es un problema que debería abordarse, cuál es el grado de generalidad adecuado para elegir tipos de proceso. Por ejemplo, podríamos elegir un grado de generalidad tan estrecho que un tipo sólo tenga un caso de instanciación. Pero entonces si la creencia es verdadera, el proceso será ciento por ciento confiable, y en caso contrario será no confiable. Ello es antiintuitivo; podríamos encontrarnos con procesos intuitivamente injustificables que en estas circunstancias resulten confiables, y viceversa.

5) Por lo dicho en 3) y 4), la teoría de Goldman constituye un "*confiabilismo histórico o genético*"; es un tipo diacrónico de fundacionismo, en el que se iluminan en primer plano los aspectos genéticos o históricos.

6) La formulación de la teoría confiabilista de la justificación, entonces, quedaría como sigue:

*Cláusula básica:* Si la creencia de  $S$  que  $p$ , en  $t$ , resulta (inmediatamente) de un proceso de formación de creencias que es *incondicionalmente confiable*, entonces la creencia de  $S$  que  $p$  en  $t$  está justificada.

*Cláusula recursiva :* Si la creencia de  $S$  que  $p$  en  $t$  resulta (inmediatamente) de un proceso de formación de creencias que es (por lo menos) *condicionalmente confiable*, y si las creencias sobre las cuales este proceso opera para producir la creencia de  $S$  que  $p$  en  $t$  están ellas mismas justificadas, entonces la creencia de  $S$  que  $p$  en  $t$  está justificada.

((1979), pp.13-14).

A lo anterior debería agregársele una cláusula de cierre estándar (del tipo "ninguna otra cosa es una creencia justificada", por ejemplo). La idea de "confiabilidad condicional" es introducida para advertir que no puede esperarse que un procedimiento produzca creencias verdaderas si es aplicado a premisas falsas.

Un poco más adelante Goldman sugiere completar la cláusula básica del siguiente modo:

"Si la creencia de  $S$  que  $p$  en  $t$  resulta de un proceso cognitivo confiable, y no hay ningún otro proceso confiable (condicional o incondicional) disponible para  $S$  tal que si hubiera sido usado por  $S$  conjuntamente con el proceso utilizado de hecho  $S$  no creería que  $p$  en  $t$ , entonces la creencia de  $S$  que  $p$  en  $t$  está justificada". ((1979), p.20).

7) En los últimos párrafos de Goldman (1967), Goldman había expresado que su análisis del conocimiento en términos causales no intenta dar procedimientos para encontrar si alguien (incluido uno mismo) conoce una determinada proposición. Indudablemente, algunas veces sabemos que las creencias de un individuo están conectadas causalmente del modo apropiado con los hechos creídos. Otras veces, en cambio, puede resultar difícil, o aún imposible en principio, averiguar si existe de hecho tal cadena causal apropiada. (Por ejemplo, ¿cómo asegurar si realmente recuerdo cierto hecho que creo recordar?) . Pero estas dificultades para hallar si un sujeto determinado conoce o no, no pueden ser consideradas según Goldman dificultades inherentes a su concepción del conocimiento. En este sentido, afirma no estar interesado en lo más mínimo en responder problemas escépticos: "Mi análisis no da respuesta al escéptico que cuestiona si yo comienzo [el análisis] a partir del contenido de mi propia experiencia y luego pruebo que conozco que hay un mundo material, un pasado, etc. No considero que sea éste uno de los objetivos de dar condiciones de verdad para " $S$  conoce que  $p$  " ((1967), p.372).

Goldman mantiene esta actitud general en sus trabajos posteriores<sup>56</sup>. Pero en el marco de la propuesta "confiabilista", surgen algunos posibles

---

<sup>56</sup>. Por ejemplo, si bien en (1986) dedica un capítulo entero para presentar el problema del escepticismo, termina dicho capítulo afirmando que el escepticismo es sólo un punto del temario que debe ocupar a una teoría del conocimiento equilibrada. Según Goldman al escepticismo se le ha dado tradicionalmente una importancia exagerada, y advierte que él intentará reparar esta desproporción, dedicándole al problema un espacio

contraejemplos de su concepción de la justificación, que resultan curiosamente afines a los planteos escépticos, y que entonces en alguna medida se ve necesitado de enfrentar. Uno de los problema es el siguiente: puesto que la teoría intenta cubrir todos los casos posibles, parece implicar que para cualquier proceso cognitivo C, si C es confiable en el mundo posible M, entonces cualquier creencia en M que resulte de C está justificada. ((1979), p.16). Pero supongamos que, objetivamente, expresar un deseo sea en nuestro mundo un proceso confiable, puesto que, aunque desconocido para nosotros, existe un demonio benigno que arregla las cosas de modo que la mayoría de este tipo de pensamientos se cumplen. Intuitivamente, sin embargo, no diríamos que es un proceso confiable. El problema tiene relación con cómo hemos de entender el externalismo de una teoría confiabilista:

Goldman termina reconociendo que lo que realmente deseamos y pedimos de una teoría es una explicación de por qué consideramos, o consideraríamos (si las examinásemos), ciertas creencias como justificadas, y otras como injustificadas. Tal explicación debe referir a nuestras *creencias* sobre la confiabilidad, y no a lo que realmente ocurre. La razón por la que consideramos una creencia como justificada es que ha sido formada por lo que *creemos* es un proceso confiable de formación de creencias. Nuestras creencias acerca de cuáles procesos cognitivos son confiables pueden ser erróneas, pero ello no afecta la adecuación de la explicación. Puesto que creemos que expresar deseos no es un proceso confiable de formación de creencias, contamos a las creencias formadas a través de tal proceso como injustificadas ((1979), p.18).

Pero, entonces, ¿dónde han quedado las aspiraciones externalistas de la teoría? ¿No es esto último inconciliable con las opiniones vertidas en 1)? En efecto, si por externalismo entendemos la idea de que la justificación de una creencia depende de la existencia de un proceso objetivo, y *no* de nuestra percepción del proceso como tal, entonces el párrafo anterior defiende una posición no externalista. Volveré sobre esta cuestión en el último capítulo<sup>57</sup>.

---

comparativamente reducido. Efectivamente, fuera de los comentarios (generales) del capítulo 2, las referencias al escepticismo en dicha obra son muy limitadas.

57. Foley (1985), por ejemplo, comenta este mismo pasaje de Goldman y opina que si la sugerencia de Goldman ha de tomarse en serio, ello supone un abandono del confiabilismo antes que un intento de reformarlo (Foley (1985), pp.193-194).

No todos los autores presentan las tesis confiabilistas de esta manera, por supuesto; en especial, no todos consideran que el confiabilismo es básicamente una teoría de la *justificación*. Armstrong, por ejemplo, lo describe como una *teoría del conocimiento* <sup>58</sup>, y otro tanto hace Dretske, como veremos en los próximos capítulos.

Las discusiones recientes sobre confiabilismo en general han versado sobre las posibilidades que tiene esta concepción de salir airoso de una serie de contraejemplos, esto es, (1) casos hipotéticos en los que intuitivamente desearíamos hablar de conocimiento, pero según el confiabilismo no debiera ser posible, o (2) casos que, a la inversa, desearíamos poder eliminar como ejemplos de conocimiento, pero que el confiabilismo sin embargo acepta. Por otro lado, también se han ofrecido análisis sobre algunos puntos complejos de la teoría, que ya he mencionado, tales como el llamado "problema de la generalidad" (cuestiones acerca de cuál es el grado de generalidad adecuado para un proceso-tipo), y el problema del espectro de confiabilidad (*the range of reliability*): ¿debe considerarse la proporción de verdades que genera un proceso en el mundo *real*, o debe incluirse una consideración acerca de las proporciones de verdad en otros mundos posibles?<sup>59</sup>. No entraré aquí en estos

---

58. Véase Armstrong (1973): una creencia constituye "conocimiento" si es "empíricamente confiable" (p.159). Véase en especial el capítulo 12, "Non-Inferential Knowledge". Allí Armstrong sugiere que lo necesario para tal confiabilidad es la existencia de una *conexión legal* entre el estado de cosas consistente en que un sujeto cree que *p*, y el estado de cosas que hace a la proposición "*p*" verdadera (p.166 y siguientes). Hay interesantes puntos de contacto entre esta idea y las que, veremos, desarrolla Dretske en (1981): para Dretske su propio análisis en términos de relaciones informacionales es más básico, pues las regularidades nómicas de las que necesita hablar Armstrong derivan su carácter modal de su naturaleza informacional. (Dretske (1981), pp. 77 a 86).

59. No me ocuparé especialmente de estas cuestiones. Tan sólo mencionaré los dos contraejemplos más influyentes que se encuentran en la bibliografía:

1) En un mundo posible en el que un demonio maligno crea experiencias visuales defectuosas, el proceso de la visión no es confiable. Sin embargo, en este mundo las experiencias visuales parecen intuitivamente estar justificadas;

2) Supóngase un individuo con reales poderes de clarividencia. Según el confiabilismo, sus creencias formadas a través de la clarividencia están justificadas (si la proporción de creencias verdaderas formadas a partir de tal proceso es suficientemente alta); sin embargo, intuitivamente las creencias del sujeto no están justificadas (tal vez inclusive él mismo las considere no justificadas). Véase "Reliabilism", Dancy y Sosa (1994), pp.433-436. Los contraejemplos mencionados, propuestos por Laurence Bonjour, junto con otros propuestos por diversos autores y comentados en Goldman (1986), se recogen y discuten entre otros lugares en Schmitt (1992), capítulo 7, "Justification and Evaluation", pp. 175-198, -Schmitt comparte las intuiciones confiabilistas de Goldman- y en Haack (1993), capítulo 7, "The Evidence against Reliabilism", pp.139-157, en este último caso en un contexto hostil al confiabilismo y a Goldman en particular. El problema de la generalidad

asuntos pues, según entiendo, me llevarían demasiado lejos de los temas que me deseo abordar en esta Tesis, y que se relacionan sobre todo con *los presupuestos que el confiabilismo necesita para poder ofrecer una respuesta satisfactoria frente al escepticismo* .

### 3. Algunas observaciones

Como ya hemos aclarado, Goldman no se ocupa especialmente de brindar alguna respuesta al escéptico. Pero esta rápida exposición de un ejemplo de teoría causal del conocimiento y, en particular, de una teoría causal confiabilista, nos ha servido para describir el marco conceptual que presuponen las estrategias contra el escepticismo que examinaremos a continuación. Antes de ello considero importante llamar la atención sobre los siguientes puntos del confiabilismo, sobre los que no he encontrado ninguna referencia explícita en la bibliografía .

- En primer lugar, quisiera disipar la posible sensación de que la formulación del confiabilismo encierra algún tipo de circularidad. Tal como hemos visto, el confiabilismo entiende que para que una creencia de un sujeto constituya conocimiento es necesario que haya sido adquirida a través de un procedimiento adecuado, esto es, a través de un proceso "confiable". Y por "proceso confiable" se entiende "proceso que genera la mayor parte de las veces creencias verdaderas". Esta formulación sería en efecto circular si la confiabilidad de un proceso de adquisición de creencias fuera el requisito que nos permite justificar nuestra consideración de cierta creencia como *verdadera* . Pero ello no es el caso. La idea, en cambio, es que la confiabilidad del proceso es lo que permite que una *creencia verdadera* pueda ser justificada como "conocimiento".

- De esto último se sigue que el confiabilismo cuenta con un gran presupuesto: *tenemos creencias verdaderas* . Y, como ya he expresado en la Introducción, si hemos de dar alguna respuesta al escéptico, debemos poder *reconocer* algunas de estas creencias verdaderas. Ahora bien, para evitar la

---

es tratado extensamente en Schmitt (1992), capítulo 6, "Reliability, Relevant Processes and Metaproceses".

circularidad, como acabamos de ver, tal reconocimiento debe ser *independiente* del proceso que generó la creencia; si hay algún criterio de verdad, éste no puede consistir en la existencia del mismo proceso confiable que ha generado la creencia en cuestión<sup>60</sup>. Antes bien, el confiabilista necesita *primero* reconocer creencias como "verdaderas", para poder luego *identificar* ciertos procesos como "confiables". El punto es que esta cuestión en general es totalmente soslayada, por lo cual no tenemos demasiados indicios de cuáles criterios supone el confiabilista útiles para reconocer la verdad de proposiciones y creencias<sup>61</sup>.

- Por otra parte, para identificar un proceso P como confiable deberíamos poder asegurar que, si P ha generado hasta hoy nueve de cada diez veces creencias verdaderas, seguirá arrojando esa proporción de verdad en el futuro. Por lo tanto, no solamente es necesario poder evaluar la verdad de una creencia con independencia del proceso que la generó; también es de vital importancia poder recurrir a mecanismos inductivos. Se hace necesario, entonces, brindar alguna justificación de la inducción, o al menos explicar por qué su uso no requeriría justificaciones de ningún tipo, y por qué, en consecuencia, el escéptico respecto de la inducción está equivocado. De todo lo anterior resulta que el confiabilismo reposa sobre un segundo gran supuesto: que podemos hacer uso de la inducción de un modo no problemático. De todas maneras, conviene aclarar con cierto cuidado *en qué sentido* considero que el confiabilista necesita de la inducción. Si la observación del párrafo anterior es correcta, el confiabilista no utiliza la inducción para afirmar: "puesto que en su mayoría las creencias generadas a través de un proceso P fueron verdaderas, la próxima creencia también lo será". La posición

---

<sup>60</sup>. Caso contrario caeríamos aquí en el círculo vicioso conocido a veces como el "problema del criterio": esto es, para identificar creencias verdaderas necesitamos un método, o criterio de verdad, pero para poder asegurar que tal criterio es en efecto bueno, debemos saber, previamente, cuáles creencias son verdaderas. Véase por ejemplo Chisholm (1977), último capítulo. (No siempre es claro en Chisholm de qué criterio se está hablando; en (1977) parece referir a un "criterio de conocimiento", aunque en otras oportunidades parece aludir, en cambio, a "criterio de verdad". Véase por ejemplo Chisholm (1973)).

<sup>61</sup>. O tal vez debiéramos decir, de un modo más general, que no tenemos indicios de cuáles son las creencias verdaderas con que contamos (suponiendo que tal vez sea posible reconocer creencias verdaderas sin contar explícitamente con algún criterio de verdad). Artículos más recientes de Dretske parecen apuntar a un criterio pragmático de verdad, algo así como "las creencias verdaderas son aquellas que nos han permitido sobrevivir como especie" (Dretske (1989)); pueden sospecharse además otros criterios posiblemente compatibles con la posición confiabilista. Véase el Capítulo cinco de esta Tesis.

confiabilista sostiene, en cambio, que "puesto que en su mayoría las creencias generadas a través del proceso P fueron verdaderas, la próxima estará también justificada".

- Quisiera esbozar también una primera evaluación de los que considero son algunos de los aspectos positivos del confiabilismo. Uno de los principales méritos del confiabilismo, según entiendo, consiste en ofrecer una caracterización del conocimiento, y en especial de lo que constituye el requisito de *justificación* del conocimiento, en términos tales que permiten reconocer como justificada aún a algunas que otras creencias falsas. Ello recoge la siguiente intuición: puesto que el conocimiento es falible, puedo equivocarme y tomar como verdadera una creencia falsa. Una buena *justificación* de que cierta creencia es conocimiento no puede, entonces, requerir certeza; no puede pretender que tengamos garantías totales de que la creencia en cuestión *no es* de hecho falsa. El escéptico que desafía a esta reconstrucción del conocimiento, entonces, no es el que dice "la próxima creencia generada por P pueden ser falsa", sino: "de la próxima serie de creencias, TODAS ellas, sistemáticamente, pueden ser falsas"<sup>62</sup>. Es decir, los procesos que se *creen* confiables pueden generar sistemáticamente falsedades; o no hay procesos confiables, o no hay forma de saber si los hay. Será importante tener en cuenta estas precisiones a la hora de considerar una respuesta al escepticismo en el marco del confiabilismo.

- Otro mérito del confiabilismo consiste, creo, en su compromiso con el externalismo. Trataré más adelante esta cuestión (en el último capítulo).

- Por último, otro aspecto positivo del confiabilismo radica en su reivindicación implícita del método científico para el ámbito de la gnoseología. La correcta evaluación de cuándo un proceso es confiable es, como vimos, un asunto delicado. En Goldman (1986), por ejemplo, encontramos una apuesta en favor de la psicología cognitiva: la investigación psicológica es la que puede decirnos cuáles procesos son confiables. Ello, por supuesto, puede ser discutible (aquí no entraré en este tipo de problemas) pero de todas maneras supone un intento por construir una gnoseología naturalizada, y en tanto tal, vale la pena tenerlo en cuenta<sup>63</sup>.

---

62. Sigo aquí la idea de que los argumentos escépticos proceden *ad hominem*, es decir que son capaces de ir adaptándose en cada caso al adversario con que cuentan (véase por ejemplo E.de Olaso (1994), p.160).

63. Véase la crítica de Susan Haack a este intento, en (1993), capítulo 7.

- Finalmente, quisiera reparar una vez más en la conexión entre confiabilismo y verdad. Como ya he expresado, no es ésta una relación que los autores se ocupen de profundizar. ¿Hay, tal vez, alguna concepción de la verdad (de lo que la verdad SEA, con independencia de la existencia de criterios para *evaluar* la verdad) que podamos reconocer como especialmente compatible con las tesis del confiabilismo? Hemos observado que Goldman supone que la verdad es no epistémica. En efecto, una teoría causal del conocimiento resulta por demás adecuada para combinarse con una noción realista de la verdad. Creo que es interesante, por ejemplo, advertir un paralelismo entre teorías causales del conocimiento y, en semántica, teorías causales de la referencia y el significado<sup>64</sup>. Estas últimas, como surge de algunas observaciones vertidas en el capítulo uno de esta Tesis, en general sostienen que es algún hecho identificable del mundo lo que permite en cada caso dar cuenta de nuestras nociones semánticas, (entre otras, de la verdad); así es que se comprometen con una interpretación "fuerte" de la idea de correspondencia y, en consecuencia, con una concepción realista de la verdad. Una teoría causal del *conocimiento*, por su parte, centra su interés en la conexión causal entre hechos del mundo y creencias de ciertos sujetos acerca de esos mismos hechos. Por lo tanto, podemos suponer, comparte este mismo "espíritu" realista respecto de la correspondencia y, si se ocupara de cuestiones semánticas (la verdad lo es) adoptaría soluciones similares a las desarrolladas por los teóricos causales de la referencia.<sup>65</sup> Ahora bien, ¿cómo procede el confiabilismo para evitar el potencial escéptico de una noción realista de la verdad? En especial, la consecuencia escéptica que aquí nos importa es la afirmación de que TODAS nuestras creencias podrían ser falsas (a esta tesis se la conoce usualmente bajo el nombre de "la posibilidad de error masivo"). Como era de esperarse, el confiabilismo simplemente elude el tema. Me ocuparé de esto en el último capítulo.

---

<sup>64</sup>. Dretske, por ejemplo, es un autor que, en (1981), ha procurado diseñar una teoría unitaria para ambos tipos de cuestiones. Como veremos, para ello aplica ideas de teoría matemática de la información, a la gnoseología y a la semántica.

<sup>65</sup>. Para un esbozo de un paralelismo entre teorías externalistas en epistemología y en otras disciplinas (por ejemplo, en filosofía de la mente), véase el artículo sobre "externalism/internalism", en Dancy y Sosa (1994), pp. 132-136.

## CAPITULO CUATRO:

### EL CONOCIMIENTO EN EL MARCO DE LA TEORIA DE LA COMUNICACION.

En este capítulo deseo analizar un intento confiabilista por argumentar en contra del escéptico respecto de la posibilidad de conocimiento empírico. Me basaré en el libro de Dretske *Knowledge and The Flow of Information* (1981), pues constituye, a mi juicio, una de las posiciones más interesantes. Dretske elabora un tipo especial de teoría causal, a saber, una "teoría informacional" del conocimiento<sup>66</sup>. Para ello extrae algunas ideas fundamentales de la teoría matemática de la comunicación.

#### 1. Una teoría informacional del conocimiento.

Intuitivamente, cualquier estado de cosas genera cierta información; la información que lleva una cierta señal, por ejemplo, es lo que la señal nos puede decir *con verdad* sobre un estado de cosas (Dretske (1981), p.50). Información aquí actúa como una suerte de primitivo, que quedará caracterizado a medida que se vaya explicitando el funcionamiento de la teoría<sup>67</sup>. La información, por otra parte, es algo cuantificable: la "teoría de la comunicación", en la que Dretske se basa, propone medir cantidades de información en *bits*, esto es, en unidades binarias. Así, por ejemplo, el

---

<sup>66</sup>. En los primeros tres capítulos, Dretske se encarga de explicar los puntos de la teoría matemática de la comunicación que considera útiles para sus fines; en los capítulos 4 a 6 se ocupa específicamente de cuestiones gnoseológicas; y en los últimos tres capítulos aborda problemas de filosofía de la mente y de semántica. Sin embargo, los temas están cerradamente entrelazados, de modo que las consideraciones que vertiré en esta Tesis recogen tratamientos que hace Dretske del tema a lo largo de los nueve capítulos. No respetaré el orden de exposición de Dretske, ni el modo en que Dretske mismo organiza y divide los problemas.

Todas las citas pertenecen a la edición castellana de la obra.

<sup>67</sup>. Hay caracterizaciones de lo que la información es en el texto de Dretske, pero son todas ellas más o menos circulares; por este motivo sostendré que "información" es un primitivo de la teoría (he aquí un buen ejemplo de las "definiciones" a las que me refiero: "la información es algo que, dado el destinatario adecuado, puede proporcionar conocimiento", p.53; como se verá, Dretske usará luego el concepto de "información" para definir "conocimiento").

procedimiento de reducir ocho posibilidades a una, genera tres bits de información. Esta posibilidad de cuantificación, aplicada a un análisis del conocimiento, da como resultado una teoría gnoseológica con ciertas características originales<sup>68</sup>.

Dretske define "conocimiento" (perceptivo) del siguiente modo: "K sabe que s es F = la creencia de K de que s es F está causada, o causalmente mantenida, por la información de que s es F" (p.90). Para comprender plenamente el significado de esta ecuación es necesario indagar en algunos puntos un poco más técnicos de la construcción de Dretske. Es importante, por ejemplo, saber qué quiere decir exactamente "llevar una información". Y para ello, a su vez, será preciso explicar primero algunos de los conceptos que maneja la teoría matemática de la comunicación, y analizar con base en ellos qué quiere decir que algo sea el "contenido" (semántico) de una señal. En I), entonces, expondré algunas fórmulas generales de la teoría de la comunicación que permiten manejar datos cuantitativos; en II) mostraré algunas estrategias de Dretske para hablar de cuestiones cualitativas (el problema del contenido de la información); en III) usaré ideas de I) y II) para explicar qué significa que una señal lleva una determinada información.

#### I)

En primer lugar, es importante distinguir entre información generada (por una fuente de información) e información recibida. Luego, dentro de ambas categorías, debiera distinguirse entre información *media* (generada o recibida), e información asociada (generada o recibida) a un estado de cosas en particular o a la ocurrencia de un evento.

Como veremos más adelante, para un análisis gnoseológico del tipo que nos interesa aquí, no es de especial interés el concepto de "información media". Más aún, habrá que encontrar el modo de enfatizar los aspectos *cualitativos* involucrados en los procesos informacionales: no basta con saber que un mensaje determinado, por ejemplo, transmite 3 bits de información y no 4 o 0,25; por lo general estamos interesados, además, en el *contenido* (semántico)

---

<sup>68</sup>. Para una introducción interesante a la teoría de la información véase por ejemplo Satoshi Watanabe, "La información", en Bar-Hillel (1978), pp.198-224. Para otros intentos de un uso no estándar de la teoría de la información ténganse en cuenta algunas construcciones de Hintikka (1973), por ejemplo el capítulo 10, "Información, deducción y lo *a priori*", pp. 256-277, donde se muestra un sentido en el que la deducción puede aumentar nuestra información.

del mensaje, y una teoría del conocimiento satisfactoria debe poder dar cuenta de este interés. Lo que intenta hacer Dretske, justamente, es transformar una teoría puramente cuantitativa, como es la teoría matemática de la comunicación, en un instrumento que permita dar cuenta de elementos cualitativos, por ejemplo, de la noción de contenido semántico de una señal.

La teoría matemática de la comunicación dispone de fórmulas para calcular:

- la cantidad *media* de información asociada con una cierta fuente generadora de información,  $I(s)$ . En ejemplos muy sencillos, la fuente puede ser el procedimiento de arrojar una moneda al aire, lo que permite dos posibles resultados, o el procedimiento de extraer una carta de un mazo de naipes.

- la cantidad de información asociada con un resultado particular, o cierto estado de cosas,  $I(s_i)$ . Por ejemplo, el hecho de que cierto tiro haya salido "cara".

- la cantidad *media* de información recibida por un receptor,  $I(r)$ ;

- la cantidad de información asociada con el mensaje determinado que se recibe,  $I(r_i)$ .

La fórmula más sencilla que podemos considerar nos permite calcular la cantidad de información generada por la reducción de un conjunto  $n$  de posibilidades (igualmente probables) a 1:

$$(1) \quad I(s) = \log n$$

donde "log" es el logaritmo en base 2, esto es, la potencia a la que hay que elevar 2 para obtener  $n$ . Por ejemplo, dada una moneda no trucada,  $n$  es igual a 2, y por lo tanto, la información media que genera la moneda es de 1 bit.

Pero, si lo que queremos es averiguar la información asociada con la ocurrencia particular de un evento (o con la existencia de un estado de cosas determinado), entonces

$$(2) \quad \begin{aligned} I(s_i) &= \log 1/p(s_i) \\ &= -\log p(s_i) \end{aligned}$$

donde  $p(s_i)$  = probabilidad de que suceda  $s_i$  en una serie de posibilidades. Debe notarse que cuanto más probable es una determinada posibilidad, tanto menos información generará su ocurrencia; a la inversa, a menor probabilidad de

ocurrencia de un evento, más información resultará asociada con dicho evento, en caso de que efectivamente suceda (se dice por lo tanto que esta fórmula mide el "valor sorpresa"). Por ejemplo, cada ocurrencia de "cara" o "ceca" a partir de una moneda no trucada tiene una probabilidad de 0.5, y  $\log 1/0.5$  es igual a  $\log 2$ , que es igual a 1 bit. Pero si, en cambio, contamos con una moneda trucada, en la cual la probabilidad de "cara" es de 0.9, entonces la información asociada con la ocurrencia de "cara" será  $\log 1/0.9$ , es decir, de sólo 0.15 bits, mientras que la información asociada con la ocurrencia de "ceca" es de  $\log 1/0.1 = 3.32$  bits.

Entonces, puesto que no todas las posibilidades tienen por qué ser igualmente probables, una manera más precisa de calcular  $I(s)$  puede consistir en lo siguiente:

$$(3) \quad I(s) = \sum p(s_j) * I(s_j).$$

Es decir, la información media asociada con una fuente es igual a la sumatoria de las probabilidades de ocurrencia de cada uno de los eventos posibles multiplicadas por la información asociada con cada una de dichas ocurrencias. En nuestro caso de la moneda trucada  $I(s)$  es igual a 0.467 bits (bastante menos que la información media que genera una moneda no trucada).

Exactamente del mismo modo se pueden construir fórmulas para  $I(r)$  y  $I(r_j)$ .

Además, puede haber equivocidad y ruido en la transmisión de información. *Equivocidad* (E) es la información generada en s que no se transmite a r. Y se entiende por *ruido* (N) la información asociada al receptor que no proviene de s. Debe poder calcularse en qué medida la información presente en la situación r refleja exactamente la información generada en la situación s -sin pérdida o distorsión de contenido-, y hasta qué punto la información en r *se debe* a la señal que partió de s. Así, si entendemos por  $I_S(r)$  una medida de la cantidad de dependencia de r respecto de s, entonces,

$$(4) \quad I_S(r) = I(r) - \text{ruido (N)}; \text{ o}$$

$$(5) \quad I_S(r) = I(s) - \text{equivocidad (E)}.$$

Estas últimas fórmulas, como veremos más abajo, son importantes, pues son las que permiten calcular la *cantidad* de información que lleva una señal: esto es, la cantidad de información que llega al receptor como resultado de cierta generación de información en la fuente. Luego nos importará atender a los aspectos cualitativos de estas fórmulas: no solamente cuánta información se ha transmitido, sino cuál, con qué contenido.

Para calcular N (ruido), (lo que es a su vez importante para (4)) hay que considerar la *probabilidad condicional de que se reciba cada una de las posibilidades (r<sub>1</sub>...r<sub>i</sub>) dado la ocurrencia en la fuente de cierta posibilidad determinada (s<sub>a</sub>)*. En primer lugar, entonces, debe calcularse la contribución de s<sub>a</sub> al promedio de ruido:

$$(6) \quad N(s_a) = - \sum P(r_i/s_a) * \log P(r_i/s_a)$$

Y, en segundo lugar, debemos sumar las contribuciones individuales y ponderarlas según su probabilidad de ocurrencia:

$$(7) \quad N = p(s_1) * N(s_1) + p(s_2) * N(s_2) + \dots + p(s_i) * N(s_i)$$

De modo similar se puede calcular la equivocidad (E):

$$(8) \quad E(r_a) = - \sum P(s_i/r_a) * \log P(s_i/r_a)$$

$$(9) \quad E = p(r_1) * E(r_1) + \dots + p(r_i) * E(r_i).$$

En este caso lo que se tiene en cuenta son las probabilidades condicionales de que, dada una cierta recepción (r<sub>a</sub>), se hayan producido en la fuente las diferentes posibilidades (s<sub>1</sub>.....s<sub>i</sub>).

Como es notorio, (8) y (9) son necesarias para (5), del mismo modo en que (6) y (7) lo eran para (4).

También es importante, según Dretske, distinguir entre *relaciones causales* entre s y r, y *relaciones informacionales*. La relación de información no debe confundirse con una relación causal *simpliciter*, sino que toda relación informacional agrega un *plus* a la mera relación de causalidad; es,

dicho de otro modo, un tipo especial de causalidad. Un individuo podría ser enfrentado a cierta experiencia perceptiva y sin embargo no recibir, por diversos motivos, ninguna información: el *dorso* de un as de espadas puede efectivamente causar una experiencia sensorial a un sujeto, pero dicha experiencia no contiene información sobre su antecedente causal<sup>69</sup>.

## II)

Puesto que Dretske necesita incorporar la noción de *contenido semántico* a la teoría de la información, en realidad está transformando dicha teoría en algo diferente. Para llevar a cabo un tratamiento exitoso de la noción de "contenido" de la información, propone, en primer lugar, un medio de distinguir entre estructuras capaces de tener contenido semántico y aquellas que no tienen tal capacidad. Para ello apela a la noción de "características intencionales" de un sistema. En segundo lugar, trata de encuadrar la idea de "estructura con contenido semántico" dentro de un análisis sobre la *codificación* de la información. Así, define a aquellas estructuras que tienen el grado de intencionalidad apropiado (lo que las identifica como estructuras con contenido semántico) como aquellas que son capaces de codificar la información de cierto *modo* particular (de modo "completamente digitalizado")<sup>70</sup>.

Veamos esto con un poco más de detalle.

Dretske define tres niveles de intencionalidad, que quedarán caracterizados, cada uno, por tríadas de enunciados:

S (una señal, suceso o estado) tiene el primer nivel de intencionalidad si  
a) Todos los F son G.

---

<sup>69</sup>. Dretske dice además que *no* toda relación informacional es síntoma de la existencia de un proceso determinista. No me ocuparé ahora de esta afirmación, a mi juicio controvertida (véase Dretske (1981), p.42).

<sup>70</sup>. El resultado de esta estrategia será una explicación, en términos de teoría de la información, de lo que son las *creencias*. Más adelante volveré sobre este tema.

- b) S tiene el contenido de que t es F
- c) S NO tiene el contenido de que t es G.

En el segundo nivel nos encontraríamos con que

- a) Es una ley natural que los F son G
- b) S tiene el contenido de que t es F
- c) S NO tiene el contenido de que t es G.

Por último, en el nivel superior de intencionalidad,

- a) Es analíticamente necesario que los F sean G
- b) S tiene el contenido de que t es F
- c) S NO tiene el contenido de que t es G.

(Dretske (1981), pp.176-177).

La idea que se quiere recoger aquí es que cuanto más discriminadora sea la estructura que lleva una información, más elevado será su nivel de intencionalidad; en el tercer nivel, finalmente, una estructura tendrá un contenido proposicional como contenido *exclusivo*. *Todo* sistema que procesa información tiene por lo menos el nivel de intencionalidad más bajo. Si una señal transmite la información de que JJ es miembro de cierto club, y hay una equivalencia extensional *casual* entre ser miembro del club en cierto momento y tener menos de 40 años, la señal *no* puede llevar, al mismo tiempo, la información de que JJ tiene menos de 40 años. La discriminación que describen los niveles segundo y tercero, en cambio, tiene lugar solamente en determinados sistemas: en los que Dretske llama "sistemas cognitivos genuinos"<sup>71</sup>. Un sistema cognitivo genuino puede tener la información de que el agua se está congelando sin tener simultáneamente la información de que ésta aumenta de tamaño (nivel 2); o puede creer que cierto número es el logaritmo en base 2 de 8 sin creer que ese número es 3; puede discriminar el

---

<sup>71</sup>. Como es evidente, Dretske habla de "intencionalidad" en un sentido bastante general. Su sugerencia es que "podemos entender la peculiar estructura intencional de nuestros estados cognitivos como síntomas de su subyacente carácter informacional. Tal vez la intencionalidad de nuestras actitudes (el modo en que tienen un *contenido* único), un rasgo que algunos filósofos consideran distintivo de lo mental, sea una manifestación de su subyacente estructura informacional" (p.81). El sistema receptor de información puede ser tanto un voltímetro, una calculadora, una rana o un ser humano (aunque no en todos los casos, por supuesto, habrá intencionalidad de tercer nivel).

significado de "Ana está sola ahora" como un significado *diferente* del de "Ana no está con su vecina ahora" (nivel 3)<sup>72</sup>.

Se podrá hablar de "contenido semántico", entonces, para referirnos a *cualquier contenido proposicional que presente el tercer nivel de intencionalidad*. Ahora bien, la pregunta es cómo dar cuenta de la propiedad de tener un tercer nivel de intencionalidad a partir del concepto de información. Para ello Dretske introduce una distinción entre información codificada de forma *analógica* vs. de forma *digital* (aclara en realidad que usará una terminología conocida pero le dará un sentido algo heterodoxo). En una codificación analógica del hecho de que t es F, la estructura lleva una información más específica, más determinada sobre t que el hecho de que sea F. Si se transmite la información de que Ana es una persona (y no un gato) a través de una foto, por ejemplo, esa información estará subsumida en muchas otras: que Ana es una mujer mayor, de raza blanca, pelo castaño, etc. Por el contrario, S lleva la información de que t es F en forma digital si y sólo si esa es la información *más específica* sobre t que lleva S; la información estará en forma *completamente* digitalizada si, además de ser la información más específica sobre t, S no lleva ninguna otra información acerca de r tal que t es F esté incluida en r es G. Ello implica que el contenido semántico se identifica con la capa informativa más externa de una señal - las capas interiores están, por definición, codificadas de modo analógico (por ejemplo, en la información digitalizada de que t es un cuadrado, la señal transmite además analógicamente que es un cuadrilátero y un paralelogramo).

Hay dos corolarios importantes de esta definición:

1) Para que una estructura lleve información con contenido semántico, tal estructura debe poder producirse como respuesta a una *variedad de situaciones informacionales diferentes*, pues si hay una única situación posible que la provoca, la señal también lleva información sobre sus medios de producción, con lo cual el contenido "t es F" estaría subsumido en algún "r es G" (= no hay codificación de modo totalmente digital sobre t) (p.192). Un ejemplo: no diríamos que la información que ha recibido la aguja de un voltímetro tiene contenido semántico (no diríamos que el voltímetro "sabe" que hay 7 voltios asociados a una cierta fuente de energía, por ejemplo) porque lo que la señal

---

<sup>72</sup>. Intuitivamente, un ser humano es un sistema cognitivo genuino; un termostato o un voltímetro no lo son. Más abajo se aclarará esta idea.

transmite en realidad son datos mucho más inmediatos; datos sobre ciertos sucesos intermedios que se producen en el interior del aparato.

2) Este último ejemplo muestra además que para que haya contenido semántico debe haber *pérdida de información*. Una estructura semántica, entonces, es una estructura *interpretativa*, pues caracteriza o ilumina un componente a expensas de los otros; un sistema puede tener diferentes modos de digitalizar la misma información (analógica) en diferentes momentos<sup>73</sup>. Un voltímetro o un tocadiscos, en cambio, transforman físicamente una cierta información (el tocadiscos, por ejemplo, en ondas sonoras) pero no la interpretan, en el sentido de que no poseen ningún tipo de sensibilidad selectiva: sus señales no llevan información con el tercer nivel de intencionalidad; no son, evidentemente, sistemas cognitivos genuinos. Del mismo modo, un sujeto puede digitalizar el contenido semántico del enunciado "Elmer ha muerto" impreso en un diario, recién cuando lo abstrae de la experiencia sensorial asociada con la visión de la página. Dretske agrega un ejemplo que creo que es sugerente e intuitivo: "El analfabetismo no es una deficiencia *perceptiva*, sino, más bien, una deficiencia en la codificación, es decir, una incapacidad para convertir la información presente en forma sensorial a una forma cognitiva, una imposibilidad de digitalizar (completamente) la información de que se dispone en la experiencia sensorial (en forma analógica) de la página impresa del periódico." ((1981), p.191).

### III)

Con los elementos con que ahora contamos podemos abordar el problema de qué quiere decir exactamente que una señal "lleve una determinada información". Considerar la información que lleva una señal, recordemos, supone tener en mente lo que ha llegado al receptor como resultado de cierta generación de información en una fuente. Es decir, debe tenerse en cuenta la medida de la independencia de la información recibida respecto de lo que ocurrió en la fuente. Para ello deberíamos considerar un modo de calcular la *cantidad* de información que una determinada señal  $r_a$  lleva sobre  $s_a$ :

$$(10) \quad I_s(r_a) = I(s_a) - E(r_a)$$

---

<sup>73</sup>. Según Dretske, "la diferencia entre ver (oír, oler, etc.) un  $t$  que es  $F$  y creer o saber que  $t$  es  $F$  es fundamentalmente la misma diferencia que existe entre una representación analógica y otra digital del hecho de que  $t$  es  $F$ ." ((1981), p.187).

$I(s_a)$  se obtiene según se mostró en (2):  $I(s_a) = \log 1/p(s_a)$

$E(r_a)$ , según (8) :  $E(r_a) = - \sum P(s_i/r_a) * \log P(s_i/r_a)$

Estas tres, por lo tanto, serán las fórmulas más importantes en vistas a nuestros objetivos.

En (5) se había presentado una fórmula análoga a (10), pero respecto del concepto de información *media*. Pero, como ya mencioné, lo que nos interesa aquí, en realidad, es la cantidad de información generada por un determinado *estado de cosas* (este último es un modo no ortodoxo de usar las ideas de la teoría matemática de la comunicación, según dice Dretske (1981), p.58), puesto que el objetivo último será entender qué significa decir que un sujeto *conoce* cierto estado de cosas.

Dretske reconoce que puede haber cierta dificultad en el intento de aplicar estas fórmulas a situaciones concretas. De todas maneras, el uso más importante que podemos prever es un *uso comparativo*, que consiste, por ejemplo, en averiguar si  $I_s(r_a)$  es igual o menor a  $I(s_a)$ , esto es, si la cantidad de información que una señal lleva sobre la fuente es igual o menor a la información que se ha *generado* en la fuente. Como veremos, si la cantidad de información es menor, no podrá en realidad hablarse de *conocimiento*.

Con base en lo anterior, y teniendo en cuenta las observaciones que aquí hemos presentado en II), Dretske define "llevar cierta información" del siguiente modo:

"Contenido informativo: Una señal  $r$  lleva la información de que  $s$  es  $F =$  la posibilidad condicional de que  $s$  sea  $F$ , dada  $r$  y  $(k)$ , es 1". (p.70).

[Es decir, la equivocidad debe ser igual a cero; habría que agregar también que la información de que  $s$  es  $F$  debe estar codificada de modo completamente digitalizado]

Sobre esta formulación hay dos aclaraciones importantes que hacer:

- 1. - *La equivocidad presente en una señal siempre es relativa (a) al conocimiento previo del sujeto; y (b) al modo en que se describe el evento.*

(a)  $k$  es aquí lo que el receptor ya sabe (si sabe algo) sobre las posibilidades que hay en la fuente. Por ejemplo: "si ya se sabe que  $s$  es rojo o azul, una señal que eliminara la posibilidad de que  $s$  sea azul (reduce a cero su posibilidad) lleva la información de que  $s$  es rojo (dado que aumenta hasta uno su probabilidad). Para alguien que no sabe que  $s$  es rojo o azul (por lo que esta

persona sabe podría ser verde), la *misma* señal podría no *llevar* la información de que *s* era rojo. (p.70). Una misma señal, por lo tanto, puede llevar información para un individuo y no llevarla para otro. Sin embargo, la mayoría de las veces según Dretske no se gana nada recordando esta necesaria "relatividad" de la noción de información, puesto que se sobreentiende un marco común de conocimientos previos. El conocimiento previo sólo interesa si afecta el valor de  $I(s)$ . Ciertas reflexiones acerca del conocimiento previo del receptor serán sumamente importantes, como luego veremos, a la hora de evaluar cuáles son las posibilidades legítimas para evaluar la información que se transmite y se recibe.

(b) Y, por otra parte, como vimos en II), si los receptores son "sistemas cognitivos genuinos", frente a una misma situación dos receptores con el mismo conocimiento previo, o aún el mismo receptor en dos momentos diferentes, pueden digitalizar contenidos diferentes. Esto ocurre porque, como ya se ha observado, la información de que *t* es *G* puede estar contenida (de modo analógico) en el hecho de que *s* sea *F*. No tiene sentido, por lo tanto, hablar *del* contenido informativo generado por una situación como si éste fuera único. De todas maneras, *dada* una cierta descripción (por ejemplo, con respecto a la descripción de *s* como un *F*) la probabilidad condicional debe ser uno; en caso contrario, la señal *no lleva* la información.

## - 2. - *Las probabilidades condicionales son características objetivas del sistema.*

Debiera distinguirse perfectamente entre la existencia de ciertas probabilidades condicionales y nuestro conocimiento de ellas. "Tal vez no *sepamos* cuáles son esas probabilidades condicionales, tal vez no *podamos* determinar cuáles son, pero eso no tiene ninguna repercusión en el valor real de la equivocidad y por tanto tampoco en si la señal lleva tanta información como genera la fuente." (p.61). En síntesis, que el valor de  $I_S(r_A)$  es independiente de lo que el receptor de  $r_A$  *crea* acerca del contenido informacional de  $r_A$ ; es independiente de lo que *crea* que ha ocurrido en la fuente.

Este último punto, a mi juicio extremadamente importante, se relaciona con el requisito de externalismo requerido por la mayoría de las teorías causales del conocimiento. Más abajo abordaré el problema, que Dretske mismo está interesado en resolver (a diferencia de Goldman, por ejemplo) de cómo puede enfrentar una posición así el desafío del escepticismo.

Según entiendo hay una cierta tensión en la teoría informacional del conocimiento entre privilegiar ciertos aspectos objetivos, o darle preeminencia a otros que no lo son tanto. Esta tensión volverá a aparecer repetidamente.

## 2. La "teoría de la información" contra el escéptico

Volvamos ahora sobre la noción de conocimiento:

K sabe que  $s$  es  $F$  = la creencia de K de que  $s$  es  $F$  está causada, o causalmente mantenida, por la información de que  $s$  es  $F$ " (p.90).

O, de otra manera, una creencia de que  $s$  es  $F$  es conocimiento si *lleva la información de que  $s$  es  $F$* , de acuerdo con el sentido de "llevar una información" que dimos anteriormente.

Dretske aclara que ésta es una ecuación recursiva. Habíamos visto que la información que una persona recibe es relativa a lo que ya conoce sobre las posibilidades de la fuente de información. Sin embargo, que una persona aprenda que  $s$  es  $F$  puede depender de lo que además conozca sobre  $s$ , pero no depende de que conozca que  $s$  sea  $F$ . Y, en último término, supone Dretske que podemos llegar a un punto "en el que la información recibida no depende de ningún conocimiento previo sobre su fuente, y esto es lo que, de hecho, permite que nuestra ecuación evite la circularidad." (p.91). Este tipo de afirmaciones parecen relacionarse con una postura fundacionista sobre el conocimiento, aunque en realidad Dretske nunca lo expresa con estas palabras, y no es del todo claro, a mi modo de ver, que una teoría como la que aquí se está describiendo sea plenamente consistente con el fundacionismo<sup>74</sup>.

El conocimiento de un hecho, según esta definición basada en la teoría de la información, es una noción *absoluta*, no admite grados. Es cierto que podemos recibir diferentes grados de información sobre  $s$ , pero la información *de que  $s$  es  $F$*  es cuestión de todo o nada, y "requiere que la ambigüedad de la señal (en relación a que  $s$  sea  $F$ ) sea cero. No *muy cercana* a cero. Sino cero."

---

<sup>74</sup>. Veremos más abajo, cuando presente su idea de un "canal de comunicación", que su teoría más bien parece describir una suerte de coherentismo.

(p.112). Semejante afirmación parece comprometer a Dretske con que hablar de conocimiento requiere certeza, y esto, según se ha observado en diversas oportunidades, es precisamente un punto de partida que favorece la aparición de argumentos escépticos. Dretske es consciente de ello: "...es seguro que se objetará que bien poco se puede conocer si el conocimiento de que  $s$  es  $F$  requiere no sólo cierta información sobre  $s$  (una cantidad "suficiente" o "adecuada"), sino la información *de que  $s$  es  $F$*  (tal como ha sido definida). Puesto que en la mayoría de las situaciones prácticas las probabilidades condicionales son siempre menores que uno. Siempre hay *alguna* ambigüedad." (p.113). Inmediatamente Dretske introduce los conceptos necesarios para poder, en sus propios términos, "flexibilizar" esta noción absoluta.

La respuesta de Dretske consiste en distinguir cuidadosamente entre lo que es *fente* de información y lo que es *el canal de comunicación*. "El canal de comunicación = el conjunto de condiciones existentes (de las que depende la señal) que 1) o bien no generan información (relevante) o 2) generan sólo información redundante (desde el punto de vista del receptor)" (p.119). Es decir, el canal de comunicación es, justamente aquel conjunto de condiciones tales que posibilita, y no altera, la transmisión de la información generada en la fuente. Si realmente hay condiciones que ofician como canal, entonces, la cantidad de información de  $I_s(r_j)$  no será menor que  $I(s_j)$ , y la equivocidad será cero.

La idea es que siempre hay algo que *debe* funcionar como canal para que alguna otra cosa pueda contar como fuente; si la señal llevara información sobre *el canal*, en realidad este último ya no sería canal para el receptor... sino fuente<sup>75</sup>. El escéptico, dice Dretske, es el que constantemente duda de la confiabilidad del canal; por lo tanto, lo que hace en última instancia es pretender a la vez información sobre la fuente y sobre el canal. Esto es, supone que "las señales que llevan información deben autovalidarse" (p.118). Pero pretender esto es un error. Es cierto que *hay* que revisar el canal en algún momento, (es decir, transformándolo en *fente* de información), para constatar que las situaciones en cuestión, consideradas en un momento posterior como pertenecientes al canal, sean en efecto informativamente redundantes. Pero es un error seguir dudando de la confiabilidad del canal

---

<sup>75</sup>. Creo que es interesante notar que la posición descrita aquí parece implicar algún tipo de coherentismo en materia de justificación de nuestras creencias.

luego de la revisión necesaria. Es un error equivalente al del ingeniero neurótico que decide calibrar sus instrumentos cada cinco minutos: cada nueva revisión es informativamente redundante.

Podría objetarse, sin embargo, que no solamente nuestro ingeniero neurótico pide garantías sobre el canal de comunicación; en algún sentido, todos lo hacemos. De hecho, cada tanto *hay* que calibrar los instrumentos; los aparatos que suponíamos confiables cada tanto fallan, o se rompen. Aquí surge inmediatamente un nuevo problema: ¿cómo identificar las alternativas que son pertinentes, esto es, las que *no* son informativamente redundantes?. La respuesta de Dretske remite, en parte, a Goldman (1976): "Una alternativa pertinente es, según Goldman, aquella que el sujeto del conocimiento debe poder excluir en base a evidencia para poder poseer realmente el conocimiento que le atribuimos" (Dretske (1981), p.133). Es decir: es una *cuestión empírica* cuándo un canal es lo suficientemente seguro como para transmitir información del tipo requerido, esto es, con ambigüedad cero, y por lo tanto no puede ser decidido ni imaginando posibilidades ni ignorándolas. "Ninguna señal puede excluir *todas* las posibilidades si las posibilidades se identifican con lo que se puede imaginar coherentemente. Ninguna señal, por ejemplo, puede eliminar la posibilidad de que hubiera sido generada, no por los medios habituales, sino por algún tipo de accidente cósmico, por un demonio engañoso o por la intervención sobrenatural. Si tales contingencias son consideradas posibilidades *genuinas*, toda señal es ambigua." (p.134). Pero, por supuesto, éstas NO son posibilidades genuinas... porque no tenemos evidencia de que sean alternativas físicamente posibles para nuestro mundo.

Es importante notar que esta teoría constituye una opción *confiabilista*, en el sentido que hemos dado en el capítulo anterior: es un tipo de teoría causal según la cual tener conocimiento depende de la existencia de un proceso confiable (en este caso, se habla de un proceso confiable *de transmisión de información*).

¿Es satisfactoria una respuesta al escéptico de este tipo? Me ocuparé detalladamente de esto en el próximo capítulo.

Antes de continuar sería bueno respondernos la siguiente pregunta: ¿Qué interés especial puede tener una teoría del conocimiento que, como la que acabamos de describir, haga uso de conceptos de teoría de la información? ¿Qué es exactamente lo que se consigue con ella? Responderé de acuerdo con mi visión personal sobre el tema:

(1) En primer lugar, la teoría informacional del conocimiento propone ciertas herramientas teóricas originales. Es decir, brinda una reconstrucción de los fenómenos gnoseológicos y un intento de explicación de por qué tienen la estructura que tienen: porque su naturaleza es informacional. Este tipo de abordaje, por otra parte, permite una interesante unificación de la gnoseología con teorías semánticas y de filosofía de la mente. Esto es, las tres disciplinas comparten la posibilidad de ser explicadas en los términos más básicos de una teoría de la información (convenientemente formulada para que sea posible su aplicación a problemas originariamente ajenos a ella). Por ejemplo, una creencia puede verse como una estructura que queda determinada, y puede ser individualizada, a partir de la relación informacional que le ha dado origen, y que tiene, a su vez, la peculiaridad de ejercer efectos sobre las salidas del sistema, esto es, sobre la conducta<sup>76</sup>.

(2) En segundo lugar, una teoría de este tipo permite manejar datos cuantitativos. Permite, como hemos visto, hablar de *cantidades* de información generada por una fuente, *cantidades* de información que lleva una cierta señal, *cantidades* de información asociada con el conocimiento de un hecho en particular. Aunque este recurso es de aplicación problemática en los casos concretos, la posibilidad del uso de datos cuantitativos, según entiendo, no es totalmente inexistente. Veamos esto con un poco más de detalle.

Dretske se ocupa especialmente del problema de la posibilidad de aplicación concreta de mayoría de las fórmulas de la teoría:

"... ¿Cómo podemos calcular, por ejemplo, la cantidad de información que genera el que Edith esté jugando al tenis? ¿Y cuánta información sobre las actividades de Edith hay en la luz que llega hasta Herman, un espectador casual?"

"Para responder a este tipo de cuestiones [según las fórmulas (i) " $I(s_a) = \log 1/p(s_a)$ " y (ii) " $I_s(r_a) = I(s_a) - E(r_a)$ " ] hay que conocer: 1) las posibilidades

---

<sup>76</sup>. Véase en especial el capítulo 8 de Dretske (1981).

alternativas (por ejemplo, que Edith se esté duchando); 2) las probabilidades asociadas a todas esas posibilidades alternativas, y 3) las probabilidades condicionales (dada la configuración de los fotones que llegan a los receptores visuales de Herman) de cada una de esas posibilidades alternativas. Obviamente, en la mayoría de los escenarios de comunicación normales no conocemos ninguna de las tres cosas. Ni siquiera está muy claro que *pudiéramos* conocerlas. ¿Cuáles son, después de todo, las posibilidades alternativas a que Edith esté jugando al tenis? Probablemente habrá cosas posibles (como que Edith vaya a la peluquería, en lugar de jugar al tenis) y otras que no lo son (que Edith se convierta en una pelota de tenis), pero, ¿cómo empezar a catalogar esas posibilidades? Si Edith pudiera estar haciendo *footing*, ¿hay que contarlo como *una* posibilidad alternativa? ¿O bien debemos contarlo como más de una, dado que podría estar corriendo, casi en cualquier parte, a velocidades muy distintas, y en casi cualquier dirección?" (p.59).

Frente a esta situación, Dretske reconoce que en la mayoría de las situaciones corrientes en las que se transmite información sólo podemos hacer *conjeturas* acerca de la *cantidad* de información generada o transmitida; inclusive, tal vez ni siquiera tenga sentido la pregunta acerca de la cantidad exacta que se transmite, pues tal vez *no haya* una gama definida de posibilidades alternativas. Sin embargo, es posible reivindicar un *uso comparativo* de tales fórmulas. En particular:

1 - "... Se pueden usar estas fórmulas del mismo modo en que se usaría un trozo de cuerda en la que no están marcadas las pulgadas ni los pies. Se puede usar la cuerda para averiguar si A es mayor que B sin determinar la longitud de A ni de B." (p.60). Si me dicen que Ana vive en Argentina, Capital Federal, calle Corrientes, obtengo más información (*respecto* de la información sobre la "dirección exacta de Ana") que si me dicen solamente que vive en Argentina, Capital Federal. Podría pensarse, cosa que Dretske no hace, que el ejemplo es lo suficientemente obvio como para que alguien se pregunte si valía sinceramente la pena desarrollar toda una teoría para llegar a semejante conclusión. Creo que la respuesta a esta reflexión consiste en que la teoría lo que nos da es el marco desde el cual comprender el *por qué* de tal obviedad: en el primer caso la información transmitida es mayor porque se ven reducidas las posibilidades de lugares diferentes (mutuamente excluyentes) en los que Ana vive, dada la información recibida; por lo tanto, por la fórmula (8), la equivocidad será menor, y, por (10), la medida de la dependencia entre información recibida e información generada será mayor. De todas maneras, es

bueno recordar que aún en el caso en el que sólo reciba la información de que Ana vive en Capital Federal, si el contenido informativo que le interesa al receptor es "la ciudad en la que vive Ana", puede que la equivocidad sea 0 respecto de este último contenido informativo, y, por lo tanto, que el receptor conozca. Esto nos lleva a un segundo punto.

2 - Para conocer, entonces, es necesario comparar la cantidad de información que produce un evento, y la cantidad de información sobre ese evento que lleva una señal. Es necesario, pues, evaluar si hay efectivamente equivocidad 0, para poder saber consecuentemente que se ha transmitido tanta información como fue generada. Este es el segundo sentido de "uso comparativo", y a la vez el más importante, que según Dretske puede darse a las fórmulas cuantitativas.

3 - Mi impresión, por otra parte, es que hay tal vez más usos comparativos de estas fórmulas que los que Dretske reconoce explícitamente, e inclusive, más casos en los cuales es posible utilizar, aunque sea de modo aproximado, datos cuantitativos. Tales casos no tienen por qué consistir solamente en ejemplos sobre dados cargados o monedas trucadas. Comentaré mi idea al respecto:

Sugiero retomar el ejemplo de Edith. La información que ha sido transmitida a un receptor determinado K es "hoy, domingo a la tarde, Edith está jugando al tenis" (se lo ha contado un amigo confiable, o tal vez K ha visto la situación personalmente). Supongamos que la equivocidad es cero, y, para ahorrarnos problemas por ahora, supongamos que K no tiene dudas al respecto, esto es, K está convencido de que la equivocidad es cero y, por lo tanto, si le preguntaran afirmaría que sabe que Edith está jugando al tenis. Ahora bien, es necesario presuponer algún conocimiento previo en K (es obvio que, en caso contrario, la situación no le indica nada relevante; ni siquiera podría percibir visualmente a Edith jugando... pues no sabría quién es Edith). Por ejemplo, supongamos que K sabe que el domingo a la mañana Edith estaba en su casa. Pudo quedarse en ella o pudo haber salido. Recuerda además que, aproximadamente, nueve de cada diez domingos por la tarde, cuando no llueve, ha visto a Edith salir de su casa. Puesto que no llueve, esto último es nueve veces más probable. De allí, es un 70% más probable que haya viajado a otra ciudad (dadas, una vez más, las costumbres de Edith). Lo que puede evaluarse aproximadamente, entonces, cuando K recibe la información (con equivocidad 0) de que Edith está jugando al tenis, es la información, subsumida analógicamente en la primera, de que Edith no permaneció en su casa pero sí

en la ciudad. Ello tiene, digamos, una probabilidad del 30% sobre 0.9, esto es, de 0.27. La cantidad de información asociada con esta probabilidad, como es evidente, (1.89 bits, según la fórmula dada en (2)) no será tan alta como hubiera resultado de haber Edith permanecido en su casa (3.32 bits), aunque es más alta que si hubiera salido de la ciudad (0.66 bits).

Una vez que hemos calculado aproximadamente la cantidad de información transmitida sobre un determinado evento/ estado de cosas, puede compararse dicha cantidad con la transmitida por señales diferentes, acerca de eventos/ estados de cosas no necesariamente relacionados entre sí. La teoría informacional del conocimiento puede dar el marco conceptual para saber cuándo una situación genera para un receptor, *aproximadamente*, más información que otra, y así, por ejemplo, evaluar grados de riqueza epistemológica (para un sujeto).

El punto es, para reconocer que es posible efectuar una serie de operaciones como las descritas en los párrafos anteriores es necesario *relativizar* la idea de posibilidades alternativas a un evento dado, las cuales se requieren para poder utilizar la fórmula (2). En un sentido absoluto, puede ser cierto que en más de una oportunidad *no haya* una gama *definida* de posibilidades mutuamente excluyentes. Pero así como Dretske sostiene explícitamente que la equivocidad de una señal es relativa al conocimiento previo del receptor y al modo en que se describe el evento, y que por lo tanto no tiene sentido preguntar por *el* contenido informativo de una señal, así también, creo yo, habría que decir simplemente que en la mayoría de los casos interesantes no tiene sentido preguntar por *la* gama definida de posibilidades alternativas a la producción de un evento, o a un estado de cosas. Esto es, podemos calcular de modo *aproximado* la cantidad de información generada, con base en nuestra evaluación de cuáles son las posibilidades pertinentes. En particular, podemos sostener que es legítimo hacer esta evaluación, que involucra básicamente el uso de (2) y (3), y, por otro lado, seguir exigiendo que para hablar de *conocimiento* la equivocidad debe ser *objetivamente* igual a cero (recordemos que este último concepto supone el cálculo de probabilidades condicionales asociadas con la recepción de una señal, esto es, probabilidades que miden la dependencia entre información generada e información recibida). En realidad, como ya hemos visto, Dretske admite que la noción de conocimiento *también* debe relativizarse en un sentido especial. Pero sobre este último punto volveré en el próximo capítulo.

En síntesis, las posibilidades alternativas son siempre posibilidades respecto de algún modo de presentar la cuestión y respecto del conocimiento previo de un sujeto. Tener en cuenta esta observación permite extender el uso de la idea de cantidades (aproximadas) de información más allá de las situaciones generadas por juegos de azar o semejantes.

## CAPITULO CINCO:

### LA SOLUCION CONFIABILISTA Y EL PROBLEMA DE LA INDUCCION

¿Es satisfactoria la respuesta de la teoría informacional del conocimiento frente al escepticismo? En este punto creo conveniente recordar las observaciones vertidas en los últimos párrafos del Capítulo tres acerca de los presupuestos que parecen ser inherentes a todo confiabilismo. Habíamos distinguido allí básicamente dos presupuestos:

- 1) podemos hacer un uso no problemático de la inducción; y
- 2) tenemos algunas creencias verdaderas.

1) y 2) plantean inmediatamente dos problemas, a saber, el problema de la inducción, en el primer caso, y el problema de cómo saber con seguridad cuáles de nuestras creencias son verdaderas (y por qué), en el segundo. En este capítulo y en el próximo me ocuparé, respectivamente, de ambos asuntos.

Recordemos la propuesta de la teoría informacional frente al escepticismo:

- Un sujeto conoce que  $p$  si su creencia en  $p$  ha sido causada por la información de que  $p$ .

- Para recibir cierta información es necesario que existan condiciones que oficien de *canal* de información.

- "Canal de comunicación" es, por definición, un conjunto de condiciones que son no informativas o informativamente redundantes respecto de la fuente de información.

- Para constatar que cierto conjunto de condiciones es en efecto informativamente redundante, hay que 1) en algún momento  $t_1$  (anterior a la recepción de la información de que  $p$ ) transformar tales condiciones en fuente, y evaluar cuál es la información que generan; 2) tomar en consideración qué otros posibles contenidos informativos pueden generar tales condiciones en un momento  $t_2$ . Pero en este último punto solamente es necesario reparar en cuáles son las posibilidades *genuinas*. Por ejemplo, NO son posibilidades genuinas aquellas acerca de las que no tenemos evidencia de que sean realmente probables (que tengan una probabilidad aproximadamente alta) para nuestro mundo.

- La teoría ofrece un diagnóstico acerca de cuál es el error que comete el escéptico:

\* El escéptico pide la posibilidad de evaluar la información generada por la fuente y por las condiciones del canal *al mismo tiempo*, esto es, pide la *autovalidación* de las señales de información. No comprende que ello es en principio imposible.

\* El escéptico no distingue entre posibilidades genuinas y aquellas que no lo son.

Dejemos de lado por el momento el problema de cómo reconocer que efectivamente una condición es no informativa; este punto se relaciona con el problema de la verdad, y lo abordaré en la segunda parte del capítulo.

Creo que es bastante claro que lo anterior presupone la inducción. ¿En qué sentido esta construcción presupone la inducción? Si en  $t_1$  las condiciones (que componen el canal de comunicación) resultaron no informativas, y si además no hay posibilidades genuinas de que esta situación cambie en  $t_2$ , es razonable pensar que en  $t_2$  las condiciones del canal continúan siendo no informativas. Con lo cual, como resulta evidente, la teoría informacional del conocimiento confía en que lo que ha sido observado hasta hoy como regularidad empírica seguirá siéndolo, y, por lo tanto, lo que resultó no informativo en  $t_1$  lo seguirá siendo en  $t_2$ .

El uso por parte de Dretske de "posibilidades genuinas" es por cierto ambiguo. No está *inequívocamente* claro que esta idea pueda amoldarse exactamente a la formulación de Goldman de lo que es un proceso confiable, dado que para Goldman tal proceso no está obligado a brindar el 100% de las veces creencias verdaderas. Supongamos, para simplificar las cosas, que efectivamente esta idea tiene su equivalente en la teoría de Dretske; supongamos que condiciones *razonablemente* confiables son aquellas que en *la gran mayoría de las veces* (pero no necesariamente en *todas* las veces) permiten transmitir la información asociada con la producción de un evento con equivocidad cero (es decir: condiciones que la mayor parte de las veces funcionarán legítimamente como *canal*) Como ya observé en el tercer capítulo, el escéptico que constituye un adversario interesante para este tipo de teoría es el que no ve razones para confiar en que haya condiciones que permitan transmitir la mayor parte de las veces información con equivocidad cero. Es decir, niega que sea posible realizar la siguiente afirmación: "en general, lo que resultó no informativo en  $t_1$ , será no informativo en  $t_2$ ".

Dretske no formula el desafío del escéptico en estos términos. Si así lo hiciera probablemente se vería forzado a reconocer que su defensa frente al escepticismo resulta en parte incompleta. No parece solucionar el problema de la evaluación de las posibilidades pertinentes. La estrategia es "considerar lo que sucede (y no sucede) de hecho (a un término suficientemente largo) como un índice de lo que puede (y no puede) suceder" ((1981), p.135). Pero ello, como es obvio, lo único que consigue es reubicar la duda escéptica en el marco del *problema de la inducción*. Un escéptico persistente se contentará con señalar que no hay porqué confiar en que las regularidades empíricas de nuestro mundo continuarán siendo tales.

En este punto creo que vale la pena recordar la respuesta de Nozick al problema del escepticismo<sup>77</sup>, ya que tiene interesantes puntos en contacto con la propuesta de Dretske y, en particular, con el problema que nos ocupa en este momento. (En realidad habría que decir mejor que tanto Nozick (1981) como Dretske (1981) son herederos de trabajos previos sobre el tema, en particular de Dretske (1971)). Tal vez en Nozick encontremos un tratamiento satisfactorio para el no demasiado claro concepto de "alternativas pertinentes".

### 1.Nozick y el escepticismo.

Trataré de exponer del modo más sintético posible el intento de Nozick en *Philosophical Explanations* por responder al desafío escéptico. Forzosamente simplificaré la construcción de Nozick en más de un sentido<sup>78</sup>.

El problema que Nozick se plantea es cómo es posible el conocimiento si el escepticismo es lógicamente posible. Su solución consistirá en mostrar por qué a partir de la posibilidad lógica del escepticismo no se sigue que de hecho no conozcamos nada. Para ello comienza por ofrecer una reconstrucción de las condiciones individualmente necesarias y conjuntamente suficientes para conocer:

(1) *p* es verdadera.

---

<sup>77</sup>. En Nozick (1981), capítulo tres.

<sup>78</sup>. Para una selección de artículos sobre el análisis de Nozick sobre el conocimiento y su refutación del escepticismo véase Luper-Foy (1987).

- (2)  $S$  cree que  $p$ .
- (3) Si  $p$  fuera falsa,  $S$  no creería que  $p$ .
- (4) Si  $p$  fuera verdadera (en circunstancias diferentes),  $S$  creería que  $p$ .

Tener conocimiento de que  $p$  consiste en tener una creencia sobre  $p$  con una propiedad especial: es una creencia que *persigue el rastro* (*tracks*) de la verdad de  $p$ . "Un individuo conoce que  $p$  cuando no solamente lo cree cuando es verdadero, sino que también lo creería si fuera verdadero y no lo creería en caso de que fuera falso. No solamente tiene una creencia verdadera real: tiene una creencia verdadera subjuntiva. (...). Será útil acuñar un término para esta situación en la cual la creencia de una persona está conectada subjuntivamente con los hechos. Digamos de una persona que cree que  $p$ , siendo  $p$  verdadero, que, si [las condiciones] (3) y (4) se cumplen, su creencia rastrea la verdad de  $p$ . Conocer es tener una creencia que persigue el rastro de la verdad." (Nozick (1981), p.178). Lo que se quiere es captar la intuición de que (1) y (2) deben estar relacionadas entre sí; esta misma intuición es la que está detrás de las teorías causales y confiabilistas, que pueden ser vistas, en este sentido, como casos especiales de la construcción de Nozick. La estructura básica de la reconstrucción de Nozick, entonces, tiene la siguiente forma:

- (1)
  - (2)
  - (3) no (1)  $\square \rightarrow$  no (2)
  - (4) (1)  $\square \rightarrow$  (2)
- (p.177)

La condición (3) explicita la "condición de variación" del conocimiento, y la (4), la "condición de adherencia": (3) especifica cómo la creencia de un sujeto debe variar con la verdad de lo que es creído, (4) especifica cómo la creencia *no* debe variar cuando la verdad de lo que es creído no varía ((1981), p.211). Sería conveniente también, para completar el análisis, aludir al método utilizado por un sujeto para llegar a adquirir una creencia determinada, pero aquí no trataré en detalle este tema. Señalaré solamente que, según Nozick, "S conoce que  $p$  si y sólo si hay un método  $M$  tal que (a)  $S$  conoce que  $p$  via  $M$  y su creencia via  $M$  en  $p$  satisface las condiciones (1) a (4); y (b) todos los otros métodos  $M_1$  a través de los cuales  $S$  cree que  $p$  y que no satisfacen las condiciones (1) a (4) están 'sobrepasados' (*outweighed*) por  $M$ ." (un método

M<sub>1</sub> está sobrepasado por otro/s método M<sub>2</sub> si cuando M<sub>1</sub> lleva a la persona a creer que *p*, y M<sub>2</sub> a creer que no *p*, el sujeto cree que no *p*). (p.182)<sup>79</sup>.

La flecha " $\Box \rightarrow$ " está aquí usada para expresar un condicional subjuntivo. El punto central del análisis de Nozick, justamente, consiste en que en (3) y (4) no se utilizan condicionales materiales, ni condicionales estrictos (relaciones de implicación lógica) sino que involucran el uso de *condicionales subjuntivos*. La semántica estándar de este tipo de condicionales, en general, apela a la noción de "mundos posibles". Un condicional contrafáctico  $p \Box \rightarrow q$  es verdadero cuando se cumple que no es cierto que *p* & no *q* para todos los mundos posibles *cercanos* al mundo real, o, de otro modo, cuando, en todos los mundos *cercanos al mundo real* en los que *p* es verdadero, *q* también lo es. Este análisis del conocimiento le permite enfrentar satisfactoriamente los contraejemplos de tipo Gettier, y abordar diversos desafíos escépticos. Por ejemplo, respecto de la hipótesis de los "cerebros en una cubeta", el análisis de Nozick permite las siguientes afirmaciones:

a) ¿Puedo afirmar que conozco que en estos momentos estoy escribiendo frente a mi computadora? Según el análisis de Nozick, la respuesta es SI, porque se cumplen las cuatro condiciones dadas anteriormente<sup>80</sup>. El mundo posible más cercano en el cual es falso que estoy escribiendo en una computadora, consiste en un mundo en el cual estoy escribiendo a mano, o tal vez leyendo, pero en ninguno de estos casos creería, por error, estar escribiendo en una computadora. En especial, no es un mundo en el cual soy un cerebro en una cubeta, recibiendo los estímulos necesarios para creer que estoy escribiendo. La última situación descrita es, desde luego, una situación lógicamente posible, pero no corresponde a ningún mundo cercano al mundo real.

b) Pero, por otra parte, no sabemos que *no* somos cerebros en una cubeta. Si fuera falso que "no soy un cerebro en una cubeta", aún lo creería, esto es, aún estaría convencida de la verdad de "no soy un cerebro en una cubeta" (según el escéptico ésta podría ser la triste situación actual). Por lo tanto, (3) no se cumple, y no podemos hablar de "conocimiento" en este caso.

---

79. Véanse para este tema las páginas 179-196 de Nozick (1981). Véase también el apartado sobre "Evidencia", (1981), páginas 248-288.

80. Notemos que en este análisis, al igual que en el de Goldman o Dretske, también se presupone la posibilidad de determinar sin problemas la verdad de una proposición.

(supuestamente, según Nozick, dar cuenta de esta intuición es un mérito de su análisis).

¿Cómo es que Nozick considera compatibles las afirmaciones a) y b)? Tal compatibilidad parece extraña, por cuanto si soy un cerebro en una cubeta, de allí se sigue que es falso que estoy escribiendo frente a una computadora. Y si no conozco que no soy un cerebro en una cubeta, ¿no debería decir que tampoco conozco que estoy escribiendo en este momento (y, en general, que no conozco prácticamente NADA)? a) y b) son compatibles, según Nozick, porque no vale el principio de cierre según el cual si un individuo conoce una implicación y conoce el antecedente, conoce también el consecuente:

( La flecha " $\Rightarrow$ " significa "implica lógicamente"). No vale que:

$[C(p \Rightarrow q) \ \& \ Cp] \Rightarrow Cq$

(Nozick (1981), p.204).

Nozick también se ve obligado a rechazar la clausura deductiva para otros principios; por ejemplo, para el principio de instanciación,  $(x) Fx \Rightarrow Fa$ , y para la conjunción, aunque la admite para la equivalencia y la generalización existencial. En ciertos casos, sus observaciones al respecto parecen bastante *ad hoc*. Sin embargo, algunos de los motivos que da Nozick para rechazar el principio de cierre son muy interesantes. Nozick muestra que la condición (3) *presupone* que no vale el principio de cierre. Por lo tanto, si es correcto incluir a (3) en un análisis del conocimiento, es correcto también rechazar el principio de cierre.

En primer lugar, entonces, habría que explicar por qué (3) supone la invalidez del principio de cierre. Definamos primero "mundos doxásticamente idénticos para un sujeto". Dos situaciones son doxásticamente idénticas para S si y sólo si S tendría exactamente las mismas creencias en ambos. (p.202). (El escéptico, por ejemplo, describe mundos doxásticamente idénticos al real en los cuales prácticamente todas nuestras creencias son falsas). Dada la hipótesis de que  $p \Rightarrow q$ , puede ocurrir que (3) se cumpla para  $p$  y no para  $q$ , puesto que lo que S creería si  $p$  fuera falsa puede ser muy diferente de lo que S creería si  $q$  fuera falsa. Por ejemplo, nacer en Buenos Aires implica lógicamente nacer en el plantea Tierra. Sin embargo, la situación de no haber nacido en Buenos Aires es muy diferente de la situación de no haber nacido en la Tierra; mis creencias en uno y otro caso seguramente también serán diferentes. Es decir: no hay razón para pensar que el mundo más cercano *no p* y el mundo más cercano *no q* son doxásticamente idénticos para S. Aún si  $p \Rightarrow q$ , no hay razón

para pensar que mis creencias en un mundo *no q* serán una subclase de mis creencias en un mundo *no p*. Ahora bien, la razón profunda de por qué ello es así, radica según Nozick en que lo que se preserva en la implicación lógica es *la verdad*. No habla de lo que ocurre cuando *p* o *q* son falsas, y poder especular con esto último es la característica esencial de la condición (3).

En segundo lugar, ¿es correcto incluir a (3) en un análisis del conocimiento? La respuesta es que (3) parece dar cuenta perfectamente de nuestras intuiciones pre-teóricas. Tanto es así que la hipótesis escéptica *necesita* a (3) para sostener que "no conocemos que no somos cerebros en una cubeta". Pero (3) y el principio de cierre forman un todo inconsistente. Con lo cual Nozick está en condiciones de arriesgar un diagnóstico sobre cuál es el problema con el escepticismo: el escepticismo es contradictorio porque presupone (3) y el principio de cierre a la vez. O mejor dicho: no es contradictoria la hipótesis escéptica (del tipo "no sabemos que no somos cerebros en una cubeta"), sino el intento por impugnar todo nuestro conocimiento a partir de ella.

Entonces, el escéptico está equivocado en creer que no conocemos nada. También está equivocado el dogmático. El resultado del análisis de Nozick, pues, es en una visión razonablemente acotada del conocimiento: conocer no significa excluir toda posibilidad lógica de creer falsedades. En esto, como podemos ver, está plenamente de acuerdo con el espíritu del confiabilismo tal como lo hemos descrito en los capítulos anteriores.

Notemos que puede hacerse un estrecho paralelismo entre el uso que Dretske da al concepto de "alternativa pertinente (o genuina)" y el uso de "mundos posibles cercanos" en Nozick. En términos del lenguaje de "mundos posibles", el mundo en el cual un genio o demonio cósmico interviene para deformar el contenido informativo que *S* recibe, no es evidentemente un mundo posible cercano; en el análisis de Nozick, ello significaría que (3) aún se cumple en este caso. ¿Tal vez el análisis de Nozick permite aclarar la noción poco explicada de "alternativa pertinente"? ¿Resulta ahora, la solución de Dretske frente al escéptico más plausible? Lamentablemente, creo que no.

La respuesta es no, porque sigue habiendo un problema de base, que Nozick tampoco se ocupa de solucionar. Esto es, Nozick incurre básicamente en la misma presuposición que Dretske. Nozick señala que podemos construir cierta jerarquía de conjuntos de mundos posibles, ordenados según sea su "distancia" respecto del mundo real. Pero en ningún momento repara en cuáles

son los fundamentos para elegir un ordenamiento y no otro<sup>81</sup>. Según entiendo, el fundamento reside en las asignaciones de probabilidad con las que nos manejamos habitualmente; Nozick también, en resumidas cuentas, presupone que puede hacer uso de la inducción de modo no problemático.

Curiosamente, Nozick sí intenta dar cuenta del problema de la inducción. Pero lo aborda.... a partir del mismo esquema de análisis con que se ha manejado hasta el momento, y por lo tanto presupone lo que quiere justificar:

"Consideremos (...) escepticismo acerca de la inducción, acerca del conocimiento del futuro o acerca de cualquier inferencia de una muestra a una población, o a hechos fuera de la muestra. Sostenemos una creencia de que  $p$  sobre el futuro con base en algunos hechos presentes y pasados. ¿Conocemos que  $p$ ? Sólo si nuestra creencia de que  $p$  rastrea el hecho de que  $p$ . Puesto que nuestra creencia en  $p$  está basada en hechos presentes y pasados, conocemos que  $p$  sólo si esos hechos presentes y pasados rastrean el hecho de que  $p$ . Esto es, sólo si: si  $p$  no fuera verdadero, esto habría sido reflejado de modo tal que algunos de los hechos presentes y pasados no habrían ocurrido. ¿Conocemos que el Sol saldrá mañana, que la Tierra continuará rotando sobre su eje durante las próximas 24 horas? Si el Sol no fuera a salir mañana, ¿no lo habríamos previsto; tal alteración en la rotación de la Tierra no habría sido presagiado en los hechos disponibles para nosotros hoy y antes? Si esto es así, entonces *conocemos* que el Sol saldrá mañana; nuestra creencia de que  $p$  rastrea el hecho de que efectivamente sucederá, debido a que se basa en hechos que, en caso contrario, habrían sido diferentes.

Pero ¿no es lógicamente posible que todo haya ocurrido como ocurrió, hasta hoy, y aún así la Tierra no continúe rotando mañana? Sí, hay tales posibilidades lógicas escépticas SK: el pan ya no nos alimenta, el Sol se frena en el cielo, un evento de cierto tipo no continúa produciendo su efecto usual. (...) El escéptico respecto de la inducción está en lo correcto al decir que no conocemos que tales posibilidades no se siguen, pero está equivocado al negar que conocemos tales resultados particulares de inferencia inductiva cuya falsedad hubiera sido reflejada en los hechos pasados, y presagiada en los hechos sobre los cuales basamos la inferencia. " ((1981), p.222-223).

---

<sup>81</sup>. Sobre esta cuestión véase por ejemplo, R. Fumerton, "Nozick's epistemology", en Luper-Foy (1987), pp.163-181. Fumerton de todas maneras no acusa a Nozick de presuponer la inducción; solamente señala algunos problemas que acarrea el uso de la noción de mundos posibles para un análisis del conocimiento (en el marco general de una crítica al naturalismo y al externalismo).

Utilizar la inducción para justificar la inducción no necesariamente está mal, pero, por supuesto, habría que reconocer que eso es efectivamente lo que se está haciendo, y dar alguna razón de por qué en ese caso no se comete circularidad (o por qué se comete, tal vez, algún tipo de circularidad legítima)<sup>82</sup>.

## 2. La teoría informacional nuevamente.

Si las observaciones anteriores son correctas, la teoría informacional del conocimiento proporciona herramientas para responder al escéptico, pero no completa su tarea. No constituye, entonces, una solución plenamente satisfactoria hasta tanto no resuelva los problemas que generan sus presupuestos implícitos.

Pese a todo, considero que en la construcción de Dretske están sentadas las bases de una posible estrategia para *eludir* el problema de la inducción (como ya he aclarado, del problema de la verdad me ocuparé más adelante). Para ello debería ponerse énfasis en los aspectos "relativistas" inherentes a la teoría informacional del conocimiento. Dretske en varios sentidos los admite, pero al mismo tiempo trata de minimizarlos; por el contrario, aquí sugiero profundizar en ellos, aunque respetando el lineamiento general de la teoría.

Creo que podemos identificar dos defensas frente al escepticismo. Una de ellas, la defensa "fuerte", presupone el uso de la inducción. Pero puede reconocerse una segunda defensa, más débil. Habíamos mencionado que las condiciones que no originan información nueva son el marco de referencia *dentro del cual* se da la comunicación (Dretske (1981), p.121). Entonces, por qué no razonar del siguiente modo: *dado* cierto canal, que *por hipótesis* consideramos no informativo, se transmitirá cierta información. La

---

<sup>82</sup>. Véase al respecto Papineau (1992). Allí Papineau sostiene, siguiendo algunas ideas clásicas que pueden encontrarse en Braithwaite y otros, que es necesario distinguir entre "circularidad de premisas" (*premise circularity*) y "circularidad de regla" (*rule circularity*). Inducciones particulares pueden justificarse con base en la afirmación (que resulta a su vez de la conclusión de un razonamiento inductivo) de que las inducciones, en general, llevan a la verdad; tal justificación sólo presupone "circularidad de regla". La "circularidad de regla" sería intuitivamente aceptable, por cuanto inclusive los razonamientos deductivos caen en ella (y, entonces, el precio de rechazarla sería exorbitante). Goldman había recordado esta misma idea en (1986), p. 394.

información, y por ende el conocimiento, quedaría así *relativizado* a un canal de comunicación.

Por supuesto, aún es necesario mantener el segundo presupuesto del confiabilismo: que es posible advertir sin mayores problemas cuándo una creencia es verdadera. Es preciso haber podido reconocer en algún momento que las condiciones que permiten la transmisión de comunicación son efectivamente no informativas, esto es, que constituyen un auténtico canal de comunicación. Pero con la suspensión, por hipótesis, del juicio sobre la confiabilidad del canal, ya no se hace necesario recurrir a la inducción. La estrategia consiste entonces en "congelar" el juicio sobre la verdad de las condiciones que conforman el canal; las condiciones que en un momento  $t_1$  resultaron evaluadas como no informativas, serán consideradas en  $t_2$ , esta vez *por hipótesis*, como condiciones confiables, o condiciones que permitirán transmitir información con equivocidad cero para la gran mayoría de las veces.

Debe notarse que la posición que resulta es en un sentido importante mucho más débil que la originariamente propuesta por Dretske. En efecto, podría decirse que no estamos explicitando, en realidad, la noción de "conocimiento", sino la de "conocimiento-en-C", o "conocimiento adquirido a través de cierto canal de comunicación". Puesto que hemos suspendido el juicio sobre la posibilidad de error en nuestra apreciación del canal, todo conocimiento se escuda en la forma de un condicional: "dado que C es un verdadero canal de comunicación (por hipótesis), la creencia de que  $p$  generada con base en C (a partir de la información de que  $p$ ) constituye conocimiento". La garantía de la confiabilidad de un canal queda justificada por la mera decisión de transformar una serie de condiciones en canal. Y revisar el canal implica simplemente... armar un *nuevo* condicional, donde otras condiciones ocuparán el lugar del antecedente; la pretensión de revisar canal e información a la vez, directamente *no tiene sentido*. Por supuesto, se puede plantear el problema de cómo justificar tales decisiones; no deseamos transformar en canal a cualquier conjunto arbitrario de condiciones. Posiblemente en este estadio ya no podemos escapar de la necesidad de recurrir a la inducción<sup>83</sup>. Pero este uso no es problemático: la conclusión (del

---

<sup>83</sup>. Puede que haya algún modo de recurrir a "buenas razones" que no presuponga inducciones (¿decisiones irracionales de una comunidad, tal vez?), pero aquí no lo consideraré. Este no es, de todas maneras, un punto que afecte mi argumentación.

razonamiento inductivo) que afirme "C es un canal (C es un conjunto de condiciones no informativas)" sólo aparecerá, en nuestro análisis de "conocer-en-C", como antecedente de un condicional.

¿Hasta qué punto esta heterodoxia concuerda con el espíritu de la teoría tal como la presenta Dretske, o hasta qué punto se desvía? Según he mostrado en el capítulo anterior, la concepción informacional del conocimiento acusa una cierta tensión entre aspectos más objetivos y aspectos que lo son menos. Por ejemplo, Dretske exige que las probabilidades condicionales no sean subjetivas, es decir, las probabilidades condicionales (las cuales, como sabemos, determinan si la equivocidad es cero o no) son las que son, con independencia de toda posible evaluación de ellas<sup>84</sup>. Pero, por otro lado, Dretske mismo admite que el recurso de identificar un canal de comunicación debe servir como un medio para "flexibilizar" la noción absoluta de conocimiento. Sin embargo, no veo que con este recurso por sí solo se consiga ningún tipo de "flexibilización". Es decir, al discernir entre fuente y canal de comunicación, lo que ocurre es que la carga de la duda recae sobre las condiciones del canal: son ellas las responsables de que haya transmisión de información (con equivocidad cero) o no. Pero puesto que se sigue exigiendo la objetividad de las probabilidades condicionales, lo que se ha conseguido simplemente es el requisito de un *canal objetivamente no informativo*. ¿Dónde está pues la flexibilización? Ella se consigue, en cambio, cuando transformamos el análisis de "conocer" en "conocer-en-C". De modo que esta última propuesta no estaría sino conservando, e intentando llevar a cabo, algunas de las intenciones originales de la teoría.

Sería importante considerar las implicaciones de contar con un análisis para (tan sólo) el concepto de "conocer-en-C". A dicho análisis puede enfrentársele una afirmación como la siguiente: " 'Conocer-en-C' constituye, en el mejor de los casos, una construcción limitada, que debería complementarse con (o tal vez reemplazarse por) una teoría que dé cuenta de nuestra noción habitual de conocimiento, que es una noción absoluta". No me expediré sobre esta última afirmación, sobre todo porque no me he ocupado de analizar intentos por justificar la inducción, los cuales tal vez permitirían evitar la relativización del conocimiento a un canal. De todos modos, creo que es perfectamente razonable suponer que nuestra concepción ordinaria de

---

<sup>84</sup>. La evaluación de las probabilidades condicionales se relaciona con el problema de la verdad, y por lo tanto lo trataré en la próxima sección.

conocimiento consista solamente en "conocer a través de un canal", y que más allá de ella no haya nada de importancia que analizar. Ahora bien, ¿esto último no supone una victoria, tal vez parcial, del escepticismo? En algún sentido, dar una respuesta al escepticismo desde una perspectiva "condicional" del conocimiento puede parecer en buena medida una admisión de nuestra impotencia: sólo podemos asegurar que, por ejemplo, *dado* que no dudo de mis capacidades auditivas, ciertos ruidos me permiten afirmar que *conozco* que alguien o algo está del otro lado de la puerta; dada una larga serie de condiciones que no discuto por hipótesis, puedo afirmar que sé que hay siete voltios asociados a cierta fuente de energía, o que sé que el organismo de A genera 37 grados de temperatura (dado, entre otras cosas, que no dudo del termómetro que utilizo ni de mis capacidades visuales). Sin embargo, pese a cierta aparente trivialidad, no me parece que la estrategia sea inútil. Los planteos del escéptico nos seducen, y nos provocan vértigo, porque *piden lo incondicionado*. Frente a esto la estrategia comentada consiste, simplemente, en recordar cuáles son los límites obvios de todo aquello que podemos pretender creer justificadamente.

## CAPITULO SEIS:

### LA SOLUCION CONFIABILISTA Y EL PROBLEMA DE LA VERDAD

Retomaré ahora el segundo presupuesto que habíamos identificado en el confiabilismo, a saber, cómo puedo asegurar que ciertas creencias son verdaderas. El confiabilismo, recordemos, decididamente no puede verse como una posición que intenta ofrecer un criterio *de verdad*, so pena de transformar su formulación en una flagrante circularidad. Habíamos dicho también que una respuesta confiabilista frente al escepticismo debe poder mostrar que HAY procesos confiables; entonces, debemos poder mostrar que contamos con creencias verdaderas; ello supone, a su vez, que podemos evaluar su verdad sin mayores problemas.

Es importante notar que el problema que debemos enfrentar en este momento no solamente concierne a la búsqueda de *criterios* de verdad, sino que está íntimamente ligado con consideraciones acerca de lo que la verdad ES. Como hemos visto en capítulos anteriores, en buena medida la raíz de las dificultades reside en el inmenso potencial escéptico de una concepción semánticamente realista de la verdad. De tal concepción de la verdad se infiere que un sujeto *siempre* puede tomar por verdadera una proposición que, de hecho, es falsa. No hay modo entonces de asegurar que NO estamos masivamente equivocados en nuestras creencias; no tenemos tampoco, por principio, ningún derecho a esperar que en el futuro se reduzcan las posibilidades de error masivo.

Ahora bien, ¿cuán fuertes son las afinidades del confiabilismo con el realismo semántico? Goldman, ya lo dijimos, concibe la verdad como una noción no epistémica. ¿Ocurre algo similar en el marco de la teoría informacional del conocimiento? Reparemos en un detalle tal vez curioso: hasta este momento no fue necesario mencionar ninguna opinión de Dretske respecto de la verdad; aún más, puesto que Dretske no formula su teoría siguiendo el esquema confiabilista más clásico ("un proceso confiable es aquel que produce un alto porcentaje de creencias verdadera") las menciones al concepto de "verdad" fueron mínimas (excepto en mis observaciones acerca de que la verdad estaba siendo presupuesta). ¿Esta situación implica tal vez que Dretske ha preferido dar una reconstrucción del conocimiento que excluya el concepto de verdad? ¿En qué sentido puede hacer esto una teoría confiabilista?

Es cierto que en la definición de conocimiento que transcribí más arriba no se menciona la verdad. Sin embargo, debemos considerarla *subsumida en la noción primitiva de información* : "Estos términos [los que encontramos en un diccionario bajo la entrada de 'información'] son sugerentes, tienen un núcleo común: todos señalan en una misma dirección, la de la verdad. La información es la que puede proporcionar conocimiento, y *dado que la verdad es necesaria para el conocimiento, también lo es para la información.* " (Dretske (1981), p.51). En ningún otro lugar reaparece esta alusión a la idea (tradicional) de que la verdad es necesaria para conocer; como sea, contamos aquí con el compromiso explícito de que el conocimiento incluye a la verdad como condición *sine qua non*. Lo que ocurre simplemente es que "transmitir la información de que *p* a un sujeto" es el equivalente de "generar en el sujeto en cuestión una creencia *verdadera* sobre *p*"<sup>85</sup>. "La información que lleva una señal es lo que nos puede "decir", lo que nos puede decir *con verdad*, sobre otro estado de cosas [esta idea ya fue mencionada en los primeros párrafos del Capítulo cuatro]. En general, la información nos puede proporcionar conocimiento, y la información que lleva [una señal] es lo que podemos saber gracias a ella. Si todo lo que yo le digo a alguien es falso, no le he dado ninguna información. O, por lo menos, no le he dado ninguna información del tipo que pretendía estar dando. Si (por otras razones) esa persona sabe que lo que digo es falso, lo que estoy diciendo le proporcionará información de todos modos, informará sobre *mí* (estoy mintiendo), pero no obtendrá la información que corresponde al significado convencional de lo que digo. Cuando digo "tengo dolor de muelas", lo que digo *significa* que tengo dolor de muelas tanto si lo que digo es cierto como si es falso. Pero, cuando es falso, no lleva la información de que tengo dolor de muelas, porque no puede llevar al conocimiento de que tengo dolor de muelas. Nadie puede saber que tengo dolor de muelas cuando no lo tengo. (...) En este sentido del término, la

---

<sup>85</sup>. Ello se conecta con un problema, específico de filosofía de la mente, que retomaré más abajo. El problema es, una creencia parece ser el tipo de cosa de la cual tenemos derecho a decir que es falsa. Pero por definición de creencia (estructura que ha sido generada por un proceso informativo, que es capaz de codificar información de modo completamente digitalizado, y que tiene control sobre las salidas del sistema), y dado que los conceptos de "información" y "verdad" están íntimamente ligados, no parece haber espacio en la teoría para hablar de "creencias falsas". Dretske soluciona el problema distinguiendo entre "creencias tipo" y "creencias caso"; solamente pueden ser falsas las segundas. (Dretske (1981), capítulo 8).

información *falsa* y la información *inexacta* no son dos tipos de información, del mismo modo que los patos de goma no son tipos de patos." ((1981), p.50-51).

¿De qué concepción de la verdad se trata aquí? Recordemos las siguientes equivalencias:

(1)-  $r$  lleva la información de que  $s$  es  $F$  si y sólo si la posibilidad condicional de que  $s$  sea  $F$ , dada  $r$  y  $(k)$ , es 1; es decir:

(2)-  $r$  lleva la información de que  $s$  es  $F$  implica que en la fuente se ha generado la información de que  $s$  es  $F$ .

Por otra parte, no parece demasiado forzado suponer que es válida la siguiente afirmación:

(3)- En la fuente se genera la información de que  $s$  es  $F$  si y sólo si  $s$  es  $F$ .

Pero además, dada la estrecha relación entre "información" y "verdad":

(4)- En la fuente se genera la información de que  $s$  es  $F$  si y sólo si la proposición de que  $s$  es  $F$  es verdadera.

De donde:

(5)- " $s$  es  $F$ " es verdadera si y sólo si  $s$  es  $F$ . (de (3) y (4))

Es decir, se acepta el esquema T.

Por supuesto, con esto solo, tal como vimos en el primer capítulo, no basta para suponer que hemos optado por una concepción realista o antirrealista sobre la verdad. En fin, no es demasiado arriesgado suponer que por lo menos Dretske admite el requisito tarskiano de adecuación material, pero ello solo no nos compromete necesariamente con algún tipo de realismo, pues el esquema T es, por sí mismo, perfectamente neutro en lo que concierne a la dependencia o independencia de la noción de verdad respecto de cuestiones epistémicas.

Creo que es saludable tener presente esta falta de definición de la teoría que nos ocupa respecto de la verdad. Sin embargo, como ya mencioné en el capítulo tres, el confiabilismo parece tener cierta natural afinidad con el realismo semántico. En el caso de la teoría informacional ello se manifiesta en el énfasis que Dretske pone en la *objetividad* de las probabilidades condicionales. Tal objetividad, como es evidente, está en función de la objetividad de la información generada en la fuente. Es decir, la información

generada en la fuente es totalmente independiente de cuestiones epistémicas. Utilizando nuevamente las equivalencias recordadas más arriba, es fácil concluir que se presupone un sentido realista de verdad. (Es decir: si la información generada en una fuente es independiente de consideraciones subjetivas, y la verdad de un enunciado consiste en la descripción de la información generada en una fuente, entonces la verdad de un enunciado es independiente de consideraciones subjetivas).

El problema del que debemos dar cuenta tiene, creo, dos grandes maneras de plantearse y responderse. Ninguna de las dos, hasta donde sé, es elegida explícitamente por algún confiabilista. Con lo cual lo que sigue implica admitir algunos cambios en la teoría, con el objetivo de hacer más efectiva la estrategia (en particular, de la teoría informacional del conocimiento) frente al escepticismo.

- Una posibilidad consiste en afirmar que no tiene sentido exigir la verdad (en sentido semánticamente realista) de nuestras creencias, sino que basta con que dichas proposiciones sean *creídas* (justificadamente) como verdades. Es decir, para conocer es fundamental disponer de algún buen *criterio, o criterios*, de aceptación de creencias como verdaderas, pero no es necesario abrir juicio sobre la *naturaleza* de la verdad, ni sobre si estamos *efectivamente* en posesión de ella. *Conocimiento* no sería así "creencia verdadera adquirida de modo confiable", sino "creencia tomada (justificadamente) como verdadera adquirida de modo confiable". En este caso, como resulta obvio, se hace preciso redefinir, o al menos precisar, el sentido en el que el confiabilista se dice "externalista".

- Una segunda posibilidad consiste en optar por una concepción no realista de la verdad. Como hemos visto en el capítulo dos, por otra parte, debe tenerse en cuenta que no *cualquier* teoría no realista servirá para nuestros propósitos.

Como espero haber mostrado, la teoría informacional no aborda ninguna de estas dos posibilidades: utiliza, aunque de modo solapado, el concepto de verdad, y, si bien no es explícita en esto tampoco, no parece que tengamos demasiados motivos para creer que se presupone una concepción no realista de la verdad; antes bien, la alusión de Dretske a la objetividad de las probabilidades condicionales, y el aire de familia que comparte toda teoría causal del

conocimiento con teorías causales de la referencia (en las cuales el compromiso con el realismo semántico es claro) hacen suponer lo contrario.

El mérito principal de cualquiera de las opciones anteriores consiste en que la estrategia contra el escéptico se torna convincente (o, por lo menos, un poco *más* convincente). Por supuesto, en ninguna de las dos faltan problemas. En lo que sigue procuraré analizar estas dificultades.

### 1. Abandono del requisito de verdad.

I)

¿Abandonar el requisito de verdad para conocer no implica abandonar, conjuntamente con ello, el externalismo? Creo que no es forzosa una respuesta afirmativa a esta pregunta. El concepto de "externalismo" no define unívocamente a una determinada posición, sino que es en varios aspectos ambiguo. En lo que sigue señalaré algunas posibles divergencias de interpretación, e indicaré cuál es la versión de "externalismo" que considero fructífera. Tal versión no es incompatible con el abandono de la condición de verdad.

En su formulación más general, el externalismo afirma que para saber no es necesario saber que se sabe<sup>86</sup>. El conocimiento es un proceso objetivo. En otra formulación, "una teoría de la justificación es *internalista* si y sólo si requiere que todos los factores necesarios para que una creencia esté epistémicamente justificada, para una persona determinada, deban ser *cognitivamente accesibles* a dicha persona, *internos* a sus perspectivas cognitivas; y es *externalista*, si permite que por lo menos algunos de los factores necesarios para la justificación no necesiten ser accesibles, de modo que pueden ser *externos* a la perspectiva cognitiva de quien posee la creencia".<sup>87 /88</sup>. (Dicho sea de paso, aquí vemos una vez más discriminar entre una teoría del conocimiento y una teoría de la justificación).

---

<sup>86</sup>. Véase Dancy (1985), capítulo 3, punto 3.5.

<sup>87</sup>. Laurence Bonjour, en Dancy y Sosa (1994), p.132.

<sup>88</sup>. L. Bonjour también señala que no debe identificarse "internalismo" con "existencia de factores de justificación que sean estados mentales". Creo que ello es importante, pues, en caso contrario, entiendo que se estarían mezclando planos de discusión. Por ejemplo, Goldman señala (en Dancy y Sosa (1994), pp.433-436) que los procesos confiables de

¿Cuál es el alcance de estas afirmaciones? En particular, parece que hay algún requisito de *menos* que pide el externalista frente al internalista: el *reconocimiento* de la existencia de justificación/conocimiento. Pero desde otra óptica ello significa, entonces, que el externalista pide algo *más*: exige que sea admitida la posibilidad de conocimiento/justificación *sin* el requisito del internalista. Ahora bien, ¿qué es exactamente lo que el externalismo considera que no es necesario para conocer? Para que *S* conozca que *p* (en un sentido objetivo), (a) ¿no es necesario que *nadie* sepa que *S* conoce que *p*? ¿O (b) lo que dice el externalismo es tan sólo que no es necesario que *S mismo* conozca que conoce? No es usual hacer esta distinción entre (a) y (b); sin embargo, creo que es un buen modo de discriminar entre una interpretación del externalismo extrema y difícil de sostener, y otra interpretación más plausible.

Recordemos algunas de las afirmaciones de Goldman (1979) que en el capítulo tres habíamos encontrado difíciles de conciliar con su autoproclamado externalismo. Goldman terminaba reconociendo que la razón por la cual intuitivamente consideramos que una creencia está justificada consiste en que *creemos* que ha sido formada por un proceso confiable; por lo tanto el confiabilismo debía reparar solamente en nuestras *creencias* sobre la confiabilidad. Ahora bien, observemos que tales afirmaciones sólo son contradictorias con (a), pero no con (b). (a) exige la posibilidad de que haya justificación aún cuando nadie (ningún ser humano) pueda ser consciente de ella. (b), en cambio, proclama la inexistencia de "acceso privilegiado" por parte de un sujeto a los factores que justifican sus creencias.

Según creo, (b) constituye una posición interesante y defendible, pero no sucede lo mismo con (a). (b) puede entenderse como un caso particular de la idea más general de una "defensa de la perspectiva de la tercera persona". Según esta perspectiva, tanto las creencias como los significados no son entidades privadas, sino públicas; ello quiere decir, básicamente, que pueden constituirse como objeto de estudio de la ciencia, pues son "observables" intersubjetivamente. Atribuimos creencias y deseos a un sujeto para poder explicar su *conducta* (observable), en lo que algunos autores llaman el uso de una "actitud intencional" como estrategia interpretativa (Dennett, por

---

justificación involucran procesos psicológicos de formación de creencias, y que en ese sentido su teoría no es externalista. Pero me parece que con esta afirmación Goldman está atribuyendo a los términos externalismo/internalismo un significado completamente diferente del estándar. Su confiabilismo ES externalista por cuanto para hablar de conocimiento no se exige que tales procesos psicológicos ("internos") sean reconocibles como factores de justificación por el sujeto que los soporta.

ejemplo<sup>89</sup>). Ciertos teóricos, inclusive, suponen que debe abordarse el estudio de las creencias y deseos propios (esto es, en "primera persona") como un caso *límite* de la tercera persona, lo cual constituye una inversión del planteamiento clásico del tema<sup>90</sup>. En este mismo espíritu, las teorías causales del significado al estilo de Kripke o Putnam han acuñado el difundido slogan de que "los significados no están en la cabeza"; ése es el sentido del famoso experimento mental de Putnam de "la Tierra Gemela"<sup>91</sup>. Supóngase la siguiente situación (lógicamente posible): existe una réplica exacta de la Tierra y de sus habitantes (la Tierra Gemela); en particular, reparemos en la existencia de Oscar, habitante de la Tierra, y de Oscar Gemelo, su *Doppelgänger* en Tierra Gemela. Tierra y Tierra Gemela, Oscar y Oscar Gemelo, son réplicas moleculares idénticas, salvo por un pequeño detalle: la composición química del "agua" en la Tierra es H<sub>2</sub>O, y en Tierra Gemela es XYZ. Tanto Oscar como Oscar Gemelo utilizan la misma palabra, "agua", para referir, respectivamente, a H<sub>2</sub>O y a XYZ. Según Putnam, intuitivamente no diríamos que "agua" en la Tierra y en Tierra Gemela "significan" lo mismo, pese a que, por hipótesis, la estructura psicológica de ambos Oscars es idéntica. El significado, entonces, es una matriz compuesta, en la cual la *referencia* de un término, esto es, su extensión, (que constituye un hecho objetivo del mundo, y de cuyo estudio debe ocuparse la ciencia) aparece como factor determinante<sup>92</sup>. Dretske, dicho sea de paso, reivindica, aunque con ciertas modificaciones, el ejemplo de Putnam; por ejemplo le agrega la idea, de su cosecha, de que la relación (causal) que media entre una referencia y el correspondiente significado es (¡obviamente!) una relación informacional<sup>93</sup>.

---

89. Véase Dennett (1989).

90. Dennett, por ejemplo, llama a su programa de investigación "hetero-fenomenología", en clara alusión a conceptos husserlianos, a los que desea despojar de connotaciones cartesianas. (Dennett (1989), cap. 5, "Beyond Belief", pp.117-202).

91. Putnam (1975).

92. Entre otras cosas, lo que hace Putnam aquí es invertir el planteo de Frege según el cual el sentido *determina* la referencia.

93. Con lo anterior no pretendo, ni remotamente, uniformar las posiciones de Dretske, Dennett y Putnam, que en numerosos aspectos son positivamente muy difíciles de conciliar (en especial, Dennett es extremadamente crítico de lo que llama "la tesis de la intencionalidad originaria", de la cual sería partidario Dretske). Tan sólo intento decir que ellas presuponen un cierto "trasfondo" común, y es el rechazo, para ciertas situaciones específicas, de una posición que podríamos llamar "cartesiana".

Así pues, la "perspectiva de la tercera persona" asegura que ningún sujeto tiene acceso privilegiado ni a sus deseos y creencias, ni a los significados de su lenguaje<sup>94</sup>. La versión de "externalismo" que considero conveniente adoptar aquí sigue esta misma línea: específicamente, se compromete con la idea de que ningún sujeto tiene *acceso privilegiado* a aquellos factores (otras creencias, procesos confiables, etc.) que *justifican* una creencia/ o que permiten hablar de conocimiento. Como es obvio, esta afirmación consigue *objetivar* los factores que constituyen requisitos para conocer, y permite que distintas disciplinas científicas (por ejemplo, según Goldman (1986), la psicología cognitiva) se hagan cargo de ellos. Por otra parte, esta afirmación es claramente diferente de (a), de la posición según la cual *S* puede conocer sin que *nadie* advierta la justificación. La óptica de la tercera persona *supone* después de todo un punto de vista: la idea es, justamente, que supone el punto de vista de cualquiera, de sujetos en principio intercambiables; la "objetividad" constituida por esta perspectiva, la objetividad requerida por la ciencia, no es sino intersubjetividad.

Si lo anterior es correcto, tal vez ahora se comprenda por qué el externalismo así entendido ES conciliable con una posición según la cual para conocer basta con poseer un buen criterio de verdad (a diferencia de, por ejemplo, el requisito de contar con proposiciones *efectivamente* verdadera). Decir que el conocimiento es un proceso objetivo, o que los procesos confiables son objetivos, equivale, en esta interpretación, a la afirmación de que son *intersubjetivos*. Es suficiente, entonces, con identificar procesos que *creemos* (intersubjetivamente) confiables; correlativamente, basta con disponer de proposiciones que tomamos (intersubjetivamente) por verdaderas. Afirmar que *P* es un proceso objetivo, entonces, no significa que sea independiente de *toda* comprensión de su justificación, sino que es independiente de cualquier comprensión por parte de un sujeto en particular.

Vale la pena recordar algunas observaciones vertidas en la Introducción de esta Tesis. Decíamos allí que el confiabilista que quería enfrentar al escéptico debía ofrecer *ejemplos* de conocimiento, ejemplos de creencias

---

94. Véase también Davidson (1984). Para Davidson significados y creencias son constituidos por el sujeto que interpreta, en la situación de "interpretación radical". El observador en tercera persona es condición *necesaria* para poder hablar de deseos, creencias y significados. Lo novedoso de la posición de Davidson, entre otras cosas, radica justamente en el hecho de que los significados no preexisten a la situación de interpretación, que no es entonces meramente *atributiva*, sino *constitutiva* de los mismos.

verdaderas adquiridas a través del procedimiento adecuado. El punto en discusión era si las características externalistas del confiabilismo no lo eximían de hacer semejante cosa. En una primera aproximación al tema ya habíamos respondido negativamente. Ahora disponemos de razones adicionales para justificar esta negativa: el externalismo bien entendido en ningún momento puede desentenderse de la *identificación* y explicitación de casos concretos<sup>95</sup>.

## II)

Ahora bien, ¿qué criterios de aceptación de creencias pueden ser reivindicados por el confiabilismo? Dretske mismo nos sugiere la respuesta, en un artículo de 1989. Allí esboza lo que podríamos considerar un criterio pragmatista de verdad, estrechamente ligado con la defensa de tesis evolucionistas.

Conocimiento es "creencia verdadera producida de modo confiable" (Dretske (1989), p.92). (Dretske aquí deja en claro que está ofreciendo una teoría confiabilista *del conocimiento*, y no de la justificación<sup>96</sup>). El conocimiento, además, es una condición necesaria para la supervivencia, de donde el argumento más inmediato que puede sugerirse contra el escéptico consiste en la observación obvia de que hemos sobrevivido, y por ende,

---

<sup>95</sup>. Por supuesto, también hay críticas clásicas al externalismo que se aplican a (b). Supóngase, por ejemplo, que, en sentido objetivo (o intersubjetivo), *S* sabe que *p*. Pero supóngase también que *S* mismo tiene la firme creencia.... de que su creencia en *p* es injustificada. ¿Diríamos en este caso, intuitivamente, que *S* conoce? Para hacer frente a casos similares a éste Kent Bach propone la interesante sugerencia de distinguir entre la justificación de un sujeto para sostener una creencia y el hecho de que la creencia misma esté justificada (Bach (1985), p. 251 y siguientes, el apartado "Creencias justificadas y creyentes justificados"). Kornblith (1985) propone para este mismo problema una suerte de conciliación entre puntos de vista externalistas e internalistas.

<sup>96</sup>. Véase por ejemplo (1989), nota 2, p.100. Dretske aclara que su opción por el confiabilismo excluye la exigencia de una condición de justificación. "Justificación" es entendido por Dretske de un modo internalista, como el autorreconocimiento de que una creencia constituye de hecho conocimiento. Para Dretske la justificación así entendida es importante sólo en tanto lo contrario nos impida actuar de un modo, digamos, exitoso desde un punto de vista biológico. "Los críticos del externalismo, aquellos que piensan que el conocimiento requiere algo *diferente de*, o por lo menos *más que*, creencia verdadera producida de modo confiable, algo (usualmente) en la forma de una justificación para la creencia de que nuestras creencias producidas de modo confiable *son* producidas de modo confiable, tiene, me parece a mí, la obligación de decir qué beneficios se supone que confiere esta justificación. (...) [la justificación] puede ser de utilidad en tanto afecte la resolución de la persona para *actuar* con base en la creencia. Pero si no la afecta (...) como creo que es claro que no lo hace en el caso de la mayoría de las creencias perceptuales (los casos paradigmáticos de conocimiento), ¿cuán valiosa puede ser la justificación?" (pp.95-96).

conocemos; luego, el conocimiento es posible<sup>97</sup>. ¿Pero qué hay del escéptico que pretenda negar el condicional según el cual "si no tuviéramos procesos representacionales confiables (esto es, conocimiento) no sobreviviríamos"? Dretske considera que ni siquiera vale la pena tomarse en serio este tipo de objeciones, y sugiere una respuesta, creo, bastante ilustrativa de su modo de argumentar: si no hay razones para dudar, no debe dudarse; la mera posibilidad de la duda no transforma a dicha duda en legítima<sup>98</sup>.

Dretske sí se ocupa, en cambio, de tres posibles críticas a su concepción, que aquí sólo mencionaré. (1) Podríamos admitir que es necesario contar con mecanismos representacionales confiables (con lo cual, la selección natural favorecerá tales mecanismos) y negar, sin embargo, que nuestras *creencias* sean productos de tales mecanismos representacionales. Dretske argumenta que la creencia es un caso muy especial de representación interna, y que, en tanto tal, *si* hay presión selectiva sobre la existencia de mecanismos representacionales fiables, entonces también la hay sobre las creencias verdaderas. (2) Por otra parte, tal vez sea suficiente la existencia de presión selectiva para tener representaciones internas *verdaderas*, pero no sea necesario exigir presión selectiva para la existencia de mecanismos confiables de conocimiento. La respuesta a esta objeción es sumamente interesante. Es cierto, dice Dretske, que la propiedad que interesa es la *verdad*, no la confiabilidad. Pero, la mayoría de las creencias importantes para la supervivencia son creencias "indexicales" (*indexical beliefs*), o representaciones perceptuales sobre el medio, y este tipo de creencias *no son heredables*. Vale decir que el único modo de seleccionar verdad es seleccionando confiabilidad. "No hay manera de dar a un individuo lo que se busca, creencias verdaderas sobre eventos futuros, sin darle *más* de lo que es buscado. Así es la verdad, al menos el tipo de verdad en este momento en cuestión. Es algo que no se puede comprar. Lo único en venta es un medio, un proceso confiable, para producir creencias verdaderas, un medio que, cuando se despliega, produce no meramente creencias verdaderas, sino *conocimiento*."

---

97. Véase un comentario a la estrategia de Dretske (1989) en E. de Olaso (1994): " 'Saber' o 'conocer' es aquí la adaptación exitosa al medio. Esos procesos de adaptación muestran que el agente (o su especie) ha sobrevivido. Si ha sobrevivido se concluye que conoce. Y así se llega a la conclusión de que el escéptico (es decir, el que sostiene que no conocemos) está equivocado porque conoce. 'Existe, por lo tanto conoce', sería la divisa del fiabilista." (pp.156-157).

98. Véase Dretske (1989), nota 1. (p.100).

(p.94). (3) Finalmente, la última posible objeción señala que aún si admitimos que hay presión selectiva para tales mecanismos confiables, ¿qué grado de confiabilidad es *suficiente* para hablar de conocimiento? ¿qué nos hace suponer que la evolución ya alcanzó el grado de confiabilidad requerido para el conocimiento? Según Dretske responder estas preguntas involucra tomar decisiones en temas de semántica y filosofía de la mente, en especial el problema de qué es lo que determina el contenido o significado de aquellos estados internos que describimos como creencias. No comentaré aquí la solución de Dretske a este punto, pero tendré en cuenta la cuestión para la próxima sección del capítulo.

De lo anterior se sigue que, una vez más, parece darse por sentado que *tenemos* creencias verdaderas. Pero en este marco creo que tal presupuesto puede justificarse como sigue. Es importante notar que de la idea de que "puesto que sobrevivimos, tenemos conocimiento" es posible derivar la afirmación "puesto que sobrevivimos, tenemos creencias *verdaderas*". "Creencia verdadera" es equivalente a "estructura representacional que da cuenta fielmente del medio". ¿Cómo saber si representa fielmente el medio? Bueno, si permite al organismo sobrevivir, entonces lo hace. He aquí un *criterio de aceptación de creencias como verdaderas*: una creencia puede ser considerada verdadera si se nos aparece como una estructura interna útil para la supervivencia. Esta formulación, inclusive, es compatible con una definición correspondentista, realista, de la verdad; por ejemplo, es verdadera una creencia *porque* representa del medio un estado de cosas real. (Evidentemente, también es compatible con una concepción antirrealista, por ejemplo, si la verdad de una creencia *consistiera* en su "utilidad para nuestra supervivencia". En este último caso el mencionado pragmatismo no sería tan sólo criterio, sino lo que la verdad de una creencia ES. Me ocuparé de ello más abajo).

Sería posible identificar otros criterios de aceptación de creencias como verdaderas, desde luego. (Por ejemplo, algún tipo de criterio coherentista: reconozco como verdadera a toda creencia que resulte necesaria para mantener la coherencia sistemática de mi conjunto global de creencias, etc.). Pero, dada la impronta naturalista del confiabilismo, creo que un criterio pragmático como el mencionado hace justicia a las aspiraciones básicas de este tipo de teorías.

## 2. La teoría informacional del conocimiento y el antirrealismo semántico.

En lo que sigue sólo esbozaré una estrategia alternativa a la ya comentada para eliminar los problemas (en particular, los problemas escépticos) que causa al confiabilista el uso de una concepción realista de la verdad. No pretendo en este punto llegar a resultados concluyentes, sino, simplemente, sugerir una vía de investigación.

Mi sugerencia consiste en complementar al confiabilismo con las tesis de lo que en el capítulo uno he llamado antirrealismo semántico. Creo que existen ciertas intuiciones pre-teóricas que de algún modo avalan esta propuesta. Me refiero a lo siguiente:

I)

Consideremos proposiciones como "el sol está compuesto mayormente por helio"; "la civilización azteca no conoció la rueda"; "ciertos virus cristalizan". En ninguno de estos casos supone un real problema la posibilidad de que, si bien hoy las creemos verdaderas, mañana tal vez aseguremos que son falsas, y que sólo erróneamente las creíamos verdaderas. En este sentido, *no hay en realidad problema del escepticismo concerniente a este tipo de enunciados*.

Consideremos ahora, en cambio, proposiciones como "existe un mundo exterior"; "existen otros seres humanos"; "esto es mi mano"; "el mundo existe desde hace bastante más que cinco minutos"; "no somos cerebros en una cubeta"; "yo no nací la semana pasada". A mi modo de ver, todas estas proposiciones tienen algo en común. Lo que ocurre con ellas es que, intuitivamente, sentimos aversión por la posibilidad de decir que son *falsas*. (Creo que esto es algo diferente de, por ejemplo, decir no tiene *sentido* siquiera plantearse si son verdaderas o falsas; encuentro arbitraria, por otra parte, la suposición de que en realidad no puedo decir que las *conozco*, por ejemplo, porque se trata de certezas no cognoscitivas). Es el tipo de proposiciones de las cuales, ante la pregunta *¿son verdaderas, o tal vez no lo son?*, nos inclinamos inmediatamente por una respuesta afirmativa. Es decir que intuitivamente rechazamos la idea de que tan sólo las *creemos* como verdaderas: *no somos*

*realistas semánticos respecto de ellas*. Ello revela, de paso, por qué nos molestan los argumentos escépticos que conciernen a este tipo de enunciados: el escéptico es el que insiste en que no tenemos derecho a *asegurar* que son verdaderas, sino tan sólo que lo creemos así; *el escéptico es el que presupone una concepción realista de la verdad, indiscriminadamente, para todo tipo de proposiciones.*<sup>99</sup>

Hay también otro tipo de afirmaciones, del tipo "lo que veo allí es un perro de carne y hueso". Tal vez, después de todo, me equivoque; tal vez no sea un perro sino un holograma tan bien hecho que no noté la diferencia. Sin embargo, la gran mayoría de las veces, digamos inclusive la *inmensa* mayoría de las veces, también sentimos aversión por la posibilidad de que nuestra proposición sea falsa. Lo que ocurre en este último caso respecto del realismo/antirrealismo semántico es un poco más delicado. Pero también, creo, en algún sentido está presente la idea de que, en *condiciones normales* (que habrá que especificar en cada caso), tales proposiciones *no pueden sino ser verdaderas*.

Pero, como hemos visto en el segundo capítulo, no cualquier tipo de antirrealismo semántico es útil para dar cuenta de las comentadas intuiciones contra el escepticismo. Ninguna concepción antirrealista que involucre nociones idealizadas, por ejemplo, podrá funcionar.

¿Qué tipo de concepción servirá, entonces? Se me ocurre, tal vez, que debería intentarse con alguna suerte de convencionalismo. Es decir: si comprendemos *el significado de los términos de nuestro lenguaje*, no puede sino ser verdadero que existe un mundo exterior, que en este momento estoy escribiendo frente a una computadora, que no soy un cerebro en una cubeta, que, en condiciones normales, eso que veo allí y que creo que es una mesa es una mesa. La verdad de ciertos enunciados es así una decisión (más o menos implícita, relacionada con los usos habituales del lenguaje) y, en ese sentido, no son falsables.

Recordemos además el criterio pragmatista de aceptabilidad de creencias (como verdaderas) comentado en la sección anterior. También podría considerarse, desde luego, una concepción pragmática *de lo que la verdad es*, aplicado al tipo de enunciados problemáticos que he señalado más arriba. Existe

---

<sup>99</sup>. No estoy suponiendo necesariamente aquí que el escéptico tenga convicciones propias sobre el conocimiento o la verdad (véase nota 61). Lo que quiero expresar podría reinterpretarse debidamente como "el escéptico es el que, equivocadamente, *supone* que todo aquel que defienda la posibilidad del conocimiento se debe ver complicado con una concepción realista de la verdad, indiscriminadamente, para todo tipo de proposiciones".

un modo relativamente interesante de incluir los criterios de aceptabilidad disponibles en una concepción epistémica de la verdad<sup>100</sup>. Así podría establecerse que, por ejemplo, *p* es verdadera, por definición, si todos los criterios de que disponemos intersubjetivamente (pragmatismo, coherentismo, etc.) permiten *evaluarla* como verdadera. Inclusive puede aquí aparecer la idea de *correspondencia* : una proposición es verdadera si (intersubjetivamente) *evaluamos* que se corresponde con los hechos<sup>101</sup>. También en este caso la verdad de ciertos enunciados consiste en una decisión: la decisión de que la verdad de ciertos enunciados queda determinada por nuestros criterios de aceptabilidad. Podría decirse, entonces, que, dado el modo en que usamos el lenguaje, para ciertas proposiciones se cumple que si todos los criterios intersubjetivamente aceptados las identifican como verdaderas, entonces *son* verdaderas. Debe notarse sin embargo que no es posible aludir, por ejemplo, a "todos los criterios de que disponemos en condiciones óptimas de investigación", pues, como ya hemos visto, una formulación de este tipo no da cuenta de las intuiciones que queremos rescatar en este momento.

## II)

Entiendo que la *teoría informacional del conocimiento es conciliable con este planteo*. En particular, algunas de las tesis de Dretske en relación con problemas de filosofía de la mente parecen, si uno las examina con cuidado, revelar algunos rasgos de antirrealismo semántico (combinados, claramente, con una posición metafísicamente realista). Recordemos ciertas consideraciones sobre lo que son las creencias. Las creencias son estructuras con cierto contenido semántico que ha sido originado a partir de una relación informacional. ((1981), cap. 8). Ahora bien, esta caracterización parece impedir que haya creencias con contenido falso. Del contenido semántico de una estructura no tiene sentido decir semejante cosa, pues ha sido determinada causalmente por una relación informacional, y por definición no hay

---

<sup>100</sup>. Una estrategia de este tipo es utilizada por ejemplo por Apel (véase Apel (1991)), y ya mencionada en el Capítulo dos en el marco de una concepción de la verdad como consenso idealizado. (Apel piensa, dijimos, basándose en ideas de Peirce, que el concepto de verdad utilizado por las ciencias naturales es una noción ideal: consiste en el consenso de una comunidad científica ideal e ilimitada sobre todos los criterios de verdad disponibles).

<sup>101</sup>. Véase el Capítulo uno de esta Tesis, la sección "Compatibilidades del antirrealismo semántico", para comentarios sobre la relación entre el antirrealismo semántico y la noción de correspondencia.

transmisión de información "falsa". Se dirá que la verdad ya está presupuesta en este primer paso: lo que sí es posible es que no haya transmisión con equivocidad 0, con lo cual lo que puede ocurrir es que no haya en realidad transmisión de información que pueda causar el contenido semántico de nuestra creencia. Sin embargo, ¿cómo puede ocurrir algo semejante? *De hecho, tenemos creencias, que tienen algún contenido semántico.*

La respuesta de Dretske creo que no hace sino confirmar nuestras sospechas. Distingue entre "estructura tipo" de cierta creencia, y "especímenes", (haciendo un uso totalmente *sui generis* de la noción de "creencia tipo"). Si bien un "tipo" puede *aplicarse erróneamente* a través de ejemplares concretos, del tipo en sí mismo no tiene sentido predicar falsedad. El contenido de una creencia *tipo* es función directa de su origen informacional; los casos se limitan a reidentificar las mismas condiciones que le han dado origen. Si a un individuo se le enseña el concepto de "petirrojo" enfrentándolo a petirrojos y a aves azules (de las cuales se le dice que *no* son petirrojos), y luego observamos que aplica "petirrojo" a gorriones, palomas, y toda ave que no sea de plumaje azul, no deberíamos decir que está aplicando mal el concepto aprendido, sino que ha aprendido algo diferente de lo que presuntamente el maestro pretendía, debido a insuficiencias en el proceso de aprendizaje<sup>102</sup>. En el mismo sentido le sirve a Dretske el ejemplo de la Tierra Gemela de Putnam<sup>103</sup>: el concepto de "agua" de un individuo en la Tierra Gemela no es el mismo concepto que maneja un individuo de nuestra Tierra; "agua" no *significa* lo mismo aquí y allá (paralelamente, según Dretske sus creencias no son las mismas), pues los contenidos semánticos de ambos conceptos difieren, dado que han tenido orígenes informacionales diferentes. Aún más: diferirían

---

102. En Fodor (1987) se presentan una serie de críticas a la posición de Dretske que estoy aquí comentando. Una de ellas consiste en que, según Fodor, toda línea divisoria entre un supuesto período de aprendizaje y un período subsiguiente de aplicación de un concepto, donde es posible el error, es siempre más o menos arbitraria. Tal vez la objeción puede responderse en parte si reconocemos que es efectivamente una tarea por cumplir (por la psicología empírica) la de averiguar empíricamente cuándo un conjunto de relaciones informacionales es necesario y suficiente para la formación de un concepto particular. De todas maneras vale la pena observar que si tal distinción entre un período de aprendizaje y otro de aplicación fuera siempre borrosa por principio, ello no invalidaría la argumentación que estoy llevando a cabo aquí. Es decir, no invalidaría la interpretación de que estos puntos de la construcción de Dretske son conciliables con el antirrealismo semántico.

103. Cf. "The Meaning of 'Meaning'", Putnam 1975; véase también el capítulo anterior de esta Tesis.

aún si en la Tierra Gemela existieran regiones con H<sub>2</sub>O y otras con XYZ por partes iguales, y nuestro Gemelo hubiera vivido en una región exclusivamente de H<sub>2</sub>O: en su conjunto, probabilísticamente, las *regularidades empíricas* de la Tierra Gemela la convierten en una fuente de información diferente de nuestra Tierra.

Los ejemplos del párrafo anterior ilustran bastante bien la situación: no es posible tener creencias-tipo falsas (ello, dicho sea de paso, elimina radicalmente la posibilidad de "error masivo" en nuestras creencias), porque siempre es posible señalar que se ha transmitido *alguna* información<sup>104</sup>; el único problema, en todo caso, es que luego podemos no reconocer cuál es exactamente la información que ha sido transmitida. Pero entonces la verdad de "esto es un petirrojo" es una cuestión lingüística: depende del *significado* de petirrojo, lo que depende a su vez de la situación de aprendizaje. Más arriba habíamos dicho que "en condiciones normales, ciertas proposiciones *no pueden sino ser verdaderas*". Creo que la teoría informacional del conocimiento podría suscribir esta idea, si por "condiciones normales" entendemos "condiciones idénticas (en los aspectos relevantes) a la situación de aprendizaje". Pero esto ES antirrealismo semántico. Y, en este sentido, la teoría informacional del conocimiento es compatible con el antirrealismo semántico.

¿Cómo podría ser falso el enunciado "esta es mi mano derecha"? Sólo si el contexto en el cual aprendí a usar estas palabras fuera radicalmente diferente del actual. Aquí, nuevamente, aparece el problema de la inducción, pero podemos dejarlo de lado por el momento. Pues observemos que el escepticismo radical del que dice "tal vez es cierto que soy un cerebro en una cubeta" *no*

---

104. Sirva como ejemplo el siguiente pasaje de Dretske (1989): "Cuando un murciélago se aproxima, la polilla inicia maniobras de fuga. Hay algo en la polilla, una representación interna de lo que la rodea, que le dice *cuándo* hacer esto. ¿Esta representación interna es la representación de un murciélago, de sus movimientos? ¿O es meramente una representación de las propiedades del input acústico [que acompaña la presencia del murciélago]? En el primer caso, podemos burlar a la polilla con un generador de sonidos de alta frecuencia. Podemos hacerla "pensar" que un murciélago se acerca cuando no hay murciélagos alrededor. Si, por otro lado, la polilla se representa el input acústico, nuestros esfuerzos por engañarla con el generador de sonidos son inútiles. (...) El input acústico tiene exactamente el carácter que la polilla se representa que tiene. La polilla no comete ningún error. (...) La polilla vive en un mundo de vibraciones acústicas. Y bajo esta modesta descripción de los esfuerzos representacionales de la polilla, ella goza de virtual infalibilidad." (p.97). (Por supuesto, Dretske no pretende decir que no haya diferencia relevantes entre nuestro aparato de creencias y el primitivo sistema de representaciones internas de la polilla, pero, de todas maneras, creo que la analogía es ilustrativa).

*depende de la inducción* ; simplemente, lo que afirma es: "siempre creí ver manos, mesas, petirrojos, etc; en especial, siempre creí que ésta era mi mano derecha, pero tal vez nada de ello ocurría *en la realidad* , porque tal vez no tengo cuerpo y soy sólo un cerebro". La teoría informacional del conocimiento, en combinación con algunas tesis semánticas sobre la formación de significados y creencias, permite rechazar la hipótesis escéptica con el siguiente argumento: nuestros conceptos y creencias son el fruto de nuestro contacto con el medio, *sea la que fuere la situación real que les dio origen* , y, en tanto tal, la aplicación de esos mismos conceptos en condiciones semejantes no puede sino resultar en enunciados verdaderos. La circunstancia misma de hacer uso de un lenguaje *provoca* que ciertas oraciones sean verdaderas. Las hipótesis del escéptico parecen preguntarnos por lo que está más allá del lenguaje. Pero ello es un grueso *error del escéptico* , en el mismo sentido en el que, en la primera parte de este capítulo, decidimos que era un error del escéptico pedir lo incondicionado, pedir por la información existente más allá de toda relativización a un canal de comunicación.

He aquí, por otra parte, una explicación del por qué de la fascinación que ejercen los argumentos escépticos: nos arrastran hasta los límites no ya de lo humanamente cognoscible, sino de lo humanamente *pensable* , y nos invitan a ir más allá. ¿No será el universo un computador gigante, y nosotros cerebros conectados a él? Dado el significado habitual de nuestras palabras, es obvio que esta proposición es falsa. La estrategia del escéptico, entonces, me parece que en el fondo consiste en decir: "bueno, ¡olvídese del significado habitual de las palabras! Yo le propongo que *piense* como son las cosas 'realmente', con independencia del uso de un lenguaje". Pero esto es imposible, por supuesto. Es un tipo de sugerencia afín a la de "piense por qué hay algo y no más bien nada". Tales propuestas aseguran cierto escalofrío, porque decididamente no pueden llevarse a cabo. Pero no constituyen una amenaza para lo que modestamente suele entenderse por "conocimiento".

### III)

Creo que es importante notar que el confiabilismo es compatible con el antirrealismo semántico sólo si aseguramos primero que el antirrealismo semántico no es inconciliable con una posición metafísicamente realista. Esto es, en efecto, lo que se ha intentado mostrar en el Capítulo uno. La teoría informacional del conocimiento, como sabemos, asegura la existencia de fuentes de información objetivas. *Nada* de lo que ocurra en la fuente, por

hipótesis, depende del sujeto receptor de la información. Inclusive cuando, en el Capítulo cuatro, sugerí que debían tomarse como posibilidades informativas de la fuente aquellas que el sujeto evaluara como tales, la única cosa que tal estrategia consigue relativizar es la *cantidad* de información recibida, pero ello no afecta al *contenido* de la información. Y cuando Dretske mismo propone un medio para relativizar el contenido, al concurriendo previo del sujeto y al modo de codificación, nada de ello afecta *lo que sucede en la fuente*; sólo afecta la índole del contenido que el sujeto es capaz de asimilar, *dado* lo que ocurre o deja de ocurrir (independientemente de él) en la situación generadora de información.

Estas mismas observaciones son las que, en la sección anterior de este capítulo, nos llevaron a interpretar a la teoría informacional de un modo semánticamente realista. Habíamos razonado de este modo: si la información generada en una fuente es independiente de consideraciones subjetivas, y la verdad de un enunciado consiste en la descripción de la información generada en una fuente, entonces la verdad de un enunciado es independiente de consideraciones subjetivas. Este, evidentemente, ES un modo posible de concebir a la teoría, pero, como espero haber mostrado, no es el único. La clave para una interpretación diferente consiste en privilegiar la idea de que la verdad es un fenómeno semántico, y como tal supone el uso de convenciones lingüísticas. (Por supuesto, ningún realista niega esto último; se trata sólo de una cuestión de énfasis). Entonces, aún si la verdad de una oración es una descripción de la información generada en una fuente (independiente) hay que tener cierto cuidado al considerar qué se quiere decir con "describir la información generada en una fuente". Tener éxito en el intento de realizar tal descripción, o fallar, es una cuestión de uso del lenguaje. Frente a esto, la reacción típica de un realista semántico consiste en decir que la *razón* por la cual la oración es verdadera es que existe (en la realidad extra-lingüística) una fuente que genera exactamente la información descrita por la oración. El antirrealista semántico no niega (necesariamente) que exista tal fuente independiente: sólo niega que tal existencia sea el factor *principal* para explicar la verdad de una oración<sup>105</sup>.

---

<sup>105</sup>. Por lo menos alguna versión de antirrealismo semántico fundamenta esta negativa en la idea de raíz quineana de que no es posible identificar relaciones causales únicas entre los términos lingüísticos y la referencia en el mundo. Véase Capítulo uno.

#### IV)

Unos párrafos más arriba volvió a aparecer el problema de la inducción. En el Capítulo cinco sugerí una estrategia para evitarlo: dar un análisis de "conocer" en términos de "conocer-en -C". ¿Es posible tratar conjuntamente la idea de "conocimiento-en-C" y de una teoría informacional semánticamente antirrealista? En su momento decidimos aceptar sin cuestionar la idea de que contábamos con algunas creencias verdaderas como punto de partida para evaluar la confiabilidad de un canal. El segundo paso podría consistir, dijimos, en suspender la duda sobre el comportamiento futuro de las condiciones del canal. Así llegamos a "conocer-en-C". Admitamos ahora, además, que las creencias que considerábamos verdaderas lo son de acuerdo con una concepción no realista de la verdad. Lo que resulta de ello es que cada nueva generación de creencias a través de C será, muy probablemente, *verdadera*, donde la expresión "muy probablemente" no alude a la imposibilidad de asegurar si una creencia dada es verdadera o no, sino al hecho de que las condiciones de transmisión *la gran mayoría de las veces* (pero tal vez no todas) serán no informativas. (Recordemos que la relativización a un canal no elimina toda posibilidad de falsedad; sólo elimina la posibilidad de que las condiciones de transmisión de información cambien radicalmente la proporción de generación de verdades sobre la generación de falsedades). Así es que la combinación de confiabilismo con antirrealismo semántico permite, por un lado, seguir reivindicando una concepción de la justificación que admita el falibilismo ("justificado" es el tipo de adjetivo que podría llegar a aplicarse a una creencia falsa, *cf. Goldman (1979)*<sup>106</sup>); por otro lado, asegura que nuestra evaluación de la verdad de ciertas creencias no puede estar equivocada.

#### V)

Reconozco que la conciliación que he sugerido tiene algunos problemas. Me limitaré a señalarlos, pero no intentaré dar soluciones. En primer lugar, tal vez resulte extraña la idea de que solamente a *algunos* tipos de proposiciones se aplica una concepción semánticamente antirrealista de la verdad como la que aquí se ha propuesto. ¿Ello significa que contamos con más de una noción de verdad simultáneamente? Aún si esta sugerencia resultara no problemática,

---

<sup>106</sup>. O tal vez: "confiabilidad" es el tipo de requisito *alternativo* al de "justificación", que permite garantizar aún alguna que otra creencia falsa.

resta una segunda observación, mucho más importante: ¿dónde trazar la línea de demarcación entre uno y otro tipos de enunciados?

Confío en que no sea imposible dar una respuesta a las cuestiones anteriores y a otras semejantes<sup>107</sup>. Por el momento, mi intención ha consistido simplemente en llamar la atención sobre los siguientes puntos:

- Una respuesta al escepticismo debe hacerse cargo del potencial escéptico de una noción realista de verdad.
- Contamos con ciertas intuiciones pre-teóricas que parecen autorizarnos para admitir cierto tipo (no idealizado) de verdad antirrealista respecto de algunos enunciados.
- La teoría informacional del conocimiento es genuinamente compatible con el antirrealismo semántico.

La teoría informacional del conocimiento también sería compatible, según las conclusiones de la sección anterior de este capítulo, con la idea de que la verdad no es, después de todo, un requisito del conocimiento. ¿Cuál de las dos estrategias resulta preferible? Toda evaluación en este punto, me parece, resultará algo arbitraria, aunque tal vez en favor del antirrealismo semántico pueda esgrimirse lo siguiente: el antirrealismo semántico favorece un diagnóstico sobre el escepticismo que, entre otras cosas, permite dar cuenta del atractivo de los argumentos escépticos. Y ésta, según sugerí en la Introducción de esta Tesis, era una de las propiedades que debía tener una solución buena frente al escepticismo.

---

<sup>107</sup>. Tal vez la clave para ello resida en recordar la importancia dada por el confiabilismo, y por Dretske en particular, al *conocimiento perceptivo*, y a todas aquellas creencias que sean directamente necesarias para la supervivencia. Esta idea posiblemente debiera ser tenida en cuenta a la hora de proponer algún criterio para demarcar tipos de enunciados o creencias.

## CONCLUSIONES

En este trabajo me ocupé de la posibilidad del conocimiento empírico y del desafío que suponen para ello los argumentos escépticos. Presenté estas cuestiones en el marco de la perspectiva gnoseológica del confiabilismo; en particular, analicé una "teoría informacional del conocimiento", la cual constituye un tipo especial de teoría confiabilista. Tal teoría ofrece la posibilidad de *cuantificar* información (si bien las fórmulas que arrojan datos cuantitativos no *siempre* podrán aplicarse a los casos concretos). Por otra parte, las herramientas de la teoría matemática de la comunicación que están a la base de la teoría informacional no solamente permiten dar cuenta de cuestiones gnoseológicas, sino también de otros fenómenos relacionados, como del problema de la significación, o del contenido mental; esta unificación explicativa supone, por cierto, una ventaja de la teoría informacional por sobre otros tipos de análisis gnoseológicos.

He intentado mostrar que las estrategias confiabilistas frente al escéptico cuentan con dos grandes presupuestos: (1) que es posible hacer un uso no problemático de la inducción; (2) que estamos en posesión de creencias/proposiciones verdaderas, es decir que es posible evaluar la verdad de una creencia/proposición sin mayores dificultades. Respecto del primer supuesto, no me he ocupado de analizar modos de justificación de la inducción, sino que, basándome en algunas ideas de Dretske, sugerí una forma de *evitar* el problema: relativizar la noción de conocimiento a "conocer-en-C", esto es, conocer a través de cierto canal de comunicación, que *por hipótesis* considero será no informativo la gran mayoría de las veces (es decir, la gran mayoría de las veces la equivocidad será igual a cero, y por ende no afectará la transmisión de la información generada en la fuente). En este punto propuse un primer diagnóstico sobre el escepticismo: el error del escéptico consiste en pedir lo incondicionado, pedir información no relativizada a ningún canal.

Respecto del segundo supuesto, asegurar que tenemos ciertas creencias verdaderas implica, en primer lugar, que es posible contar con un *criterio* (o criterios) de verdad satisfactorios. Pero esto sólo no es suficiente, sino que son necesarias algunas aclaraciones. Una concepción realista de la verdad, por ejemplo, sostiene que la verdad de un enunciado es independiente de cuestiones epistémicas. De ello se infiere que siempre, por principio, podría

ocurrir que lo que *creemos* verdadero en realidad no lo sea. Frente a este problema he identificado dos formas alternativas de proceder: (a) podemos considerar el abandono de la verdad como requisito indispensable para el conocimiento. Para conocer no es necesario *asegurar* que contamos con proposiciones verdaderas, sino que sería suficiente con disponer de *criterios* de aceptación de creencias; (b) podemos adoptar una concepción no realista de la verdad, por lo menos en lo que concierne a ciertas proposiciones. Ninguna de las dos opciones es reivindicada explícitamente por ningún confiabilista. Sin embargo, he argumentado que ambas son compatibles con el confiabilismo. Para asegurar tal compatibilidad ha sido necesario, en el primer caso, proceder a una reinterpretación del externalismo. Y, en el segundo caso, ha sido necesario mostrar primero que el antirrealismo semántico no necesariamente tiene implicaciones idealistas, es decir que es posible conciliar antirrealismo semántico con una posición realista en terreno metafísico.

Toda elección entre ambos tipos de respuesta al problema de la verdad tiene posiblemente algo de arbitrario. Sin embargo, podría intentar justificar mi preferencia por el antirrealismo semántico sobre la base de ciertas intuiciones respecto de la verdad de algunos enunciados: de enunciados del tipo "no somos cerebros en una cubeta" o "yo no nací ayer" no pensamos que sólo los *creemos* verdaderos; antes bien, tenemos la fuerte intuición de que *no pueden ser falsos*. Ello permite, de paso, otro diagnóstico (complementario con el anterior) sobre cuál es el error del escéptico: el escéptico comete la equivocación de presuponer una concepción realista de la verdad, indiscriminadamente, para todo tipo de proposiciones.

Es importante notar que, sin las modificaciones propuestas, la teoría informacional del conocimiento fracasa a la hora de dar una solución al escepticismo. Fracasa porque:

- Reubica la duda escéptica en el contexto del escepticismo respecto de la inducción, y no ofrece argumentos para refutar a este último.
- No encara el problema del potencial escéptico inherente a una concepción realista de la verdad. Al menos en una primera lectura, no hay razones que nos impidan atribuir al confiabilista una posición semánticamente realista.

Sin embargo, si las observaciones de esta Tesis son correctas tales problemas son superables, y la respuesta confiabilista puede ser considerada satisfactoria. Lo interesante de esta respuesta, según creo, reside en su diagnóstico sobre cuáles son los errores del escepticismo. Ello nos habilita para

comprender a qué se debe el innegable atractivo de los argumentos del escéptico. Los argumentos escépticos presuponen que para conocer son necesarios ciertos requisitos, que luego resultan ser imposibles de satisfacer *por principio* . Pero al pedir tales requisitos imposibles el escéptico nos enfrenta con los límites de lo humanamente pensable, con el vértigo de lo que ni siquiera podemos concebir.

Los resultados alcanzados por este trabajo revelan, según creo, que la teoría informacional del conocimiento es una concepción gnoseológica con buenas perspectivas. Tiene posibilidades de manejar datos cuantitativos, utiliza un marco explicativo que permite dar cuenta de otros problemas filosóficos, se compromete con algunos puntos típicamente externalista y naturalistas, y, finalmente, permite elaborar una respuesta frente a la amenaza del escepticismo. Es cierto que no he abordado otros problemas específicos de la teoría, sino sólo aquellos directamente relacionados con el último punto. Sin embargo, algunas de sus características más relevantes permiten suponer que vale la pena continuar trabajando en ella.

## BIBLIOGRAFIA:

Siempre que se indica la edición castellana, las citas del texto corresponden a esa traducción.

- APEL, Karl-Otto, (1991), *Teoría de la verdad y ética del discurso*, Barcelona-Buenos Aires- México, Paidós.
- ALSTON, William P., (1985), "Concepts of Epistemic Justification", *The Monist*, Vol. 68, Nro 1, pp.57-89.
- ARMSTRONG, David, (1973), *Belief, Truth and Knowledge*, Cambridge, Cambridge University Press.
- ARMSTRONG, David M., (1978), *Universals and Scientific Realism*, Cambridge, Cambridge University Press.
- AUDI, Robert, (1988), *Belief, Justification, and Knowledge*, Belmont, California, Wadsworth.
- BACH, Kent, (1985), "A Rationale for Reliabilism", en *The Monist*, Vol. 68, Nro. 2, pp. 246-263.
- BAR-HILLEL, Y. y otros (comp.), (1978), trad. cast. (1983): *El pensamiento científico*, Madrid, Tecnos-Unesco.
- CHISHOLM, Roderick, (1973), *The problem of the Criterion*, The Aquinas Lecture 1973, Milwaukee, Marquette University Press.
- CHISHOLM, Roderick, (1977), *Theory of Knowledge*, Englewood Cliffs, Prentice-Hall, trad. cast: (1982), *Teoría del Conocimiento*, Madrid, Tecnos.
- DANCY, Jonathan, (1985), *An Introduction to Contemporary Epistemology*, Oxford, Basil Blackwell, trad. cast: *Introducción a la epistemología contemporánea*, Madrid, Tecnos, 1993.
- DANCY, Jonathan y SOSA, Ernest, (eds.), (1994), *A Companion to Epistemology*, Oxford, Blackwell.
- DAVIDSON, Donald, (1984), *Inquires into Truth and Interpretation*, Oxford, Oxford University Press.
- DENNETT, Daniel C., (1987), *The Intentional Stance*, Cambridge, Massachusetts, The MIT Press.
- DEVITT, Michael, (1984), *Realism and Truth*, Oxford, Basil Blackwell.
- DRETSKE, Fred, (1971), "Conclusive Reasons", *Australasian Journal of Philosophy*, 49, pp. 1-22

- DRETSKE, Fred, (1974), "Epistemic Operators", *The Journal of Philosophy*,  
Nro 24.
- DRETSKE, Fred, (1981), *Knowledge & The Flow of Information*, Cambridge, Mass,  
The MIT Press, trad. cast., (1987), *Conocimiento e información*,  
Barcelona, Salvat.
- DRETSKE, Fred y ENÇ, Berent, (1984), "Causal "Theories of Knowledge" en  
*Midwest Studies in Philosophy, Vol.IX, Causation and Causal  
Theories* , Minneapolis, University of Minnesota Press, pp.517-528.
- DRETSKE, Fred, (1989), "The Need to Know", en M. CLAY y K. LEHRER,  
*Knowledge and Skepticism* , Boulder, Colorado, Westview Press,  
pp.89-100.
- DUMMETT, Michael, (1978), *Truth and Other Enigmas* , London, Gerald  
Ducnwoth, trad.cast., (1990), *La verdad y otros enigmas*, México,  
Fondo de Cultura Económica.
- FELDMAN, Richard, (1985), "Reliability and Justification", *The Monist*, Vol. 68,  
Nro 2, pp.159-174.
- FIELD, Hartry, (1972), "Tarski's Theory of Truth", en *The Journal of  
Philosophy*, Vol.LXIX, No13, Julio 1972, pp.347-375.
- FODOR, Jerry, (1987), *Psychosemantics* , Cambridge Mass, The MIT Press.
- FOLEY, Richard, (1985), "What's Wrong with Reliabilism?", en *The Monist* ,  
Vol.68, Nro 2, pp. 188-202.
- GETTIER, Edmund, (1963), "Is Justified True Belief Knowledge?", trad. cast: "¿Es  
conocimiento la creencia verdadera justificada?", en PHILLIPS  
GRIFFITHS, A., trad. cast: (1974), *Conocimiento y creencia* , México,  
Fondo de Cultura Económica, pp.221-224.
- GINET, Carl, (1985), "Contra Reliabilism", *The Monist*, Vol. 68, Nro 2, pp. 175-  
187.
- GOLDMAN, Alvin, (1967), "A causal theory of knowing", *The Journal of  
Philosophy*, 1967, Vol.64, Nro 12, pp.355-372.
- GOLDMAN, Alvin, (1976), "Discrimination and Perceptual Knowledge", *The  
Journal of Philosophy*, 1976, Vol. 73, Nro 20, pp. 771-791.
- GOLDMAN, Alvin, (1979), "What is Justified Belief?", PAPPAS, George S., (1979),  
*Justification and Knowledge* , Dordrecht, Holland, Reidel  
Publishing Company.
- GOLDMAN, Alvin, (1986), *Epistemology and Cognition*, Cambridge, Harvard  
University Press.
- HAACK, Susan, (1980), *Lógica Divergente*, Madrid, Paraninfo.

- HAACK, Susan, (1993), *Evidence and Inquiry - Towards Reconstruction in Epistemology*, Oxford, Basil Blackwell.
- HINTIKKA, Jaakko, (1973), *Logic, Language-Games and information*, Oxford, The Clarendon Press; trad. cast. (1976): *Lógica, Juegos de Lenguaje e Información*, Madrid, Tecnos.
- HORWICH, Paul, (1982), "Three forms of Realism", en *Synthese* 51, 1982, pp.181-201.
- HORWICH, Paul, (1990), *Truth*, Oxford, Blackwell.
- KIRKHAM, Richard, (1984), "Does the Gettier Problem Rest on a Mistake?", *Mind*, XCIII, 1984, pp.501-513.
- KIRKHAM, Richard, (1992), *Theories of Truth. A critical Introduction*, Cambridge, The MIT Press.
- KORNBLITH, Hilary, (1985), "Ever Since Descartes", *The Monist*, Vol. 68, Nro 2, pp. 264-276.
- LEEDS, Stephen, (1978), "Theories of Reference and Truth", en *Erkenntnis*, 1978, pp.111-129.
- LIZ GUTIERREZ, A.Manuel, (1988), "Las semánticas del realismo interno", en *Gavagai*, Vol.I, Nro.2, 1988, pp.21-37.
- LUPER-FOY, Steven (ed.), (1987), *The Possibility of Knowledge- Nozick and His Critics*, Totowa, Rowman & Littlefield.
- McDOWELL, John, (1978), "Physicalism and primitive denotation: Field on Tarski", en *Erkenntnis*, Vol.13, pp.131-152.
- NOZICK, Robert, (1981), *Philosophical Explanations*, Cambridge, Harvard University Press.
- OLASO, Ezequiel de, (1981), "Nota crítica sobre una refutación del escepticismo", *R.L.F.*, Bs. As., Vol.VII, Nro 3, Noviembre 1981.
- OLASO, Ezequiel de, (1994), " 'Yo le aseguro que el hombre no existe' ", en *Dialogo di Filosofia*, Roma, Nro 11, 1994, pp. 151-161.
- OLASO, Ezequiel de, (1995) "Certeza y escepticismo", en *Enciclopedia Iberoamericana de Filosofía, Vol. Teoría del Conocimiento*, Madrid, Trotta, (en prensa).
- PAPINEAU, David, (1992), "Reliabilism, Induction and Scepticism", *The Philosophical Quarterly*, Vol.42, Nro 166, January 1992, pp.1-20.
- PUTNAM, Hilary, (1975), "The Meaning of 'Meaning'", trad.cast. "El significado del 'significado'", en VALDES VILLANUEVA, Luis, (ed.), (1991), *La búsqueda del significado*, Madrid, Tecnos, pp.131-193.

- PUTNAM, Hilary, (1978), *Meaning and the Moral Sciences*, London, Routledge and Kegan Paul, trad.cast., (1991), *El significado y las ciencias morales*, México, UNAM.
- PUTNAM, Hilary, (1978 a), "Significado y Conocimiento (Conferencias John Locke 1976)", en PUTNAM (1991), pp.17-97.
- PUTNAM, Hilary, (1978 b), "Referencia y Comprensión", en PUTNAM (1991), pp.115-137.
- PUTNAM, Hilary, (1978 c), "Realismo y Razón", en PUTNAM (1991), pp.141-158.
- PUTNAM, Hilary, (1981), *Reason, Truth and History*, Cambridge, Cambridge University Press, trad. cast., (1988): *Razón, verdad e historia*, Madrid, Tecnos.
- QUINE, W.V., (1960), *Word and Object*, Cambridge, Mass., The MIT Press, trad. cast. (1962), *Palabra y objeto*, Barcelona, Ariel.
- QUINE, W.V., (1986), *Teorías y cosas*, México, UNAM, (1ra. ed. en inglés, 1981).
- REICHENBACH, (1938), *Experience and Prediction*, Chicago, The University of Chicago Press.
- SCHMITT, Frederick F., (1992), *Knowledge and Belief*, London, Routledge.
- STROUD, Barry, (1984), *The Significance of Philosophical Scepticism*, Oxford, Oxford University Press; trad. castellana: (1990), *El escepticismo filosófico y su significación*, México, Fondo de Cultura Económica.
- TARSKI, Alfred, (1933), trad. inglesa: "The Concept of Truth in Formalized Languages", en TARSKI (1956), (comp.): *Logic, Semantics and Metamathematics*, New York, Oxford.
- TARSKI, Alfred, (1944), trad.cast.: "La concepción semántica de la verdad y los fundamentos de la semántica", en VALDES VILLANUEVA, Luis (ed.), (1991), *La búsqueda del significado*, Madrid, Tecnos, pp. 275-313.
- VILLORO, Luis, (1982), *Creer, saber, conocer*, México, Siglo XXI.
- WITTGENSTEIN, Ludwig, (1969), (ed. bilingüe alemán-castellano, 1988): *Sobre la Certeza*, Barcelona, Gedisa.

UNIVERSIDAD DE BUENOS AIRES  
 FACULTAD DE FILOSOFÍA Y LETRAS  
 Dirección de Bibliotecas