



01083
6
250

**UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO**

FACULTAD DE FILOSOFIA Y LETRAS

**LA EXPLICACION DE LA ACCION
INTENCIONAL:
RAZONES Y CREENCIAS**

FALLA DE ORIGEN

T E S I S

**QUE PARA OBTENER EL GRADO DE:
DOCTORA EN FILOSOFIA**

P R E S E N T A:

SALMA SAAB HASEN

MEXICO, D. F.

1995



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Resumen en inglés de la tesis **LA EXPLICACIÓN DE LAS ACCIONES INTENCIONALES: RAZONES Y CREENCIAS**

SALMA SAAB HASSEN

In this thesis I discuss the problem of the explanations of intentional actions. I consider that they are directed to an end, which is what makes them differ from physical explanations. I divide the thesis in four parts. In the first, I discuss the mental and physical causal relations, in their logical, semantical and epistemological aspects. Many authors consider that the laws play an important role in the causal and explanatory relations. One of my main thesis goes against these claims. On the contrary, I defend that the singular causal relation is basic. In the second part, I combine this discussion with the problem of the model that is adequate for explaining actions. I conclude that in it, you have to combine the causal and normative requirements. I defend a biological-functional model for explaining actions, where in fact one has a double normativity: one that emerges from the notion of "function" and the normativity involved in inferences. In the last two parts of the thesis, I concentrate on Belief. I subscribe the thesis that beliefs are dispositions that intervene between different levels. Finally, I examine two cases of irrationality: self-deception and Moore's paradox. I explain these phenomena in the context of the biological-functional model.

AGRADECIMIENTOS

*Nosotros no nos realizamos nunca.
Somos un abismo que va hacia otro abismo
-un pozo que mira al cielo.*

F. Pessoa

Quisiera agradecer, en primer lugar, a mis asesores Sergio Martínez y Raúl Orayen por haberme ayudado a hacer que los abismos se hayan hecho un poco más pequeños. También quisiera agradecer a los participantes de los proyectos de "Actitudes Proposicionales" y de "Modelos de argumentación y racionalidad" por sus comentarios a diversos trabajos que leí en los seminarios y que antecedieron a esta tesis, sobre todo a Maite Ezcurdía. A Eduardo Flichman por sus comentarios al primer borrador del material que aquí aparece en los primeros dos capítulos. Finalmente quisiera agradecer a Raymundo Morado por su amistad y paciente ayuda durante los muchos tropiezos que tuve durante la elaboración de esta tesis.

INDICE

AGRADECIMIENTOS

INTRODUCCION.....	i
-------------------	---

I. LA RELACION ENTRE LAS RAZONES Y LAS ACCIONES EN LAS EXPLICACIONES INTENCIONALES: ¿RELACIÓN CAUSAL Y/O NORMATIVA?

1. La causación física.....	7
1.1 Relación singular entre sucesos	9
1.2 La extensionalidad de la relación causal singular	11
1.3 Necesidad: Poderes inherentes vs. leyes	25
2. TIPOS DE VÍNCULOS ENTRE RAZONES Y ACCIONES: CAUSALIDAD MENTAL Y/O NORMATIVA	33
2.1 Reduccionismo y causación física	35
2.2 Niveles de causalidad: propiedades supervenientes	38
2.3 Causación racional	48

II. MODELOS DE EXPLICACIÓN INTENCIONAL

3. Posturas ontológicas y epistemológicas o pragmáticas	51
3.1 Modelo nomológico deductivo	54
3.2 Modelo de causalidad singular	57

3.3 Modelo de explicación intencional:	
Posiciones intermedias	65
i) Davidson	69
ii) Dennett: Instrumentalismo	80
3.4 Modelo de explicación biológico-funcional	92
III. CARACTERIZACIÓN DE LAS CREENCIAS	107
4. La creencia como disposición	115
4.1 Fenomenalismo	116
4.2 Realismo	118
4.3 Racionalismo	120
5. Ontología de las creencias	
5.1 Análisis disposicional	124
5.2 Niveles de creencias	129
5.3 El asentimiento	134
IV. ADSCRIPCIONES Y AUTO-ADSCRIPCIONES DE CREENCIAS	
6.1 Usos expresivos y descriptivos de los enunciados de creencia: Wittgenstein	145
6.2 La paradoja de Moore	151
V. CONCLUSIONES	161
BIBLIOGRAFIA	177

INTRODUCCION

Filósofos de todos los tiempos se han interesado en explicar y entender los fenómenos que se dan en el mundo que nos rodea y en desentrañar el papel que juega nuestra presencia humana en los desarrollos y cambios que en él se dan. Al igual que nos transformamos con el mundo, lo transformamos. A veces lo hemos hecho de manera constructiva aprovechando e incrementando sus bondades pero, desafortunadamente, la mayoría de las veces intervenimos destruyéndolo y destruyéndonos. En el mundo se dan simultáneamente y en diferentes niveles una gran variedad de fenómenos entre los que se encuentran los que aquí nos ocuparán: los que se gestaron de acuerdo con una finalidad, plan o propósito.

Ha sólido considerarse que las acciones propositivas o intencionales constituyen un tipo de fenómenos que tienen el rasgo distintivo de estar orientados hacia un fin. La explicación de las acciones intencionales suele englobarse dentro de las llamadas explicaciones teleológicas, que en algunas épocas han incluido a las explicaciones en términos de *entelequias* (como en Aristóteles y en Leibniz) y las explicaciones en términos de las causas finales de los fenómenos. Al querer colocar las explicaciones intencionales dentro de los patrones de explicación de los fenómenos naturales nos exponemos a erosionar esa dimensión constitutiva.

Históricamente, Hume fue quien revirtió la costumbre de separar las explicaciones de los fenómenos naturales de las de los fenómenos humanos llevándolos todos al terreno de las explicaciones naturales. El único punto caso el cual los antecesores de Hume, en particular los racionalistas, acercaron las explicaciones naturales y las humanas fue al considerar que los cambios eran producidos por "poderes" internos o "fuerzas" en los objetos que repercutían en otros objetos. Con ese modelo explicaban las relaciones entre la mente y el cuerpo, concebidas como dos entidades que conforman al hombre que, a

través de la "fuerza" de la voluntad, deseos, etc. pone en movimiento el cuerpo. Los vínculos, tanto a nivel físico como mental, se concebían como vínculos causales, con la diferencia de que los vínculos entre la mente y cuerpo eran más estrechos e íntimos. Es bien sabido lo que Hume opina al respecto: no tenemos un conocimiento claro de la relación de causación física.

A partir de Hume la tradición empirista abandona la referencia a "poderes inherentes" en los objetos como causas del cambio y, aunque no todos siguen a Hume en negar la objetividad de la causación, se remiten a las leyes para explicar las relaciones entre los objetos. En el escenario en el que debaten positivistas y anti-positivistas se vuelve a discutir el problema de las explicaciones teleológicas. Los positivistas erigen el modelo nomológico de explicación como criterio de "cientificidad" y consideran que las explicaciones teleológicas no lo cumplirían, poniendo así en entredicho su calificación como explicaciones. El foco de atención en este debate son las explicaciones históricas y, quienes tratan de salvaguardar la legitimidad de las explicaciones teleológicas se rehúsan a tomar como criterio de las explicaciones el que se sustenten en leyes. Más bien, hacen hincapié en que los acontecimientos que se explican en los ámbitos sociales y humanos son singulares e irrepetibles de modo que, si explicarlos consistiera en verlos como instancias de una ley y, perderían su especificidad, que es lo que finalmente busca este tipo de explicaciones. Hempel recoge, en nuestra era contemporánea, esta idea de explicar los acontecimientos particulares subsumiéndolos en leyes, en su conocido modelo nomológico de explicación. Hempel intenta unificarlos a todos bajo ese modelo de explicación.

En la actualidad se ha impugnado el modelo hempeliano, esta vez desde las propias filas de los filósofos de la ciencia y no de las filas de los filósofos de las ciencias sociales. De todas las críticas que se le han dirigido, en este trabajo destaco la que cuestiona que se haga descansar la validez de la explicación en las leyes. Ciertamente no es unánime el rechazo de la fuerza explicativa que se les asigna a las leyes en los modelos físicos, como

tampoco es unánime su rechazo en los modelos de explicaciones intencionales. Un modelo alternativo, que intenta usurpar el lugar monopólico que venían ocupando las leyes, propone la existencia de una multiplicidad de valores mediante la introducción de consideraciones contextuales. Este desplazamiento de valores explicativos supone un viraje que abre la brecha para que las ciencias físicas puedan unificarse con las explicaciones en las ciencias sociales y humanas, alentando la búsqueda de *desiderata* para todas las explicaciones. En este trabajo no me ocuparé de la pretendida unificación de las explicaciones y sólo mantengo que el viraje hacia la línea pragmática va en la dirección correcta. Conforme a esta línea, ocupan un primer plano las explicaciones singulares, en lugar de las nomológicas, y se ponen de relieve las consideraciones contextuales. Pero, en el caso de algunas explicaciones, entre las que se encuentran las explicaciones de las acciones, hay que agregar algo más: los diseños del sistema. Sin referencia a ellos no acertaríamos a explicar adecuadamente el comportamiento de los sistemas. En el contexto de las explicaciones singulares se mantiene la diferencia entre las explicaciones de los acontecimientos físicos y las acciones, pero se traza de una manera distinta a la acostumbrada, en donde la manera de introducir la normatividad juega un papel decisivo.

A lo largo del trabajo presento algunas maneras de dar cuerpo a la idea de que las acciones intencionales son comportamientos guiados hacia un fin. Centro su evaluación en torno a dos cuestiones básicas: la estructura lógica de los enunciados explicativos y la manera en que en las explicaciones de las acciones intencionales intervienen la causalidad y la normatividad.

En lugar de entrar directamente al tema de la causación mental, en el primer capítulo examino el concepto de causación física. Esto lo hago con la causación física con el objeto de establecer con claridad algunas objeciones y dificultades que se le han señalado al modelo causalista en cuanto forma de vincular las razones y las acciones. A partir de los escritos de Davidson se ha vuelto común sostener que las razones (creencias, deseos, intenciones, etc.) pueden ser causas de las acciones que inducen. Es importante

dilucidar el sentido en el que las razones contribuyen causalmente al desenvolvimiento de las acciones. La comparación entre la causación física y la mental también me permitirá afinar el sentido en el que la causalidad interviene en las explicaciones intencionales: si la causalidad es *complementaria* -i.e. es un requisito adicional- a las explicaciones o si forma parte de la investidura misma de la explicación.

En la primera sección analizo la relación causal, donde la considero como una relación que se establece entre sucesos particulares. Examinó dos aspectos de la causalidad, uno lógico y el otro semántico: el de extensionalidad y el de necesidad, respectivamente.

Se han presentado algunos contraejemplos a la tesis que considera los contextos causales como extensionales. Después de referirme a diferentes estrategias que se han adoptado para contestar estos contraejemplos, argumento que la manera como Davidson la defiende, a saber, como una relación binaria inanalizada que se establece entre sucesos es esencialmente correcta. La respuesta de Davidson es la misma que da cuando habla de una causación mental, a saber, que cuando entramos en la cuestión de las propiedades en virtud de las cuales se establece el vínculo causal, en realidad ya nos hemos trasladado al dominio de las explicaciones. A esta respuesta de Davidson agregó que es importante que nos refiramos al diseño de nuestro sistema mental, que introduce la referencia a ciertas funciones biológicas evolutivas. Si logramos establecer que las funciones conllevan un elemento normativo, esto nos permitirá resolver la dificultad que se genera en la postura de Davidson, de que se escindan en las explicaciones intencionales los aspectos causales de los aspectos normativos. Si la solución de tomar la causalidad mental como una relación explicativa es viable, esto hace innecesario, como algunos han propuesto, recurrir a un tipo de causación distinto del físico para hablar de las acciones.

La discusión de la necesidad me lleva a suponer que en caso de que se acepte hablar de una causación singular que se establezca a través de las leyes, tendría que pensarse en términos de una causalidad probabilista, entendida en sentido propensista. De

este modo, se puede mantener la idea de que el vínculo causal singular encuentra sustento en las leyes, pero se negaría que la relación causal fuera una relación necesaria. Otra manera de defender la idea de una causalidad singular es considerarla como una relación primitiva. En este trabajo tomo una postura de este último tipo en la cual los casos singulares adquieren una primacía tanto ontológica como explicativa.

En el segundo capítulo considero el problema de las explicaciones y me pregunto por el modelo que más conviene para hablar de las acciones intencionales. Discuto los siguientes modelos: los modelos nomológicos de Hempel y el vínculo de éstos con el modelo disposicional, con base en la teoría de las decisiones, el modelo de explicación singular, el modelo intencional y el modelo funcional. De los dos primeros modelos, *qua* intentos de acomodar las explicaciones intencionales dentro de las explicaciones científicas, considero que su defecto mayor consiste en asimilar el patrón normativo, indispensable para las explicaciones de las acciones, a un patrón nomológico. Considero que el modelo de explicación que mejor recoge los aspectos normativos y causales es un modelo biológico-funcional. En un modelo de este tipo se defiende que de hecho intervienen dos formas de normatividad -una asociada a la idea de "función biológica evolutiva" y la otra la forma de normatividad asociada a un patrón de racionalidad. En la normatividad funcional se introduce una noción de contenido -que no se entiende en sentido semántico- y que es la que permite que la conducta de ciertos organismos además del humano pueda caracterizarse teleológicamente.

En los dos últimos capítulos, una vez elegido el modelo de explicación para lo mental, circunscribo la discusión de lo mental a la creencia. Analizo las diferentes maneras en que se ha entendido el estado mental de la creencia y lo conecto con el estudio de los diferentes niveles de creencias. Inicio este análisis con algunas de las características que más comúnmente se han asociado a las creencias, como lo son la de estar vinculadas con la verdad y la manera específica en que se vinculan con la conducta verbal. En particular, me detengo en la manera en que la voluntad y la deliberación intervienen en las creencias.

Al igual que autores como Mellor y Dennett distingo dos niveles de creencias y señalo algunas rasgos que caracterizan cada nivel y las maneras en que se relacionan. En esta discusión surge el problema de qué tan correcto es identificar, como hace Mellor, las creencias de segundo orden con el nivel donde surge la conciencia. O, qué tan correcto es, como hace Dennett, desglosar las conexiones entre los diferentes niveles en términos de relaciones de diferentes sub-sistemas. La cuestión es si el primer nivel -el nivel que Dennett denomina sub-personal- es un nivel básico que pueda cubrir el comportamiento o modo de funcionamiento de seres y objetos más simple que el de los humanos. McDowell, en particular, sostiene que en las discusiones de estos temas suelen confundirse dos procesos: el de la captura de ciertos datos que nos vienen del medio ambiente y el del procesamiento mismo de información. McDowell introduce esta diferencia con el propósito de señalar que en el nivel sub-personal se da el primer proceso en el cual hablar de creencias y de contenidos como parte de esos procesos es metafórico y que sólo en relación con el último proceso se puede hablar de ellos de manera genuina. La selección de información, para McDowell, no la lleva a cabo ningún sub-sistema, ni se trata de una relación mediatizada entre sub-sistemas, sino que la lleva a cabo el sistema mismo como receptor, estableciendo un vínculo directo con el entorno. Intento defender que aunque se pueda trazar una diferencia entre ambos procesos, no es claro que en la captura de elementos ambientales no esté ya interviniendo alguna forma de contenido al igual que alguna forma de percatarse de él.

Con el desarrollo de ciertas capacidades lingüísticas y de pensamiento hemos alcanzado una enorme variedad y flexibilidad en nuestra conducta que, para comprenderla, necesitamos la postulación de un nivel adicional de creencias. La intervención del lenguaje nos obliga a identificar la conducta humana de una manera mucho más precisa y diferenciada de la que se necesita para explicar la conducta de seres menos complejos que nosotros. En este nivel es donde aparece la noción semántica del contenido.

El comportamiento de los sistemas evolutivamente menos desarrollados, o de objetos diseñados para desempeñar funciones más específicas, puede explicarse apelando a las creencias de primer orden, las cuales se vinculan funcionalmente con ese comportamiento. Entendemos el vínculo como funcional donde otros lo han considerado como disposicional. Algunos de esos seres incluso podrían modificar sus respuestas según el entorno en el que se encuentran. También hay otros seres que, por el tipo de disposiciones que tienen o que pueden desarrollar, amplían el margen de posibilidades de sus respuestas, lo que permite que las creencias, articuladas con otros estados de creencia, deseos, intenciones, etc., se pongan al servicio de diferentes proyectos o intereses. Muchos autores se han empeñado en negar que este mecanismo sea sólo más complejo que el mecanismo que hay en los organismos más simples, sino que mantienen que su funcionamiento refleja que estamos frente a un mecanismo cualitativamente distinto. La posibilidad de valemos de nuestros estados mentales para los más variados fines es prácticamente el único argumento en el que se basan autores como Davidson, Evans, Peacocke y Williams a sostener que es erróneo hablar de "creencias" en los comportamientos más simples. Pero es igualmente factible sostener que hay muchas semejanzas en los comportamientos que nos permiten ordenarlos en grados de complejidad que nos permita hablar de un mismo mecanismo que en ciertos organismos ha alcanzado un mayor desarrollo evolutivo. Esta continuidad no se advierte si hablásemos simplemente de un mecanismo *sui generis* que sólo existe, hasta donde sabemos, en el hombre y que el desarrollo de las capacidades lingüísticas y de autorreflexión fuesen las que posibilitaran la aparición de los contenidos.

Si las razones fungen como causas de las acciones, entonces éstas tienen que hacer referencia a los contenidos semánticos de esas razones. Sugiero que esa referencia a los contenidos semánticos puede recogerse apelando a los *modos* en que se presentan ciertas propiedades en el sujeto. Como es insuficiente que nos refiramos sólo a las propiedades, podríamos agregar que se trata de propiedades de segundo orden. Así, considero que, en

el caso de los sistemas que tienen la capacidad de formar creencias de segundo orden, desarrollan un tipo de disposición relacional, una disposición de segundo orden o funcional. Podemos suponer que las creencias básicas tendrían un contenido, más no un contenido semántico, y que este último sólo aparece cuando se tiene el segundo nivel de creencias. Una vez que se posibilita la existencia de las disposiciones de segundo orden, cuya existencia depende de la existencia de disposiciones más básicas, debe ser posible que se pueda llevar a cabo el proceso inverso, de modo que estas disposiciones más complejas puedan influir y tal vez modificar las creencias más simples, en las cuales esos contenidos semánticos enriquecen los contenidos de las disposiciones más básicas. De no ser posible este camino inverso, la referencia a contenidos semánticos, en el caso de las creencias más básicas, se reduciría a un mero "modo de hablar".

Si bien con las explicaciones intencionales intentamos ofrecer un patrón que consideramos racional, en los capítulos finales me ocupo de la manera en que podemos explicar ciertas situaciones que consideramos irracionales. Dos de los patrones de irracionalidad que se analizan con mayor frecuencia son el de la *akrasia* y el del auto-engaño. El primero, como un ejemplo de irracionalidad que se produce entre lo mental y las acciones, y el del auto-engaño, como una incoherencia producida entre creencias o entre diferentes estados mentales del sujeto. De estas formas de irracionalidad me ocupo únicamente del auto-engaño y cómo es posible explicarlo en un modelo de niveles de creencias.

Para finalizar, me refiero a la denominada "paradoja de Moore", como otro ejemplo de incoherencia. Este caso me sirve para ilustrar la manera en que las consideraciones pragmáticas pueden afectar la manera de explicar cierto enunciado. En este sentido, las aportaciones de Wittgenstein al respecto son invaluable. En el caso particular de las creencias es palpable la diferencia que surge según se trate de las auto-atruciones o de las atribuciones de los otros. En cada caso hay que desentrañar el uso particular que hace un sujeto cuando dice creer *p*. En algunas circunstancias, puede afirmar una creencia con

el propósito de crear expectativas en sus oyentes en cuanto a que se comportará de acuerdo con lo que dice. En otras circunstancias puede tener en mente otros propósitos. En unos casos, cuando dice que cree p o dice simplemente " p ", el empleo del término de creencia puede ser redundante, en otros no. En cuanto a la paradoja de Moore, presento una interpretación de lo que Wittgenstein dice y señalo en qué difiere de otras interpretaciones que se han hecho, tanto de la paradoja de Moore, como de los pasajes en los que Wittgenstein se refiere a las creencias.

I. LA RELACION ENTRE LAS RAZONES Y LAS ACCIONES EN LAS EXPLICACIONES MENTALES: ¿RELACION CAUSAL Y/O NORMATIVA?

En nuestra vida cotidiana explicamos con naturalidad algunos de los comportamientos de nuestros semejantes -y el propio- refiriéndonos a las creencias, deseos e intenciones que le antecedieron. También diferenciamos con naturalidad esta forma de explicar los acontecimientos humanos de las explicaciones de los acontecimientos en el mundo. Entablamos así un marcado contraste entre hacer y acontecer -contraste que consignamos en sus respectivas explicaciones- y, consecuentemente, reaccionamos, juzgamos y evaluamos las situaciones según las veamos.¹ Esta distinción se refleja en nuestra manera de entender los flujos y reflujos de, y hacia, nuestros semejantes y nuestro entorno. En particular, en los casos en los que sabemos -o creemos- que una situación es producida o provocada por la intervención de la voluntad de algún agente lo responsabilizamos de ella. De igual manera, medimos las acciones de un agente en cuanto a sus fallas o aciertos en función de lo que pretendió, o creemos que pretendió. Nuestras convicciones respecto de las diferencias que entrañan lo que se hace y lo que acontece permean íntegramente nuestra vida, su valor y su significado.

Algunas de nuestras formas de pensamiento pre-teórico se han conservado en las teorías que elaboran tanto filósofos como científicos. Muchas de nuestras concepciones pre-teóricas, sin embargo, una vez sometidas a una reflexión más detenida pierden consistencia. La distinción pre-teórica entre hacer y acontecer de la que nos valemos para explicar los diferentes comportamientos tampoco ha escapado a semejante cuestionamiento. En las últimas décadas venimos presenciando una fuerte tendencia hacia

¹ El campo semántico del "hacer" es más amplio que el de "actuar" que en latín, como señala Mosterín (1987), corresponden a *agere* y *facere*, pero que nuestro uso castellano incluye a ambas en "hacer". La diferencia consiste en que no todo lo que hacemos, por el hecho de hacerlo, es una acción.

la modificación o evaporación de la distinción, en sentido contrario a nuestras intuiciones más arraigadas. Es ardua la batalla que están librando tanto filósofos como científicos y expertos en psicología cognitiva para convencernos de que nuestra distinción entre hacer y acontecer es equivocada. Cualquier propuesta de disolverla o retrazarla se enfrenta al reto de no perder en el intento la concepción de nosotros mismos como agentes racionales, libres y responsables.

La nueva topología de la distinción ha buscado orientación en los modelos de la física (de las ciencias naturales), ya que en esos campos la preocupación por la cuestión de la explicación ha sido, hasta ahora, más atendida. Una vez aplicados esos modelos al ámbito de lo mental se establecen semejanzas, diferencias y limitaciones y se requiere cautela para no caer en extrapolaciones apresuradas. Si se toman las precauciones necesarias constituye un buen comienzo. Sin embargo, hay que advertir que en las ciencias físicas los modelos de explicación al igual que la naturaleza de los fenómenos a explicar se encuentran a su vez en discusión. En la física contemporánea se proponen modelos teóricos que nos hablan de fenómenos estocásticos y teorías en donde imperan las leyes indeterministas y, desde la biología, emergen propuestas novedosas de modelos evolucionistas.

Los filósofos han acuñado el término de explicaciones intencionales para referirse a nuestras acciones. Frente a estas explicaciones se perfilan básicamente dos posiciones: a) la de quienes defienden que las explicaciones de las acciones, en general, las explicaciones que provienen de las disciplinas sociales y humanas, son estructuralmente distintas y autónomas de las explicaciones en las ciencias físicas. b) La posición de quienes sostienen que todas las explicaciones comparten una misma estructura. La homogeneidad explicativa suele hacerse tomando como paradigma las explicaciones científicas pero, más recientemente, se ha revertido la situación de manera que cada vez es mayor el número de filósofos que intenta hacer extensivas las estructuras de las explicaciones en las disciplinas sociales y humanas a las explicaciones científicas. La

mayoría de los participantes en el debate considera la cuestión de las leyes como decisiva para sus posiciones. Así, surge la pregunta de si en las disciplinas sociales y humanas puede haber leyes semejantes a las que se encuentran en las ciencias físicas y otros, más bien, se preguntan por el papel que desempeñan las leyes en las explicaciones, si es que consideran que desempeñan alguno.

En este trabajo considero que el prisma de las leyes que ha servido de base para la discusión de las explicaciones es muy estrecho. Como veremos más adelante, la existencia de leyes dista mucho de ser la virtud principal de las explicaciones. Si orientamos la discusión hacia la dimensión pragmática de las explicaciones es más fácil apreciar esta multiplicidad de valores. La comprensión de nosotros mismos con todo y la riqueza y complejidad de nuestra conducta, se resiste a ser apresada, aún en sus trazos más generales y esquemáticos, en una estructura nomológica. De allí que la inclusión de una gama de virtudes no debe verse como una mera manera de enriquecer nuestras explicaciones sino como parte de sus condiciones necesarias.

No me ocuparé aquí de la cuestión de si existe un modelo unificado de explicación. Mi objetivo básico y central es el estudio de las explicaciones intencionales. Aludiré a las explicaciones científicas, a manera de contrapunto, para destacar algunos rasgos que se han considerado característicos de sus modelos más representativos, con el propósito de ver en qué medida también se encuentran en las explicaciones intencionales. En este trabajo sostengo que al menos en las explicaciones intencionales la referencia a la causalidad es ineludible, sin con ello hacer que recaiga el peso explicativo en las leyes o en generalizaciones. La posición que aquí defiendo se opone tanto a las posturas que aceptan el modelo nomológico de explicación como a las posturas que introducen las leyes en las explicaciones a través de la noción de causa. La postura que aquí defiendo, más bien, le otorga primacía a las explicaciones causales singulares sin el respaldo de leyes o sin que forzosamente se evoquen leyes. Por otra parte, el compromiso que contraigo al admitir la irrenunciable referencia a la causalidad no conlleva que sea la única información

explicativa pertinente. Como verá en el segundo capítulo, también tenemos que hacer referencia al diseño de aquello a lo cual destinamos la explicación. La aproximación a lo mental se hace a través del vínculo biológico-funcional y no sólo a través del vínculo causal. Estos vínculos no se encuentran en el mismo nivel y el vínculo biológico-funcional o el sistema visto a partir de su diseño depende del vínculo causal. De las diversas funciones que estructuran a un diseño o sistema, algunas especifican el vínculo funcional no sólo etiológica o causalmente, i.e. especificando cómo de hecho funciona, sino normativamente, i.e. especificando cómo debe funcionar. El patrón normativo se constituye analizando la función del sistema diacrónicamente y no meramente en un individuo sino a través de su continuidad a nivel de la especie.

En este capítulo me referiré al concepto de causación física que eventualmente me servirá para aclarar el sentido de este concepto cuando aparece en relación con las explicaciones intencionales. Abordaré el problema de la causalidad en su doble aspecto, metafísico u ontológico y epistemológico, sin dejar de mencionar algunas consideraciones lógico-semánticas. Distingo diferentes tipos de causación: causación física determinista o probabilista, mental y racional o normativa, para después explicitar lo que considero característico de la causalidad mental.

En la primera sección me concreto a examinar dos rasgos de la causación: la extensionalidad de los contextos causales y el concepto de necesidad en tanto noción semánticamente conectada a ella. Antes de entrar en la discusión de la extensionalidad de los enunciados causales me refiero a la cuestión de la estructura lógica de los enunciados causales singulares y de allí paso a considerar la conexión que existe entre la estructura lógica de los enunciados causales singulares y la discusión de la extensionalidad. Estimo, como sugiere Davidson, que la relación causal puede tomarse como una relación binaria primitiva, cuyos *relata* son sucesos, en combinación con una defensa de su extensionalidad. Respecto al segundo rasgo, considero erróneo el vínculo semántico que se establece entre la noción de necesidad y la de causa, sea en la forma de poderes

inherentes en los objetos o en la forma misma del vínculo. Davidson admite únicamente la existencia de la causación física y a ésta le agrega la idea de un mundo que se rige por leyes deterministas. Para muchos, uno de los puntos más frágiles de la postura de Davidson radica justamente en su defensa de la conexión entre causación y leyes estrictas. En este trabajo, además de cuestionar este nexo, también cuestiono el apoyo de la causación en leyes del tipo que fueren, que muchos suponen correcto. Dicho brevemente, defendiendo la alternativa que da primacía a la causación singular. En lugar de un modelo nómico suscribo un modelo del tipo que propone Woodward para los enunciados causales singulares, en el cual se asocian con contrafácticos. La sugerencia supone resistir la tesis de que los enunciados causales singulares se *analicen* en términos de contrafácticos y tomarlos más bien como presupuesto de que de no haberse dado ciertas condiciones, eso habría hecho una diferencia en cuanto a la obtención del suceso-efecto.

En el siguiente capítulo me valdré de la discusión de la extensionalidad de la relación causal para diferenciar la relación explicativa de la relación causal, de modo que la relación explicativa no se tome como extensional sino como intensional. Su no-extensionalidad modifica también sus *relata*, que en lugar de aludir a sucesos aluden a enunciados.

En el aspecto epistemológico hago referencia a la tradición humeana en la cual se introducen las generalizaciones como apoyo o base para la relación entre sucesos particulares. Examinó varias posturas que hacen intervenir a las leyes en la discusión, tomando como eje la postura de Davidson. Cuando en el siguiente capítulo pasemos a la discusión de las explicaciones veremos la manera en que en algunos modelos, como el modelo nomológico deductivo de Hempel, nuevamente se juzga primordial la referencia a las leyes. También en los contextos explicativos defenderé que las leyes no son determinantes.

En la segunda sección entro a la discusión de la causación mental y aludo a tres propuestas que se refieren al vínculo entre razones y acciones: la propuesta reduccionista

que admite la existencia de un sólo tipo de causación: la causación física; la propuesta que alude a diferentes niveles de causación y que toma a las razones como propiedades mentales entendidas como propiedades supervenientes; y, finalmente, la propuesta de un tipo específico de causación, distinto del físico, para dar cuenta de lo mental. Me propongo explicitar, de una manera más clara de lo que suele hacerse, los diferentes compromisos que conllevan estas propuestas, en particular, las dos últimas. Quiénes defienden la causalidad que se da entre propiedades mentales supervenientes, generalmente la acompañan de sus correspondientes leyes, probablemente no-estrictas. En cambio, quienes conciben lo mental en términos de una causalidad racional consideran que los principios que la rigen son normativos y se resisten a considerarlos como el equivalente de las leyes en el ámbito de lo físico. Una forma que suele tomar el debate se recoge en términos de si lo psicológico puede "naturalizarse". La búsqueda de una causación mental distinta a la física responde al rechazo de asimilar la normatividad que define lo mental a un patrón nómico. La diferencia de patrones se detecta en la forma en que cada uno de esos patrones opera. Incluso si aceptamos, como señala Wittgenstein, que puede haber varias formas de normatividad y que en alguna de ellas su función fuese semejante a la de las leyes, no sería la normatividad que se tiene en mente cuando se habla de lo mental. Al respecto, la discusión de Wittgenstein de "seguir una regla" es esclarecedora. La normatividad de éstas es comparable a la normatividad que caracteriza a lo mental. En un tipo de funcionalismo biológico incluso se habla de dos tipos de normatividad: una normatividad que se introduce a través de la noción de función biológica evolutiva y otra normatividad asociada más específicamente a lo mental, que tiene que ver con principios de racionalidad y de inferencia. Se pretende que en ambas formas de normatividad se haga referencia a los contenidos y para una postura no reductivista es crucial no confundir ni eliminar alguna de las formas de normatividad. Esta es la tesis que discutiré y defenderé en el siguiente capítulo.

Quando se habla de una causación mental, al aludir a lo mental, la relación se hace

más compleja -deja de ser una relación binaria- y se requiere que las propiedades mentales se introduzcan como parte de la relación. Esta situación nos abre dos caminos: o sostenemos que la relación causal mental sigue siendo extensional pero que constituye un tipo de causación distinto al físico, o consideramos que cuando intervienen las propiedades mentales, en realidad estamos pasando al ámbito de las explicaciones y la relación pasa a ser intensional. Opto por la segunda vía ya que considero que incorpora de manera más satisfactoria el elemento normativo que caracteriza a lo mental.

1. La causación física.

Se puede abordar la cuestión de la causalidad desde varias perspectivas: genética, epistemológica, en cuanto a la forma lógica de sus enunciados, ofreciendo un análisis semántico, o como noción metafísica. Si bien pueden destacarse en el estudio de esa noción diferentes intereses, de hecho, suelen encontrarse en un mismo autor combinadas tesis que corresponden a diferentes perspectivas. Así Hume, por ejemplo, se interesó tanto del aspecto genético-epistemológico del problema como del aspecto metafísico. Muchos autores han seguido la misma ruta epistemológica trazada por Hume, sin ocuparse de la cuestión de la génesis de la idea de causalidad y, en cambio la conectan con cuestiones lógicas y semánticas. En este trabajo, primeramente me refiero de manera rápida a la cuestión de la forma lógica de los enunciados causales singulares y del análisis semántico de la noción de causalidad para después pasar a la cuestión epistemológica, a través de las explicaciones causales.

Es insoslayable la penetración del discurso causal en nuestro lenguaje, tanto en nuestros contextos cotidianos como en los contextos científicos y filosóficos. El empleo de terminología causal forma parte integral e incisiva de nuestra apreciación cotidiana de los cambios que percibimos en el mundo pero, sobre todo, en relación con las acciones en las que intervienen agentes. Nos apegamos a ella, indistintamente, al hablar de sucesos, procesos, hechos, estados de cosas, o cualquier término que los filósofos hayan forjado

como los *relata* de la causación. Como advierte Anscombe, su presencia se sugiere fácilmente en el uso de un sinnúmero de verbos como empujar, tirar, golpear, sumergir, disolver, quemar, etc.. Con el empleo cotidiano de estos verbos se admite, sin cuestionamiento alguno, la objetividad de la relación causal. Y, sin cuestionamiento, lo conservan muchos filósofos y científicos por igual. Sin embargo, en el proceso de desentrañarla, figuras de la talla de Hume (según algunos intérpretes), Russell y Quine, entre otros, han terminado prescindiendo de ella. Así, Russell nos insta a que la noción de causalidad se expurgue del vocabulario filosófico. Considera que la ley de causalidad sólo constituye "una reliquia de una era olvidada, que sobrevive, como la monarquía, sólo porque se supone, erróneamente, que no causa daño."² Quine, dentro de su favorecida - y estrecha- concepción de la ciencia, juzga que la noción de causa, en la medida en que no puede resolverse en términos de predicados, cuantificadores y funciones de verdad, debe correr igual suerte que el discurso intensional, a saber, que "la ciencia en su forma más austera prescinda de la noción y se decida por nociones concomitantes".³

¿Qué hay detrás de la noción de causalidad que ha suscitado reacciones tan polares como la declaración de la obviada de la existencia de esa relación y la eliminación total de su existencia? ¿Acaso su existencia obvia equivale a que se la tome como un hecho bruto, o más bien debemos tomarla como una relación compleja que nos fuerza a que la descompongamos o analicemos en los elementos más básicos que la integran y definen? Parece desmedida la reacción de quienes propugnan por su eliminación, sobre todo, dado lo insatisfactorias que han sido las formas de la relación que han venido a ocupar su lugar. En este trabajo suscribo la idea de que existe la causación singular en tanto vínculo objetivo entre sucesos.

La preocupación por el problema de las relaciones causales puede abordarse

² Russell (1963), p. 133.

³ Quine (1990) p. 76.

separadamente del problema de las explicaciones causales, aunque no ha sido lo más usual. Desde Aristóteles ha sido común responder a la pregunta del por qué de algo remontándose a su causa. En el caso concreto de Aristóteles el término "causa" significa diferentes cosas o diferentes maneras posibles de explicar un mismo hecho, a saber: material, formal, eficiente y final.⁴ Comparto la idea de raigambre aristotélica de que el problema de la causación es importante en la medida en que se vincula al ámbito de las explicaciones pero, para efectos expositivos, me referiré a ellas por separado, iniciando en esta sección con la causación física. A pesar de que me concierne la discusión de la causalidad aplicada a lo mental, me referiré a la manera de entender el vínculo causal que se produce entre acontecimientos físicos, ya que es en comparación con ésta que se ha visto la causación mental como problemática.

1.1 La relación causal singular.

Davidson propone una caracterización de la forma lógica de los enunciados causales singulares que resulta intuitivamente atractiva. Su sugerencia es que se trata de un enunciado relacional binario primitivo, i. e. inanalizado. Concebir a la relación causal como primitiva sólo compromete a Davidson a decir que se trata de un término inanalizado pero no a que sea *inanalizable*.⁵ En "Causal Relations" Davidson explícitamente señala que la propuesta de la forma lógica de la relación causal es más modesta que la de ofrecer un *análisis* de ella. Más aún, considera que la cuestión de la forma lógica de la relación causal y el análisis de la noción de causa no deben confundirse, como algunos lo han hecho. De la propuesta de tomar a la relación causal como un predicado binario deriva que los *relata* que les convienen a los enunciados

⁴Para una discusión más detallada de la postura de Aristóteles véase Guariglia (1982).

⁵Cfr. P. 147 y 161, Davidson (1982). Respecto a esta observación al igual que las que se refieren a la relación entre la forma lógica y el problema de la extensionalidad agradezco la ayuda de Scott Soames y Raymundo Morado.

causales singulares sean *sucesos*. Davidson agrega que la cuestión de la no extensionalidad de los contextos causales no surge ya que sus enunciados se expresan en un lenguaje extensional de primer orden.⁶ Por otra parte, Davidson identifica a los sucesos en términos de sus antecedentes y consecuencias causales. Así, decir que el *succeso-causa* es condición necesaria y suficiente para el *succeso-efecto* es un resultado inmediato de su manera de identificar los sucesos. De estas tesis, no triviales, como dice Platts, "se sigue (trivialmente) que la causa, en tanto suceso particular, es tanto necesaria como suficiente para el efecto. Si el *succeso-causa* no hubiera producido el *succeso-efecto*, no habría sido el suceso que fue. Asimismo, dado que el *succeso-efecto* ocurrió, el *succeso-causa* debe haber ocurrido y debe haber causado el *succeso-efecto*, ya que de otro modo el *succeso-efecto* habría sido un suceso diferente."⁷

Si en los enunciados causales singulares se vinculan sucesos, los términos descriptivos que empleemos para capturarlos van a tener una función eminentemente referencial; esto es, suponemos que lo determinante para la verdad del enunciado causal estriba en fijar o elegir adecuadamente los sucesos entre los que se establece la relación. No es indispensable que se incluyan las propiedades en virtud de las cuales los sucesos producen sus resultados. Pero, en caso de que se incluyeran en el enunciado, si no nos olvidamos de que su función es eminentemente referencial y no descriptiva, la verdad del enunciado no tiene por qué alterarse al hacer sustituciones correferenciales. Cuando se viola el principio de sustitutividad es porque en realidad la función del enunciado también

⁶ La forma de la relación que Davidson favorece no deja tan claro que la pregunta por la no extensionalidad de la relación no surja: dependería del tipo de predicado binario de que se trate. Podría haber predicados binarios tanto extensionales como intensionales. Por ejemplo un predicado binario no extensional sería el que aparece en el enunciado "Los griegos adoran a Zeus", pues si se sustituyera "Zeus" por un nombre que tuviera su misma extensión como "Afrodita" pasaríamos de un enunciado verdadero a uno falso. La extensión sería la clase vacía. Es una cuestión abierta que el predicado causal binario sea o no un predicado extensional de primer orden, de la misma manera que sería una cuestión abierta el tipo de argumentos que integran la relación.

⁷ Platts (1992) p. 297.

se ha alterado.

La lectura estándar de la tesis de la extensionalidad se aplica a los términos referenciales, de manera que las sustituciones de descripciones en sus términos singulares, si es que el contexto es genuinamente extensional o transparente, no debe alterar el valor de verdad del enunciado; en otras palabras, el principio de substitutividad *salva veritate* se cumpliría en los contextos causales.

1.2 La extensionalidad de la relación causal singular

Así como hay una gran cantidad de enunciados causales que apoyan la tesis de la extensionalidad de la relación causal, otros parecen desmentirla. En la literatura del tema se han ofrecido muy diversas maneras de defenderla frente a esos casos. En esta sección me referiré a algunas de ellas considerando que, de todas, la solución que ofrece Davidson, en la cual se recurre al expediente de la noción de explicación causal como salida a los contraejemplos, es más satisfactoria que sus rivales. El argumento de fondo en el que me apoyo para defender la extensionalidad de la relación causal es que tenemos un enunciado causal completo verdadero con sólo seleccionar adecuadamente los sucesos que así se encuentran vinculados.

La mayoría de los filósofos finca la asimetría entre las relaciones y las explicaciones causales singulares en torno al requisito extensional, haciéndola depender virtualmente del hecho de que pueda hacerse efectiva la independencia descriptiva o lingüística de los enunciados causales singulares.

La forma lógica de las relaciones causales es diferenciable de la forma lógica de las explicaciones causales pero tenemos que decidir si la cuestión de la objetividad de la relación causal es lo que da sustento y validez a nuestras explicaciones. Como veremos en el siguiente capítulo, existen versiones tanto ontológicas como epistemológicas de la explicación que dan una contestación distinta a esta cuestión.

La liberalidad en el manejo de descripciones en relación con los sucesos, que se

desprende del criterio amplio con el que Davidson los identifica, hace de su postura un

blanco fácil que amenaza la transparencia referencial de los sucesos.⁸ Así, la catarata de contraejemplos no se ha hecho esperar. Como muestra de algunos de los más usuales,

podemos mencionar los que emplean adverbios o adjetivos calificativos:

- 1) El que Lucía girara tan abruptamente causó que se le dislocara la rodilla.

podría ser verdadero a la vez que ser falso que

- 2) El que Lucía girara causó que se le dislocara la rodilla.

Pese a que

- 3) El que Lucía girara tan abruptamente = El que Lucía girara = El que Lucía girara tan abruptamente en la cancha de tenis al mediodía.

⁸ Para algunos autores es simplemente falso que todos los infinitos modos en que puede describirse un suceso haga apropiado que individúen al mismo suceso, sobre todo si se individúan los sucesos por sus causas. Es fácil imaginar casos en el que un mismo objeto, por ejemplo una esfera que simultáneamente gire y cambie de color en donde, con base en que tienen diferentes causas, no tendríamos ningún reparo en decir que el que la esfera gire sea un suceso y el que la esfera cambie de color sea otro suceso.

Por lo tanto, es posible inferir un enunciado falso de uno verdadero, mediante la sustitución en 1) de términos correferenciales. Otro ejemplo sería "El impacto del Titanic contra un iceberg causó que se hundiera" podría ser verdadero, sin necesariamente serlo que "El impacto del Titanic contra un iceberg causó que se hundiera rápidamente", ya que el hundimiento rápido del trasatlántico podría haber sido causado por una carga excesiva.

Se podría también enfatizar alguna parte o aspecto del suceso que ocupa o la posición de la causa y/o del efecto. El énfasis se destaca subrayando la expresión que, en un enunciado causal, hace suponer que lo subrayado es el factor causal clave. Un ejemplo sería:

4) El que Sócrates bebiera cicuta en la prisión causó su muerte.

Si suponemos que 4) es verdadero y suponemos que el enunciado es referencialmente transparente, de modo que

5) El que Sócrates bebiera cicuta en la prisión = El que Sócrates bebiera cicuta en la prisión = El que Sócrates bebiera cicuta en la prisión.

De 4) y 5) podemos inferir

6) El que Sócrates bebiera cicuta en la prisión causó su muerte.

Se podría decir que 6) es falso ya que sugiere que la parte del suceso que está enfatizada es lo causalmente eficaz; esto es, que el suceso aconteciera en la prisión fuera lo que causara su muerte.

Este tipo de casos parecería forzar la decisión en la dirección de renunciar a la extensionalidad de los contextos causales y a desvanecer la diferencia entre los contextos

causales y los contextos causales explicativos. En los últimos, en general, se da por sentado que la sustitución de descripciones no preserva la verdad de las explicaciones, i.e. constituyen contextos intensionales. Un multicitado ejemplo explicativo de este tipo sería, "El que Edipo se casara con su madre explica que enloqueciera" podría ser verdadero bajo esa descripción y ser falso bajo otra descripción, como por ejemplo, "El que Edipo se casara con la viuda de Layo explica que enloqueciera".

Frente a las dificultades que acechan a la extensionalidad de la relación causal, las opiniones se han dividido entre las siguientes disyuntivas: la de quienes, se empeñan en salvaguardar la extensionalidad de la relación causal singular y, la de quienes no tienen ningún reparo en catalogarla como una relación intensional.

Existen varias estrategias para conservar la extensionalidad de los contextos causales, que tienen en común el interés de asegurar el contraste entre los contextos causales y los contextos explicativos. Una sería considerar que tanto los términos de sucesos como los términos que especifican aspectos de sucesos denotan. Una segunda estrategia sería modificar la forma lógica de los enunciados causales, por lo general, haciéndola más compleja. Por último, se podría optar por precisar la tesis misma de la extensionalidad. En este caso se podría adoptar la sugerencia de Marcus⁹, de que hay principios de extensionalidad más fuertes o más débiles, que nos permiten restringir la sustitutividad según la equivalencia que se quiera establecer: Dretske, por ejemplo, toma el primer sendero, Kim el segundo, mientras que Beauchamp y Rosenberg, el tercero. D. Lewis y Woodward protagonizarían la vía de preservar la extensionalidad de los contextos causales, con la modalidad de que también la defenderían en los contextos explicativos causales singulares. En este sentido la postura de los dos últimos autores irían a contrapelo de las tendencias predominantes. Ambas posturas se caracterizan por defender una posición en la cual la explicación causal singular tiene la función de seleccionar información acerca de los sucesos que se vinculan causalmente.

⁹ Marcus (1971).

Pese a la gran cantidad de autores que han escrito en su contra y pese a que no logre dar cuenta satisfactoria de todos los casos, me parece que la sugerencia de Davidson tiene mayores méritos que sus alternativas. Su solución consiste en dejar a la relación causal como extensional y los casos en los que se tenga la intención de destacar algún aspecto del suceso, podría rephrasearse en términos de una explicación. Pasemos entonces a las diferentes maneras de defender que los enunciados causales singulares son extensionales.

Dretske pretende salvar la idea de que los contextos causales singulares son extensionales, sugiriendo que podemos distinguir entre sucesos y aspectos de sucesos -que denomina "sucesos-alomorfos"- y permitir que los aspectos de los sucesos se utilicen como términos referenciales. Dretske juzgaría que al pasar de un enunciado como 4) -i.e. sin aspectos enfatizados- a un enunciado como 6) -i.e. con algún aspecto enfatizado-, al hacer la sustitución de términos, los enunciados cambian de valor de verdad, debido a que la denotación de sus expresiones también cambia. Así, Dretske no admitiría las identidades señaladas en 5) y, por ende, no concedería que se trata de la sustitución de términos correferenciales.¹⁰

La idea de Dretske de convertir a las descripciones que se refieren a los aspectos de los sucesos en denotativas es problemática. Por nombrar una de sus dificultades, se individualarían de manera muy estrecha y se evaporaría la conexión íntima que puede haber entre diferentes aspectos de un mismo suceso. Por ejemplo, el aspecto rápido de la corriente de un río y el aspecto del río que se desborda no serían independientes. Achinstein ataca la propuesta de Dretske sosteniendo que su razonamiento es falaz. De la verdad de que el enunciado causal depende, o es verdadero en virtud, de la eficacia causal del aspecto del suceso, no se sigue que ese aspecto del suceso sea un término referencial. Para Achinstein la selección de un aspecto del suceso, mediante el énfasis, es una manera de expresar que es un factor causalmente operante, e implica que tiene

¹⁰ Dretske (1977).

relevancia semántica. De ser así, el cambio de énfasis conllevaría un cambio en el significado de la oración y, de ese modo, posiblemente el valor de verdad del enunciado.¹¹ Si en el contexto de afirmar un enunciado causal singular se destaca alguna de las características del suceso-causa o del suceso-efecto es bastante natural suponer que se le está dando relevancia causal. Cuando la afirmación misma destaca alguno de sus términos cobra otro sentido: le da importancia a algún rasgo o propiedad del suceso y tendría repercusiones en la verdad del enunciado.

Como segunda estrategia, Kim sugiere modificar la tesis de la forma lógica de los enunciados causales de manera que los sucesos que contienen palabras enfatizadas se entiendan como formas elípticas de una conjunción cuyo primer miembro es el término-causa sin términos enfatizados -que sería un término referencial- y un segundo conyunto que es o un enunciado contrafáctico, o una oración que incluye una ley, u otra oración causal.¹² En la oración 6) ("El que Sócrates bebiera cicuta en la prisión causó su muerte") el término-causa "El que Sócrates bebiera cicuta en la prisión" no se referiría a nada. El enunciado 6) debe entenderse -según Kim- como una conjunción del término-causa no enfatizado que si sería un término referencial. Su forma completa sería:

6') El que Sócrates bebiera cicuta en la prisión causó su muerte y,
si él hubiese bebido cicuta en algún otro lugar no habría muerto.

Con esta solución Kim destaca que las sustituciones en el primer conyunto -en principio- se licenciarían, pero que con el segundo conyunto se verían bloqueadas en tanto que aludirían a diferentes propiedades. Kim negaría la identidad de los sucesos en 3) y

¹¹ Además del uso del énfasis, en sentido semántico, Achinstein sostiene que habría otro sentido que se emplearía para "negar" algo antes dicho, o para introducir nueva información, o para destacar algún interés especial de la oración empleada, en donde no se cambiaría el significado del enunciado. Cfr. *ibid.* p. 236.

¹² Kim (1977).

en 5) ya que sus propiedades constitutivas serían distintas.¹³ Además, Kim acompaña esta tesis con la tesis de que los sucesos están estructurados de modo que un objeto concreto ejemplifica una propiedad o relación en un tiempo dado.

Achinstein critica la posición de Kim sosteniendo que no da una verdadera solución a los contraejemplos, sobre la base de que no se libra de manera efectiva del problema del énfasis. Según Achinstein, en el enunciado sin el énfasis se crea una ambigüedad en cuanto a que se podría pensar, en una de sus lecturas, que todos los aspectos mencionados son causalmente eficaces, generando así "una afirmación causal que o carece de un valor de verdad único o tiene uno que no es necesariamente el mismo que el de la oración original".¹⁴ De hecho, Achinstein utilizará este último argumento para también criticar la postura de Davidson. Si el enunciado no enfatiza ninguna de sus expresiones o sería ambiguo o implicaría la eficacia causal de todos los aspectos que se mencionan. La crítica de Achinstein a Kim me parece correcta en tanto que, efectivamente, en la paráfrasis que ofrece del enunciado causal -que a su vez se analiza en términos de individuos y de propiedades en ciertos momentos [a,P,t] causó [b,Q,t']- la ambigüedad persistiría en su primer conyunto. Pero como veré más adelante, esta crítica de Achinstein a Davidson me parece incorrecta.

Con la tercera estrategia, Beauchamp y Rosenberg se proponen conservar la extensionalidad de los contextos causales permitiendo que los contextos causales admitan tanto términos singulares como oraciones como sus *relata*. Sugieren, así, una nueva prueba de extensionalidad para el caso de contextos causales que toman oraciones como sus *relata*, que denominan "extensionalidad nominal". Los enunciados explicativos contrastarían con los contextos causales en tanto que los primeros seguirían sin pasar esta prueba de extensionalidad. Para Beauchamp y Rosenberg la referencia de los sucesos a los

¹³ Cfr. Kim (1973).

¹⁴ Achinstein (1983) p. 242.

que aluden las nominalizaciones gerundivas de las oraciones tiene que ser la misma. Así, por ejemplo, la nominalización de "El Titanic se hundió" sería "El hundimiento del Titanic".¹² Dichas nominalizaciones suponen un funcionamiento análogo al de los términos singulares y, similarmente, puede esperarse que las sustituciones en las nominalizaciones de predicados coextensos, al retener la misma referencia, las oraciones en las que aparecen también conservarían su valor de verdad. El criterio de extensionalidad nominal aplicado a las explicaciones singulares podría no preservar el valor de verdad de los enunciados en los que se efectúa la sustitución. Así, en el ejemplo de Edipo antes mencionado, la sustitución de "haberse casado con la viuda de Layo" por "haberse casado con su madre" no preserva la verdad de los enunciados explicativos en los que aparece, pese a que cumple con el requisito de que sus respectivas nominalizaciones gerundivas tienen la misma referencia.¹³ Mi comentario respecto de la solución de Beauchamp & Rosenberg es que recurren a una complicación innecesaria. Para qué hablar de diferentes *relata* de la relación causal y asociarles diferentes criterios de extensionalidad. Con la estrategia de Davidson se resolvería de una manera más simple, dejando que los *relata* de la relación causal sean sucesos y reservar las oraciones para los enunciados explicativos causales.

Finalmente, paso a la postura de Woodward. No me referiré a la manera en que Lewis entiende las relaciones causales y la defensa de la extensionalidad para esos contextos. Lo característico de ambas posturas -la de Woodward y la de Lewis-, en comparación con las otras posturas que hemos relatado- es que consideran también a los

¹² Por ejemplo podrían establecerse las siguientes identidades: "El hundimiento del Titanic"= El suceso acerca del cual Walter Lord escribió un libro que constituyó su primer éxito de ventas= El suceso que le costó al Lloyd's de Londres más dinero que cualquier otro en 1912= El suceso que resultó en la única tumba masiva en el cementerio de Halifax. Beauchamp & Rosenberg (1981), pp. 264-269.

¹³ Cfr. op. cit., pp. 267-68.

enunciados explicativos causales como extensionales. Presuponen que es inteligible que en ciertas circunstancias indagemos por la causa de la existencia de un suceso *tout court* y de que el sentido de un buen número de explicaciones consiste en pedir sus causas. Para Lewis, en particular, resulta primordial la idea de un mundo en el que se tejen una infinidad de historias causales y que, cuando nos detenemos en alguna de estas historias, el suceso que nos concierne explicar constituye su último eslabón. Para Lewis, por así decirlo, las explicaciones ideales y completas serían aquellas en las que se tuviera la causa, toda la causa, y nada más que la causa. Lewis considera que "explicar un suceso consiste en proporcionar alguna información acerca de su historia causal."¹⁴ En un mundo determinista, preguntar por qué se dio un suceso, nos lleva a buscar la diferencia que hay en su historia, en comparación con las posibles historias alternativas a la situación real.

En Woodward las explicaciones causales singulares -que especificaré en un momento- comparten con las relaciones causales singulares el hecho de que relacionan sucesos y los relacionan respecto de cierta "clase de contraste". Woodward defiende la tesis de que los enunciados causales, al igual que ciertas explicaciones causales singulares, pueden relacionar sucesos: tendrían la misma estructura lógica y satisfarían el mismo principio de extensionalidad. Es decir, defiende la posición inversa a la de Beauchamp y Rosenberg en tanto que ahora son los enunciados explicativos al igual que los enunciados causales singulares los que pueden relacionar sucesos. Woodward considera que la pregunta por la causa de algún efecto es ambigua o incompleta, mientras no especifiquemos su foco de contraste o "clase de contraste". Así, según él mismo indica, retoma la sugerencia de Dretske y de Van Fraassen en cuanto a introducir en sus enunciados la clase de contraste. Para Woodward la estructura lógica de los enunciados causales singulares (y la de ciertos enunciados explicativos) no es una relación binaria sino ternaria (o de cuatro plazas si se introduce el índice temporal). El principio de substitutividad se ejerce sobre los términos que especifica la clase de contraste.

¹⁴ Lewis, D. (1993) p. 185.

Woodward incluye la clase de contraste como parte tanto de la pragmática como de la semántica de los enunciados. En este sentido diferiría de la posición de van Fraassen que lo inspira, ya que en este último la idea del foco de contraste se introduce exclusivamente como parte de la pragmática de las explicaciones y no encara el problema de los contextos causales como tales. Y, por la razón contraria, Woodward diferiría de la posición de Dretske: Dretske se ocupa de la semántica de los enunciados causales y en relación a éstos introduce su propuesta de la clase de contraste y se olvida de la pragmática.

Aparentemente, Woodward parece estar diciendo lo mismo que Kim y Achinstein: que el enunciado causal es ambiguo y que hay que introducir la clase de contraste para desambiguarlo y que la clase de contraste tiene la función de señalarnos con cuáles términos es intersustituible sin alterar el valor de verdad del enunciado. Sin embargo, Woodward conecta la distinción entre usos atributivos y referenciales de la descripción definida -utilizándola como distinción semántica- con la introducción del foco de contraste. Para Woodward el papel de la clase de contraste es desambiguarlos enunciados causales. Al desambiguar el enunciado, de hecho, estaría utilizando tanto el principio débil como el fuerte de substitutividad. Si el contraste se especifica en términos de cierta propiedad que despliega el suceso, entonces la sustitución de términos se limitaría a propiedades sinónimas. Si la clase de contraste se refiere a la ocurrencia del suceso se podría hacer la sustitución con cualquier descripción que se aplique al suceso.

Mi desacuerdo con Woodward -al igual que con Kim y Achinstein como ya mencioné antes- es que no considero que el enunciado causal sea incompleto o ambiguo por el solo hecho de que no especifiquemos la propiedad o propiedades causales pertinentes. Su estructura y las sustituciones que pudieran hacerse en ella dependerían de cómo se individuaran los sucesos, haciendo innecesaria la especificación de una clase de contraste en particular. En el reporte de la relación causal singular suponemos que nos interesa recoger los sucesos que participan en la relación causal. Pero Woodward agrega

que el foco de contraste se marca respecto de otros sucesos alternativos, esto es, *vis à vis* que no se hubiera dado o que se hubiera dado otro suceso. Sin embargo, estos contrafácticos no se referirían a esos mismos sucesos sino a otros. Con el enunciado causal buscamos fijar una referencia; su verdad dependerá del acierto en recoger a los sucesos adecuados, independientemente de los términos descriptivos que empleemos para llegar a ellos. Me parece que si aceptamos que la función básica del enunciado causal es referencial no procedería la construcción del contraste, como quiera que lo establezcamos.

Woodward se equivoca también de otra manera, al no hacer un uso adecuado de la distinción semántica entre usos atributivos y referenciales. Cuando se ha intentado defender que la distinción es semántica, se supone que la descripción definida es ambigua y que el enunciado cambiará sus condiciones de verdad según se tome atributiva o referencialmente. Me parece que en el caso del enunciado causal singular no habría tal fluctuación, si es que aceptamos que su uso es referencial. Si fluctuara, querría decir que la descripción definida no ocurre de manera esencial en el enunciado. Pero habría una razón más de peso por la que al enunciado causal no podría asignársele un uso atributivo a la descripción definida. Cuando Donnellan en su conocido artículo "Reference and Definite Descriptions" alude al uso atributivo de la descripción definida y la contrasta con el uso referencial tiene la función de recoger cualquiera suceso que satisfaga la descripción.¹⁵ Pero en Woodward en el uso atributivo del enunciado causal singular se presupone que ya se ha fijado cuál es el suceso al que nos referimos. La distinción en Woodward entre los diferentes usos a los que se puede prestar la descripción definida funciona de otra manera; no pretende introducir la distinción para resolver -como sostiene Kripke- el problema de que la referencia de quien emite el enunciado y la referencia semántica no correspondan.¹⁶ Este -para Kripke - es un problema de los actos

¹⁵ Cfr. Donnellan (1966) p: 258.

¹⁶ Cfr. Kripke (1977).

ilocucionarios y no un problema semántico. Por otra parte, si en Woodward la clase de contraste de por sí ya tiene la función de elegir el aspecto del enunciado que nos interesa, no parecería necesaria la distinción entre los diferentes usos de las descripciones definidas que finalmente sirve para lo mismo.

Ninguna de las estrategias para defender la extensionalidad de los contextos causales hasta aquí esbozadas está exenta de dificultades. Veamos si la propuesta de Davidson sale mejor librada.

Davidson pretende paliar la dificultad que surge en los casos en donde aparecen adverbios, proponiendo que esos enunciados se entiendan mediante la conectiva oracional "explicar causalmente". Los contrajemplos tienen el efecto de que el valor de verdad del enunciado causal se pueda alterar según se utilice una descripción u otra. Sin embargo, pese a que la estrategia de Davidson puede parecer un tanto forzada, me parece viable. Cuando Davidson recomienda que en esos casos pasemos a la terminología de "explicar causalmente", hace dos cambios importantes: a) los *relata* se convierten en oraciones y b) hace que la verdad del enunciado dependa de la manera como se describen los sucesos expresados en las oraciones. Con ello, las sustituciones que se hagan en las oraciones pueden cambiar el valor de verdad del enunciado explicativo.

Achinstein presenta un par de críticas a la postura de Davidson, ninguna de las cuales me parece correcta. La primera crítica ya la mencioné antes al referirme a la postura de Kim. La crítica es que el enunciado causal es ambiguo y en una de cuyas lecturas las propiedades mencionadas son causalmente relevantes. Si bien esta crítica me pareció correcta contra Kim, no me lo parece en el caso de Davidson. No hay que olvidar que las descripciones se usan para *fijar* los sucesos involucrados causalmente: no habría ambigüedad ni tampoco tendría que tomarse cualquiera de sus expresiones como causalmente eficaz. Utilizando el de Davidson, Achinstein sostiene que si afirmáramos que "El que Flora se secase con una toalla al mediodía explica causalmente el hecho de que aparecieran manchas en su piel" implícitamente estaríamos destacando que contamos a

todos esos aspectos como parte de la explicación. Pero entonces, tendríamos que considerar que la explicación ofrecida sería falsa porque introduce en ella elementos que son irrelevantes para la explicación. Me parece que el señalamiento de este aspecto de "relevancia" que sí tendría sentido en el contexto de las explicaciones es justamente una de las aristas del contraste con los enunciados causales singulares. Si, por otro lado, la oración explicativa contuviese elementos enfatizados, me parece que tendríamos que decir que su significado en ese contexto sería distinto al significado de la oración sin el énfasis. El propósito de recurrir a los contextos explicativo-causales es para darle relevancia semántica a todo lo que se señala en la explicación. En nuestra práctica de dar explicaciones sería tan inusual que se incluyeran todos esos aspectos sin el énfasis, como que se usaran enfatizadas. Por lo general sólo incluimos los aspectos en los que fincamos la explicación. Es más fuerte nuestra tendencia a simplificar las explicaciones en lugar de complicarlas introduciendo aspectos irrelevantes.

La otra crítica que Achinstein le dirige a Davidson consiste en que no considera que el paso de Davidson de una oración causal a una relación de "explicar causalmente", que equivale al paso de una relación extensional a una intensional, logre resolver el problema. Según Achinstein, de todas maneras se pretendería que "explicar causalmente" fuera relacional y, por tanto, que "la causa y el efecto tendrían que ser referenciales". Achinstein considera que Davidson no tiene abierta la opción que otra postura podría adoptar, a saber, que los enunciados que flanquean a la relación "explicar causalmente" no sean referenciales, ya que Davidson explícitamente dice que las explicaciones relacionan enunciados. Pero -según Achinstein- para ser genuinamente relacionales, las posiciones que ocupan la causa y el efecto deben ser referenciales.¹⁷ Sin embargo, si bien es cierto que la relación explicativa causal es genuinamente relacional, no sería correcto derivar que del hecho de que sean referenciales se tenga que aceptar que sean transparentemente referenciales. Si se acepta que el contexto explicativo es intensional, no

¹⁷ Cfr. *ibid.*, pp. 243-44.

se podrían sustituir las proposiciones que ocupan la posición-*causa* o la posición-*efecto*, en este caso por otras proposiciones. Esto refleja que su ocurrencia no es *puramente referencial*. Como dice Quine: "La cuestión acerca de las comillas [en un enunciado] no es que debe destruir la ocurrencia referencial, sino que puede (y por lo común así sucede) destruir la ocurrencia referencial."¹⁸

La propia solución que contempla Achinstein es un claro ejemplo que corresponde al segundo bloque de propuestas que opta por convertir a los contextos causales en intensionales. También supone que no se corre un serio peligro si los contextos causales son intensionales. Achinstein considera los aspectos de los sucesos a la manera "*de dicto*", que nos inclinaría a suponer que los hace depender de aspectos descriptivos, convirtiendo así a los contextos causales que hacen referencia a ciertos aspectos de los sucesos en opacos o intensionales. Pero en lugar de llegar a esta conclusión, Achinstein afirma que la terminología entre causas y sucesos no es la adecuada para los enunciados causales singulares.¹⁹ Achinstein sostiene que en los contextos causales la referencia a los sucesos es opaca. Sostiene que la noción de causa, de manera semejante a la de explicación, opera seleccionando las expresiones de los aspectos de las situaciones que tienen importancia semántica. Su función es seleccionar algún rasgo del suceso. Esta solución de Achinstein me parece desmedida ya que no deja sitio para que hubiera términos en la relación que funcionarían referencialmente. Tampoco creo necesario reiterar que no comparto la idea de que la función central del enunciado causal sea la de seleccionar alguna propiedad del suceso, sobre todo si se parte de que la ontología básica es de sucesos, mientras que sí creo que lo sea en el caso de la explicación. Esa fue una de las razones por las que consideré que la clase de contraste no tiene por qué incluirse en la estructura del enunciado causal singular.

¹⁸ "Reference and Modality" (1961), p. 141.

¹⁹ Ibid. cap. 6.

Hasta aquí hemos señalado algunas variantes que se abren para quienes consideran preferible dar una vuelta de tuerca en nuestra caracterización de la relación causal que nos permitiera manejar los contraejemplos, antes que renunciar a la tesis de su extensionalidad. En conclusión, a mi juicio, lo más viable o preferible de todas estas maneras de dar cuenta de los enunciados causales singulares sería una caracterización de la relación causal como lo hace Davidson. Sin embargo, no excluyo que se podría proponer una manera más diferenciada de la que propone Davidson para identificar a los sucesos, sin menoscabo de la tesis de la extensionalidad de los contextos causales. Por ejemplo, se podría utilizar un criterio distinto para identificar sucesos -que no fuera en términos de las mismas causas y los mismos efectos- o que excluyera a los adverbios y los adjetivos y se les diera a éstos un tratamiento lógico distinto o, tal vez, que sólo incluyera a sus propiedades esenciales.

1.3 Necesidad: Poderes inherentes vs. leyes.

En esta sección discutiré la cuestión de si la necesidad es una característica semánticamente asociada a la de la causalidad. Ambas nociones - causalidad y necesidad- han solido acompañar a la concepción determinista del universo. Se puede recurrir a la necesidad como parte del análisis semántico -descomposicional- de la noción de causa o desde el punto de vista epistemológico, en tanto forma de diferenciar cognoscitivamente secuencias causales de las que no lo son. Me referiré a algunas posturas que intentan conservar una noción de causalidad que salvaguarda la necesidad via contrafácticos al igual que otras posturas que intentan prescindir de la noción de necesidad sin prescindir de los contrafácticos. Hay quienes recurren a las leyes en apoyo de los casos singulares mientras que en otros se parte del caso singular. Como he venido diciendo, en este trabajo defiendo el punto de vista de que las relaciones causales singulares se tomen como primitivas, en particular, que no tienen que descansar en leyes y no tienen que analizarse de manera contrafáctica.

Según la más difundida interpretación de Hume, él es el primero que tuvo el

atrevimiento de negar la objetividad de la causación, ya fuera como algún poder que reside en los objetos o como una conexión entre diferentes objetos. En esta lectura se considera que Hume introduce como parte del "análisis" de la noción de causación la idea de conjunción constante entre dos objetos distintos, sustituyéndola por la de necesidad. De ser correcta esta interpretación, Hume estaría -ontológicamente- renunciando a la causación como relación natural entre los fenómenos físicos y exilándola al ámbito de lo psicológico. La causación, en particular, caería bajo el dominio de la facultad de la imaginación (y no de la razón). Así, en esta versión que para muchos es la oficial, se atribuye a Hume el haber reemplazado la noción de necesidad (objetiva) por la de conjunción constante y convertido a la necesidad en una impresión interna (subjctiva) que el sujeto proyecta al mundo. La causación pasa de ser una característica intrínseca del mundo a una relación que depende de la naturaleza humana. Esta tesis se glosa como su respuesta escéptica al problema de la causación y a las inferencias inductivas.

De ser correcta esta manera de interpretar a Hume, a raíz de sus escritos se estarían estableciendo dos caracterizaciones antagónicas respecto de la causalidad: la de Hume, que sería reductiva, en tanto que la "analiza" o descompone en términos de una regularidad o uniformidad natural y, la que conserva la noción de necesidad como parte esencial de la noción de causalidad; esto es, en la segunda caracterización la noción de causalidad se trataría de un concepto eminentemente modal. Esta última sintetizaría la posición defendida por los predecesores de Hume, sobre todo la de los racionalistas.

Sin embargo, esta interpretación reduccionista de Hume es cuestionable si se entiende como tesis ontológica, más no en cuanto explicación de cómo nos formamos la creencia o la "idea" de la causación. Sumándose a la interpretación de Anscombe y a dos de las interpretaciones de Hume menos ortodoxas, la de Stroud²⁰ y la de Beauchamp y Rosenberg²¹, en cuyo apoyo esgrimen razones distintas, considero que la atribución de

²⁰ Stroud (1977) y (1978).

²¹ Beauchamp & Rosenberg (1981).

la versión oficial a Hume es equivocada.

Para Beauchamp y Rosenberg, Hume suscribe -como resultado de dos proyectos distintos- ambas caracterizaciones de la causalidad arriba mencionadas y que es posible formar con ellas "una sola teoría unificada". Concluyen que Hume sólo quiere criticar la manera en que los racionalistas entienden la causalidad que intenta negar que la necesidad que entraña sea una necesidad lógica que puede establecerse de manera *a priori*. Es decir, Hume sólo cuestionaría que el problema de la causación que introduce elementos empíricos, se pueda determinar de manera *a priori*.²² Si bien es cierto que hay apoyo textual para atribuir a Hume la referencia a ambas caracterizaciones, no estoy segura de que puedan conformar una tesis unificada. De cualquier forma, es innegable que el proyecto central de Hume es un proyecto eminentemente epistemológico.

Por su parte, Stroud considera que Hume: "i) no ofrece un análisis de la causalidad "en términos de regularidad", ii) más que negar que además de la regularidad tenemos una idea de necesidad o de conexión necesaria, afirma que tenemos esa enigmática idea y iii) pasa la mayor parte del tiempo intentando explicar cómo la obtenemos."²³ Según Stroud, la crítica de Hume a la noción de causa no pretende disolver el vínculo semántico entre esta noción y la de necesidad, ni tampoco negar o afirmar que en el mundo externo se relacionen sucesos de manera causal. Me parece que esta interpretación que hace Stroud de Hume, según la cual no se compromete con ninguna tesis ontológica, es esencialmente correcta. Hume se abstiene de ello, por ejemplo, cuando escribe en el Tratado:

... la necesidad es algo que existe en la mente, no en los objetos; tampoco es posible para nosotros formarnos la más remota idea de ésta, considerada como una cualidad en los objetos. O no tenemos la idea de necesidad, o la necesidad no es más que una determinación del pensamiento de pasar de causas a efectos y de efectos a causas, según su unión que hemos

²² Cfr. *ibid.*, Capítulo 2.

²³ Cfr. Stroud (1978) p. 39.

experimentado.²⁴

Como se apresta a agregar Stroud, de este pasaje no podemos extraer que la causalidad sólo opera en la mente y no en el resto de la naturaleza. Hume sólo enfatiza el hecho de que tenemos esa idea, no porque descubramos directamente su presencia cuando observamos la conjunción entre dos objetos particulares, sean mentales o físicas, sino a través de ciertos sucesos en nuestra mente cuando nuestra experiencia exhibe ciertos rasgos.²⁵ La sugerencia de Stroud es que la necesidad, que se fragua en la mente, la proyectamos como relaciones objetivas entre objetos. *Creemos*, a partir de lo que ocurre en nuestra mente, que en el mundo los objetos se vinculan de manera necesaria. Ciertamente la idea de necesidad se establece entre ideas (en el sentido humano) y Hume no dice mucho respecto a si existe ese rasgo en el mundo independientemente de los sujetos cognoscentes. Stroud hace estas observaciones con el propósito de que no perdamos de vista el proyecto genético-epistemológico en el cual Hume inscribe su discusión. Para reforzar esta misma idea de Stroud podemos recordar que Hume tampoco sostiene que en el mundo gobiernan el azar o las probabilidades objetivas. Las nociones de azar y de probabilidad señalan la manera en que repercuten en nuestras creencias. Hume considera que lo azaroso o fortuito, propiamente dicho, no es más que la negación de la causalidad y que influyen en la mente de manera opuesta: así como la causalidad "se impone" a nuestro pensamiento, el azar lo deja "totalmente indiferente".²⁶ En Hume, en el imperio de la contingencia, la mera idea de una causalidad probabilista en tanto rasgo objetivo, como hoy en día reivindican algunos autores, era impensable.

En Hume, al igual que en los filósofos de su época, hay una marcada inclinación por desentrañar la génesis de la idea de causalidad y a partir de allí traza consecuencias

²⁴ Hume (1978), pp. 165-66.

²⁵ Stroud (1978) pp. 43-44.

²⁶ Cfr. Hume (1958) Libro I, Sección XI.

epistemológicas. Y, como bien señaló Stroud, no se le interpretaría justamente si se le atribuyese el interés y la pretensión de dar un análisis semántico de la noción de causalidad. También resulta claro que Hume trata de llegar a la noción de causación indirectamente a través de la regularidad de sucesos, ruta que ha servido de inspiración a uno de los modelos de explicación más influyentes, el modelo nomológico deductivo de Hempel. Hume propugna por la tesis de la generalidad implícita en las instancias causales particulares, o la existencia de leyes implícitas en sus instancias, coincidiendo con Mill, en que las secuencias particulares son instancias de sus respectivas leyes.

Hoy en día estamos lejos de la preocupación humeana en torno a la génesis de nuestras ideas en general, y de la causalidad en particular, y su apego irrestricto al modelo empirista de los diferentes tipos de ideas. Sin embargo, siguen vigentes algunos de los enigmas que se planteó en torno a la causalidad. Sigue preocupándonos cómo se distinguen las relaciones causales singulares de las relaciones que no lo son y sigue preocupándonos si realmente tenemos una idea clara de la necesidad fáctica que entraña. A estas preguntas se les puede dar un giro epistemológico u ontológico. Se puede también combinar el problema ontológico con una caracterización lógico-semántica de los enunciados causales singulares y/o generales. Hume no reparó mucho en las interrogantes ontológica y lógico-semántica, y de manera generosa admite una ontología muy heterogénea (que incluye impresiones, objetos, acciones y movimientos, cualidades, entidades matemáticas, el yo). Igual habla de la relación causal singular física que mental, que de principios generales o leyes.

Al explicar las relaciones causales singulares en función de sus leyes, Hume les quitó primacía causal a las relaciones causales singulares, revirtiendo así la tendencia que venía imperando en sus predecesores. Como dicen Beauchamp y Rosenberg: "Los predecesores de Hume supusieron que la relación causal se analiza en términos de poder inherente, eficacia o por acción de algún agente -o tal vez de la transmisión de alguna cantidad como energía, que un poder inherente hacía posible. También creyeron que las

leyes causales se derivan y se establecen a partir de la repetida experiencia de las secuencias particulares de fenómenos, independientemente reconocidas con el carácter de causales.²⁷

Pese a Hume, seguimos preguntándonos si la relación causal debe considerarse como la relación, o quizá una de tantas, que da cuenta de la estructura del mundo, y la lista de filósofos convencidos de que sí es una relación objetiva sigue incorporando nuevos simpatizantes.

La discusión contemporánea en torno a la causación, además de plantearse como una discusión ontológica, al igual que en la época de Hume, siguen abordándola como una cuestión epistemológica, pero esta vez a través de la discusión de las explicaciones. Por un lado, están quienes sostienen que la causación nos remite a las propiedades o poderes inherentes en los objetos, comprometiéndose a darle primacía a la existencia de la relación en los casos particulares. Por otro lado, se encuentran los que consideran que las leyes o generalizaciones sirven de apoyo a los casos particulares en tanto instancias de ellas. No es sino hasta muy recientemente, y ya no sólo en boca de algunas voces disidentes, que se ha visto seriamente cuestionado el modelo nómico y se ha vuelto la mirada a la causación singular como prioritaria.

En esta sección suscribo la postura que le concede primacía a los casos particulares y que entre ellos puede establecerse un vínculo causal, y que la satisfacción nómica es adicional y no constitutiva del concepto de causación. La manera de defenderla es mediante contrafácticos, a la manera como sugiere Woodward. Combino esta idea de Woodward con algunas de las tesis que defiende Davidson, a la vez que me distancio de la postura de este último en cuanto al papel que le asigna a las leyes. Davidson establece que las leyes son condición necesaria y suficiente de la causalidad. También apela a las leyes en el contexto de las explicaciones, aunque fuese sólo para las explicaciones físicas básicas. No me parece que la ausencia de leyes psicológicas o psico-físicas sea el mayor

²⁷ Beauchamp y Rosenberg (1981) p. 80.

impedimento para asimilar el modelo de explicación psicológico al físico.

Davidson es uno de los perpetuadores contemporáneos de la tesis de que el mundo está causalmente estructurado. Asimila en su postura el supuesto, asumido sin mayor argumentación, de que la causación se vincula con leyes estrictas. Da por hecho que la causación es una relación fundamental en el sentido de que una caracterización completa del universo estaría dada en términos causales que contendrían las propiedades "suficientes para la formulación de un sistema cerrado de leyes estrictas".²⁸ Davidson expresa esta idea mediante el principio que él denomina "El carácter nomológico de la causalidad", según el cual cada enunciado causal singular verdadero está respaldado por una ley estricta que conecta sucesos del tipo al que pertenecen los sucesos mencionados como causa y efecto.²⁹

Muchos autores han considerado una debilidad de la postura de Davidson que conserve una noción de causalidad asociada a una concepción determinista del universo. (Anscombe, Woodward) Con ellos sostengo que el reconocimiento de la presencia e importancia de las teorías indeterministas en la ciencia. Quienes no quieren renunciar a la relación causal ni tampoco a su objetividad, consideran imperioso revisarla para legitimar su aplicación en estos contextos. Si las leyes que rigen en un mundo indeterminista fuesen de carácter irreductiblemente probabilista, se tendría que renunciar a la idea de una necesidad fáctica y suplirla por la de causalidad probabilista. La exploración de la noción de probabilidad en combinación con la de causalidad puede tomar una vertiente frecuentista o una vertiente propensista. En este nuevo escenario se reedita la pugna entre el modelo nómico *versus* el que da prioridad a los fenómenos singulares. En el modelo nómico se podría optar por reducir la noción de causalidad y

²⁸ Cfr. Davidson (1993), nota 10, p. 17.

²⁹ Cfr. "Mental Events", (1980), p. 223. No es difícil juzgar por qué Davidson suscribe estas tesis. Su énfasis en que a nivel físico haya leyes estrictas es clave para preservar el contraste entre el discurso físico y el discurso mental y, consecuentemente, su defensa de su conocida tesis del "monismo anómalo" respecto a lo mental. Me referiré a esta cuestión más adelante.

"descomponerla" en términos estadísticos o, alternativamente, se podría pensar en la existencia de un tipo de propiedades disposicionales de carácter propensista. Podría decirse que en la última, a diferencia de la primera, se propugnaría no sólo por una noción epistémica sino por la existencia de probabilidades reales -ontológicamente hablando- asociadas a un tipo particular de propiedades. En este caso, la referencia a su valor probabilista indicaría que se trata de un proceso azaroso.

Podemos defender una postura que dé primacía a las relaciones causales singulares en la que no resulte problemático aceptar la existencia de ciertas propensiones en los objetos, que al interactuar con el entorno posibiliten la presencia de cierto desenlace.

Una de las críticas que más insistentemente se ha señalado respecto a los contextos indeterministas es que al quitarle a la causalidad su sesgo necesario se corre el peligro de que sólo se podría dar cuenta de qué sucedió, y en donde la probabilidad mayor no incrementa la posibilidad de que en este caso particular se obtenga cierto resultado: podría haber sucedido lo improbable.

Tal vez una vía sería darle nuevamente prioridad a la relación causal singular, tomando la sugerencia de Anscombe, que al centro de la relación causal singular está la idea de la derivación del efecto a partir de la causa: que el efecto proviene o surge de su causa. Este repensar en el "poder" intrínseco en los objetos podría servir para dar primacía a la relación causal singular en donde, el que el efecto surja o emerja de la causa no implica que similares efectos surjan de similares causas. El olvido de este aspecto de la causalidad se explica por la inercia heredada de Hume y de Mill que conlleva a la concentración en el aspecto de "necesidad" o algún aspecto semejante asociado a las leyes. Si le damos primacía a la causalidad singular entonces, aunque asumiéramos que el mundo es esencialmente indeterminista, el índice de probabilidad de que ocurriera deviene una información adicional y no parte de lo que explica que se dé cierto efecto. La relación causal podría considerarse como básica o primitiva.

Si se da primacía a los casos causales singulares otra alternativa -que no se opone

a tomarla como una relación primitiva- sería que se les asociara con contrafácticos. El problema se trasladaría a la cuestión de cómo distinguir los contrafácticos que se asociaran a una relación causal y los que no, por ejemplo los casos de contrafácticos como "Si mi hermana no hubiera tenido un hijo yo no habría sido tía", o "Si a no se hubiera ubicado a la izquierda de b , b no habría estado a la derecha de a ". Me parece viable y sugerente la caracterización de la causalidad mediante contrafácticos, aunque éstos no constituyan o formen parte de su forma lógica ni tampoco de su análisis. No me voy a comprometer con ninguna caracterización de las condiciones de verdad para el contrafáctico. Esta vía es distinta de la de autores como Lewis que recurren a la idea de contrafácticos como análisis de la causalidad, introduciendo la semántica de mundos posibles. En otros autores se recurre a propensiones o disposiciones como por ejemplo Mellor.

En este trabajo no intentaré argumentar en favor de alguna de las caracterizaciones de la prioridad de relación causal singular. Me basta con asumir, à la Davidson, que se trata de una noción lógicamente primitiva. Lo crucial para mi proyecto es que hay bases para sustentar una noción de causalidad sin que nos casemos ni con una concepción determinista del universo ni con un modelo nómico que la sostenga.

2. TIPOS DE VINCULOS ENTRE RAZONES Y ACCIONES: CAUSACION MENTAL Y/O NORMATIVA

En la sección anterior, hemos señalado algunas de las cuestiones más apremiantes en relación con el vínculo causal entre sucesos físicos, que me permitirán introducir la causación mental. Cuando se recurre a la causalidad para dar cuenta del vínculo entre sucesos mentales o del vínculo entre lo mental y lo físico, se hacen más profundas las dificultades en torno a ella. La razón por la que se piensa que la causación mental es más enigmática se debe, por un lado, a que se piensa que lo mental y lo físico apuntan a diferentes tipos de fenómenos y que por tanto no es claro cómo pueden influirse mutuamente de manera causal. Por otro lado, también se piensa que el funcionamiento de

la esfera de lo mental es, en sí, más complejo en tanto que los estados mentales intervienen de manera holista. Se podría agregar que la estructura misma de algunas de las actitudes mentales es más compleja en otro sentido: sus estados se especifican haciendo referencia a sus contenidos. Estas consideraciones nos hacen pensar que la causación física necesita al menos suplementarse o quizá hasta requerir otro concepto de causación. Para que el enunciado causal sea verdadero basta con que se dirija a los fenómenos adecuados. Pero en el caso de la causación mental los términos de la relación tienen que capturarse de manera más específica. En la sección anterior sostuve que en el caso de la estructura misma del enunciado causal físico se podía prescindir de la especificación de la propiedad o propiedades responsables causalmente. El enunciado causal entendido como una relación binaria entre sucesos es en sí completo. Por el contrario, en el caso de la causación mental podemos sostener que la especificación o descripción apropiada de los sucesos exige que se incluyan las propiedades mentales para que el enunciado causal sea completo. ¿De qué diferentes maneras puede acomodarse ese dato respecto de la causación física? En la postura de Davidson se sostiene que el discurso mental, por el hecho de apoyarse en descripciones, pertenece ya al contexto de las explicaciones. Para Davidson hablar de causación mental constituye una de las muchas formas en que se pueden redescibir los hechos. El suceso-efecto se redescibe como una acción, deja de ser visto como un mero acontecimiento físico. Ese giro nos presiona a que se aluda a descripciones adecuadas a la dimensión o aspecto que se quiere destacar. En sus escritos más recientes Davidson ha ido cambiando su posición -o tal vez sólo la ha ido precisando más-, de modo que ahora agrega las propiedades mentales. Estas seguirían subordinadas a su base física. Ninguna tesis monista o dualista respecto de lo mental puede dejar de reconocer -al menos- la pertinencia del tipo de discurso para poder asignar responsabilidad a los actos. Para algunos la importancia de la evaluación de los actos pediría algo más real o firme que el discurso.

Me parece que la solución de tomar al enunciado que hace referencia a la causación

mental como descriptivamente incompleta y la solución de Davidson de mover los enunciados causales mentales al ámbito de las explicaciones son bastante afines. Las diferencias se ahondarían si se pensara que la dependencia descriptiva convierte al enunciado causal en intensional. No considero que la dependencia descriptiva de la causalidad mental convierta a sus enunciados en intensionales, como también creo que más que preocuparnos por la relación causal misma, la discusión importante de lo mental se da en el contexto explicativo. Existe una inmensa literatura filosófica que, sin embargo, lo discute en el plano ontológico. En la siguiente sección me referiré a las que considero más prominentes.

2.1 Reduccionismo y causación física

Poner al descubierto las dificultades en torno a la causación mental ha tenido como consecuencia que las más extremas posturas dualistas se hayan hecho inviables; ya no parece haber retorno a la concepción de la mente cual "fantasma en la máquina", según la contundente descripción de Ryle. Si a este hecho le sumamos la gran cantidad de datos que arrojan las investigaciones científicas, no es de extrañar que la balanza se haya inclinado a buscar soluciones en la dirección de una cierta matriz física para lo mental.

Algunos filósofos han intentado paliar la carga del misterio alrededor del vínculo mente-cuerpo sugiriendo algún tipo de conexión sistemática entre lo mental y lo físico. Pero, como señala Davidson, cualquiera de sus formas, incluyendo la de superveniencia, implica la aceptación de una tesis monista que se realiza en lo físico.³⁵ Las diferentes conexiones van desde posturas reduccionistas de lo mental a lo físico hasta las que apelan a una forma de dependencia de lo mental en lo físico, pasando por formas que convierten a lo mental en un mero epifenómeno de lo físico. Surgen entonces preguntas como las siguientes: ¿Qué estatuto ontológico tienen las propiedades mentales? ¿Son propiedades

³⁵ Cfr. "Thinking Causes", (1993) p. 5.

que sobrevienen en las propiedades físicas? ¿Cómo debemos explicitar la relación de superveniencia de modo que las propiedades mentales tengan una potencia causal propia? ¿Hay varios tipos de causación o sólo uno? Y, de cara a estas interrogantes ¿cómo debemos entender el proyecto de naturalización del dominio de lo mental?

En esta parte del trabajo examino posturas que intentan responder a estas interrogantes, que introducen matices diversos a las posturas extremas, a la cabeza de las cuales se encuentra la de Davidson. Estas posturas -de alguna manera intermedias entre las reductivistas y las dualistas- navegan a contracorriente de la tendencia reduccionista extrema -i.e. como reducción ontológica y descriptiva-, en tanto que pretenden poder mantener un monismo físico frente a lo mental, sin naufragar en el epifenomenalismo. La delicada maniobra requiere que las propiedades mentales tengan alguna repercusión causal. ¿Qué razones esgrimen quienes persisten en señalar que los escollos en ese tipo de posturas son insalvables? ¿Cómo puede una postura como la de Davidson articular armónicamente que sólo haya sucesos físicos, ontológicamente hablando, los cuales son susceptibles de describirse mediante una amplia gama de propiedades, entre ellas las mentales, de tal suerte que las propiedades mentales de los sucesos no pierdan el poder de generar (ciertas propiedades de) otros sucesos? En una postura -como la de Davidson- en la que sólo se admite la causación física. ¿de qué manera se insertan en ella las propiedades mentales? En una visión fisicalista del universo ¿están destinadas a desaparecer -o a no tener cabida- las categorías mentales o intencionales?

Davidson, como ya vimos, deslinda el problema de las explicaciones de los problemas ontológicos, más específicamente, el problema de las explicaciones causales del problema de las relaciones causales mismas. Se pueden ofrecer otras salidas como la que ofrece Dennett, en la cual se asocia a cada nivel de explicación una cierta ontología. Aquí se estrecharían más las cuestiones ontológicas con las explicativas. O podría hacerse, como hace Lewis, que se identifiquen las explicaciones con las relaciones causales, de modo que una explicación consistiera en seleccionar segmentos de la cadena causal completa y que

su corrección o incorrección entrara en la pragmática de las explicaciones. ~~No es~~ un parámetro absoluto de corrección; algunas son más técnicas y especializadas que otras y cambian según la información que tengan los que ofrecen las explicaciones y según los destinatarios.

Mi conclusión será que el encasillamiento de diferentes tipos de propiedades dentro del modelo único de causación física no es inverosímil. Para poder recoger las propiedades mentales, no parece necesaria la referencia a un tipo distinto de causación, sea en la forma de propiedades supervenientes o de otro tipo, por ejemplo funcionales. Si bien, por una parte, las propiedades mentales guardan un cierto tipo de dependencia respecto de las propiedades físicas, por otra parte, la causación mental tendría que tener su propia modalidad.

Si se acepta que la relación causal singular se apoya en leyes y se acepta algún tipo de causación que no sea la física, según la relación que se establezca entre lo mental y lo físico, sería difícil evitar la existencia o de leyes psicológicas o de leyes psico-físicas. Si, por lo contrario, se acepta un reduccionismo fuerte, tendríamos un modelo como el que propone Churchland en el cual se defiende una tesis materialista-eliminacionista respecto de lo psicológico.

Churchland prepara el camino para la eliminación de lo mental quitando del paso una supuesta diferencia entre las leyes (o explicaciones) físicas y las leyes (o explicaciones) psicológicas, a saber, que las últimas contienen cláusulas *ceteris paribus*. De alguna manera, la presencia de este tipo de cláusulas impediría que habláramos de leyes. Estas cláusulas indicarían que podrían interceder muchísimos factores, aunque no de manera sistemática, que obstaculizaran que se dieran ciertos resultados. Las leyes, en una concepción determinista del universo, toman la forma de leyes estrictas que, en algunas interpretaciones, quiere decir que sus leyes tienen la forma de condicionales universales irrestrictos, i.e., que no contienen cláusulas *ceteris paribus*; no enuncian condiciones negativas que inhiban la influencia de los factores causales. Churchland no

está de acuerdo en que exista un contraste entre las leyes físicas y mentales en ese respecto ya que considera que también en la ciencia podemos encontrar leyes que contienen cláusulas *ceteris paribus*. Churchland presenta como un ejemplo físico que contiene cláusulas *ceteris paribus* el de la frotación de un cerillo. Si un cerillo se frota, en ausencia de factores que lo perturben, el cerillo se prendería. Sin embargo puede señalársele a Churchland que aunque sus leyes fuesen *ceteris paribus*, de cualquier forma subsistiría el contraste entre el ámbito mental y el físico (McGinn³⁶ y Moya³⁷). La fuerza explicativa o el carácter nómico de la generalización física derivaría de que puede vincularse con leyes estrictas, en las cuales se sustituyen los términos que se emplean en ella por términos correspondientes a la ley -que no hablaría de cerillos, frotación, etc.. En el caso mental, en cambio, o no sería posible la eliminación de términos mentales de las leyes, o, si se eliminaran, por ejemplo mediante leyes fisiológicas, un aspecto propio de lo mental, a saber, sus contenidos, y su normatividad se perdería. Estas dos razones me parecen decisivas en contra de las propuestas eliminacionistas. Más adelante volveré al aspecto normativo que caracteriza a los vínculos en los que interviene lo mental. La pregunta es si en un modelo de niveles de causación podemos encontrar una manera satisfactoria de dar cuenta de lo mental y logra superar las dificultades que abruma al eliminacionista.

2.2 Niveles de causalidad: Propiedades supervenientes

Concentrémonos ahora en un modelo no eliminacionista. Un modelo estructurado en niveles, como el que discute Kim³⁸, es uno de estos modelos. En este modelo, cada una de las capas superiores, si bien dependen de las inferiores, cobran una cierta autonomía.

³⁶ McGinn (1979).

³⁷ Cfr. Moya (1990), cap. 8.

³⁸ Cfr. Kim (1993).

La descripción de este modelo supone que los niveles superiores no colapsen en los inferiores y, a la vez, permita que las propiedades que pertenecen a cada nivel puedan tener una contribución causal novedosa. El núcleo de este tipo de modelo es que tanto en los diferentes niveles de relación como en sus explicaciones se tendrían sus propias leyes. Un modelo estratificado tendría una forma como la siguiente:

... un nivel básico, que consiste en lo que la micro-física nos dirá que son las más básicas partículas físicas, de las cuales está compuesta toda la materia (electrones, neutrones, *quarks* o lo que fuere). Y estos objetos, lo que fueren, se caracterizan por ciertas propiedades y relaciones físicas fundamentales (masa, *spin*, *charm*, o lo que fuere). Conforme ascendemos a niveles superiores, encontramos estructuras conformadas por entidades que pertenecen a los niveles inferiores y, más aún, se conciben las entidades en cada nivel como conformaciones a partir de un conjunto de propiedades distintivas de ese nivel. Así, en cierto nivel, encontraremos un conglomerado de moléculas de H₂O, con propiedades tales como transparencia, poderes de disolución de azúcar y sal, una densidad y viscosidad características, etc. En niveles jerárquicamente más altos encontramos células y organismos con sus propiedades "vitales", y más alto aún, organismos dotados de conciencia e intencionalidad. Más lejanamente, encontramos a los organismos que forman grupos sociales y tal vez grupos de esos grupos.³⁹

Cada nivel -según Kim- contiene un dominio de conjuntos de particulares y un conjunto de propiedades que se refieren a ese dominio. Cada propiedad que se yergue sobre los niveles inferiores tiene como punto de referencia básico un sistema físico, de modo que todas las propiedades de los niveles jerárquicos se realizan o se instancian en ese sistema físico. Kim sugiere que la estructuración de esa formación es mereológica (una relación de parte-todo). La cuestión es resolver cómo se relacionan las diferentes propiedades de los diferentes niveles y cómo se relacionan las propiedades en su respectivo nivel. Kim propone lo siguiente:

³⁹ *Ibid.*, 190.

Cuando P "realiza" M en un sistema s , P debe especificar una propiedad micro-estructural de s que proporciona un mecanismo causal para la implementación de M en s ; más aún, en los casos interesantes, -de hecho, si hablamos con sentido de "implementación" de M - P será un miembro de la familia de propiedades físicas que formarán redes de estados micro-estructurados y nomológicamente conectados, que proporcionan un mecanismo micro-causal.⁴⁰

Para Kim la relación de realización o instanciación no es explicativa; más bien los estados del micro-sistema forman una base explicativa para las propiedades supervenientes y las relaciones nomológicas entre ellas. No debe confundirse la relación de explicación con la relación de realización. La relación explicativa que también puede formar parte de la relación de realización no es constitutiva de ella. Para que la propiedad M (propiedad mereológica) sobrevenga en la micro-propiedad es nomológicamente necesario que si un objeto s tiene la micro-propiedad entonces, s tiene M . Pero para que se dé la posibilidad de que esas propiedades mereológicas nuevas tengan una fuerza causal respecto de otras propiedades del mismo nivel, sería una condición suficiente i) que haya una ley (posiblemente no estricta) que conecte a las propiedades de ese nivel y ii) que en cada caso en el que se dé la relación causal en ese nivel, de manera correspondiente, haya micro-propiedades de las cuales sobrevienen, que a su vez se relacionan causalmente.⁴¹

Davidson, de hecho, a pesar de compartir con Kim la idea de que las propiedades mentales sobrevienen en las propiedades físicas, diferiría de la postura de Kim en que no aceptaría que para que las propiedades mentales tengan poderes causales es necesario establecer leyes para ese nivel que correspondan a las leyes físicas. Esas correspondencias permiten establecer vínculos entre tipos de propiedades. En ello radica la diferencia central entre la postura de Kim y de Davidson. Davidson no acepta ni leyes ni relaciones entre

⁴⁰ Ibid., 197.

⁴¹ Cfr. Sober & Segal (1990).

tipos de propiedades para el nivel mental.

Veamos cómo articula Davidson su postura. Davidson suscribe las siguientes tesis: a) que la relación causal puede mantenerse como relación entre sucesos, b) que se acepten propiedades mentales como un tipo de relaciones supervenientes y c) que no se acepten leyes a nivel psico-físico o psicológico.

Se señala con frecuencia que en la posición de Davidson se aprecia una tensión generada por su caracterización de la causalidad en la cual se defiende una causación física única, con su subsiguiente soporte en leyes con la defensa de la injerencia causal de lo mental. Esta tensión se daría en todos los casos, exceptuando las relaciones causales que recogen los sucesos en la terminología que corresponde a las ciencias físicas básicas y sus correspondientes leyes. Todas las otras descripciones que aluden a propiedades o relaciones que no sean micro-propiedades estarían en la misma situación que las propiedades mentales.

Antes de entrar a las objeciones veamos primero -de manera muy apretada- lo que sostiene Davidson en cuanto a las diferencias entre las relaciones causales y las explicaciones causales.

Pese a que Davidson considera que la relación causal es una noción primitiva o inanalizada que se establece entre sucesos particulares, también asume, sin mayor argumento, que la relación causal supone la existencia de leyes. Y, como ya dijimos, la relación causal es una relación binaria que se obtiene, independientemente de cómo se describan o se recojan esos sucesos, esto es, es una relación extensional. La relación causal es una relación fundamental para Davidson en el sentido de que una caracterización completa del universo estaría dada en esos términos y que son "suficientes para formular un sistema cerrado de leyes "estrictas".⁴² Así, estas leyes incorporarían sólo cierto tipo de propiedades que serían de carácter físico.

Entre las muchas posibles maneras de aludir a los sucesos físicos es empleando

⁴² Davidson (1993) Cfr. nota 10, p.17.

términos mentales. El empleo de términos mentales no obedece a la presencia de algún tipo de suceso especial, sino que, ontológicamente hablando, sólo hay sucesos físicos. Para Davidson, pese a la identidad entre los sucesos físicos y los sucesos mentales, debido a la ausencia de leyes, no permite la reducción de las propiedades mentales a las físicas. La relación se establece entre instancias o ejemplificaciones ("token"), y no entre tipos de propiedades. Llegamos así a su conocida tesis del monismo anómalo según la cual hay una reducción ontológica de lo mental a lo físico, más no una reducción conceptual, y por ende, explicativa. Dice Davidson: "Las entidades mentales (objetos y sucesos particulares sujetos a espacio y tiempo) son entidades físicas, pero los conceptos mentales no son reducibles, por definición o por ley natural a los conceptos físicos."⁴³ Los sucesos mentales tales como percepciones, recuerdos, decisiones, y acciones se resisten a ser capturados en la red nomológica de la teoría física. Para Davidson todas las generalizaciones que pudieran apoyar a las explicaciones que se expresan en términos mentales, o que combinaran términos mentales con términos físicos, distan mucho de ser leyes estrictas; esto es, no hay leyes psicológicas o psico-físicas. En suma, Davidson deriva la tesis del monismo anómalo de tres tesis:

- 1) Los sucesos mentales se conectan causalmente con sucesos físicos.
- 2) Las relaciones causales singulares se apoyan en leyes estrictas.
- 3) No hay, ni puede haber, leyes psico-físicas estrictas.

Davidson, movido por las muchas críticas y malentendidos que ha suscitado su postura, ahora incluye como una premisa adicional a las anteriores, la doctrina de la superveniencia. Define la relación de la siguiente manera: "No puede haber dos sucesos iguales en todas sus aspectos físicos pero que difieran en algún aspecto mental", o alternativamente, "un predicado p es superveniente en un conjunto de predicados S si y

⁴³ Davidson (1993), p. 3.

sólo si *p* no distingue ninguna entidad que no puede distinguir *S*.⁴⁴ Las propiedades mentales dependen de las propiedades físicas pero no se reducen a ellas, de modo que si se modifican las propiedades mentales también se modifican las propiedades físicas base o micro-propiedades.

Dada la dependencia de la causación mental de sus descripciones, sus respectivos sucesos se especificarían a través de sus contenidos, generando contextos incluíblemente intensionales.⁴⁵ Así, las relaciones causales que refieren a lo mental nos remiten al ámbito de las explicaciones. Las explicaciones, por su parte: a) relacionan oraciones, b) son intensionales, ya que su verdad depende de la manera en que se describen los sucesos involucrados, c) son sensibles a intereses; d) el término "explicación" es un concepto epistemológico, que como tal es un concepto intencional, que tiene como propósito producir en quien inquiera un entendimiento de la situación. Así, establece Davidson que los sucesos que se vinculan causalmente admiten de múltiples explicaciones, tantas como aspectos que nos interesen de los sucesos. Las explicaciones psicológicas, además de las características anteriores, dependen de los sucesos físicos que se vinculan causalmente. La dependencia es tal que, dada la relación de superveniencia, si se diera alguna alteración en las propiedades mentales de un suceso, conllevaría una alteración en sus propiedades físicas. Empero no podríamos inferir que el mismo cambio se daría en otros sucesos.⁴⁶ Así, el ámbito de lo mental no constituye un sistema cerrado y no puede contar con leyes estrictas que lo gobiernen; sus leyes contarían con muchas excepciones y cláusulas *ceteris paribus*. En innumerables escritos Davidson reitera que para que haya apoyo en leyes,

⁴⁴ Ibid. p. 4.

⁴⁵ Algunos autores como Kim distinguen entre propiedades intrínsecas y propiedades relacionales y explican la causación mental sólo en términos de las primeras. Excluye las propiedades relacionales, como las propiedades-contenido. Las propiedades relacionales dependerían, en parte, del hecho de que el agente esté integrado en su medio ambiente. Algunos autores como von Gulick y Baker, entre otros, cuestionan la distinción entre los dos tipos de propiedades. Cfr. Baker, (1993).

⁴⁶ Cfr. ibid. p.7.

éstas deben de ser leyes estrictas.

Pasemos ahora a las críticas. ¿Es justa la acusación de que la postura de Davidson conlleva una inconsistencia entre sus diferentes tesis? Entre estas críticas cabe mencionar las de Churchland, Stich, Fodor, Kim, Sosa y Follesdal.

Una manera de expresar la dificultad -según algunos de estos autores- es que la caracterización de la superveniencia que Davidson defiende es insuficiente para sostener que sus propiedades contribuyan causalmente. Me parece que lo determinante para que la crítica surta efecto dependerá de qué piensa Davidson de las propiedades -sean físicas o mentales- y la manera en que la referencia a ellas reverbera en su tesis causal ontológica. Si como dijimos antes, las propiedades se asocian con las descripciones, es claro que Davidson no haría depender la verdad de los enunciados causales en ellas. Pero esto no lo obliga a negar que a través de las propiedades se produzca el vínculo causal. Como bien dice Kim, es perfectamente consistente mantener una tesis de la causalidad, como relación binaria y extensional que se produce entre sucesos con una tesis que diga que las propiedades generan y fundamentan esas relaciones entre sucesos, de manera que "para hablar del papel de las propiedades en la causación no necesitamos introducir la locución "qua" o cualquier otra relación causal con múltiples términos."⁴⁷ Pese a la separación de estas dos tesis, Kim juzga que en las respuestas a sus críticos Davidson parece confundirlas, como por ejemplo cuando dice: "dado mi punto de vista extensionalista de las relaciones causales, no tiene literalmente sentido ... hablar de que un suceso causa a otro en tanto mental, o en virtud de sus propiedades mentales, o en tanto descrito de una manera u otra."⁴⁸ No me queda claro que Davidson sea víctima de la confusión que le señala Kim, sobre todo si la acusación se sustenta en citas como la anterior. Lo que sí me parece claro es que pondríamos un énfasis equivocado en esta cita de Davidson si

⁴⁷ Kim (1993), pp. 22-3.

⁴⁸ Davidson *ibid.*, p. 13.

suponemos que tanto las propiedades mentales como las físicas son causalmente impotentes, y que el asunto es directamente entre sucesos, como hace McLaughlin.⁴⁹ Dado que esta interpretación que hace McLaughlin a Davidson aparece en muchos otros autores, me parecen aclaratorias las siguientes palabras de Davidson:

Si se da la superveniencia, las propiedades psicológicas hacen una diferencia para las relaciones causales de un suceso, ya que importan para las propiedades físicas, y las propiedades físicas importan para las relaciones causales.⁵⁰

Me parece que el acento debe ponerse -como ya se ha señalado en varias ocasiones a lo largo del trabajo- en que la relación causal se sostiene independientemente de las descripciones que se empleen para referirse a los sucesos así relacionados. Me parece que Davidson tiene razón cuando reitera que esta aserveración no lo compromete a sostener la superfluidad causal de lo mental y menos de lo físico. Davidson sólo se resiste a hacer aparecer tanto las propiedades físicas como las mentales en la relación causal misma; hace dispensable la introducción de más términos en la relación causal.

Si aceptamos que ni en el caso físico ni en el mental la referencia a las propiedades del suceso convierte al enunciado causal en intensional ¿cómo se reconcilia con la tesis de que la alusión a las propiedades mentales en la causación mental es inevitable? Ya mencioné que es inevitable el vocabulario mental para juzgar los sucesos físicos como acciones: a nivel físico no podemos hablar de acciones. En este sentido es que Davidson dice que la causación mental es intensional, que equivale a decir que la causalidad mental es explicativo-causal. Compárese el caso mental al caso de un disparo ruidoso en donde lo ruidoso funciona como las propiedades mentales. Lo ruidoso sobreviene al disparo como sobrevienen otras propiedades específicamente mentales. Lo ruidoso en ciertos casos repercute causalmente y en otros no, por ejemplo cuando el ruido del disparo produce en

⁴⁹ Cfr. McLaughlin (1993).

⁵⁰ Davidson (1993), p.6.

un sujeto x un desmayo o un paro cardíaco. Como vimos en la sección anterior, Kim no tiene las resistencias que tiene Davidson para introducir una multiplicidad de términos en su propio análisis causalista y en este caso el ruido aparecería en el enunciado causal. En lo particular nosotros, en acuerdo con Davidson, no vemos ninguna motivación de peso que nos induzca a hacerlo, antes bien lo contrario: queremos permitir que se establezcan enunciados causales verdaderos aun cuando se deje de señalar o se desconozca en virtud de cuáles de sus propiedades. Parece ya letanía, pero su verdad depende de que los sucesos se fijen o apresen correctamente.

Al margen de si Davidson confunde la cuestión del número de términos que se relacionan causalmente con la cuestión del papel causal que desempeñan las propiedades, de cualquier forma Kim no queda convencido de los intentos de Davidson de reconciliar las tesis arriba mencionadas. Kim persiste en que la tesis de que los fenómenos mentales pueden ser causas o efectos de fenómenos físicos (la premisa 1) es inconsistente con la tesis del monismo anómalo. Brevemente, la crítica de Kim a Davidson es que su posición lo enfrenta a un dilema que lo obliga ya sea a que lo mental no desempeñe ningún papel causal o, a renunciar a la idea de que no puede haber leyes psico-físicas (premisas 2 y 3). En cualquiera de los casos la "anomalía" desaparecería: si se rechaza 1) tendríamos un monismo físico en donde lo mental se convierte en su epifenómeno de lo físico y si se rechaza 2) se renunciaría a la anomalía de lo mental. Por la breve exposición que hicimos de la propuesta de Kim, él evita el dilema negando que lo mental sea anómalo y admitiendo leyes en el nivel mental.

¿Se puede sintetizar el desacuerdo entre Davidson y sus críticos, como a veces lo expresa Davidson, como una cuestión de si es correcto hablar de leyes no estrictas o si se reserva el derecho a llamarse "leyes" sólo a las estrictas? En la lectura de Kim, por ejemplo, si Davidson admite que las propiedades mentales tienen una contribución causal, con base en su relación de superveniencia y la existencia de leyes no estrictas, tendría que sacrificar la anomalía de lo mental; la relación causal se da, ya que se trata de sucesos de

cierto tipo o a que tienen ciertas propiedades.⁵¹ Si no fuese así, considera Kim, se tendría que tomar a la relación causal como un "hecho bruto" de los sucesos, posición que él rechaza. Para Kim, la tesis de la superveniencia -a no ser que se tome como lo que él denomina una superveniencia "débil"- supone la existencia de leyes psico-físicas, ya que de no ser así, las generalizaciones carecerían de la necesaria fuerza nómica que tendría que acompañar a la relación causal. Indefectiblemente, la posición de Davidson, lleva, según sus críticos, a que lo mental no haga ninguna aportación causal y que ahora se señala la dificultad con el siguiente dilema: o las leyes no estrictas tienen la fuerza nómica necesaria para lograr apoyar a las relaciones causales, o carecen de fuerza nómica y entonces la superveniencia es muy débil para poder sustentar la contribución causal de lo mental. En el primer caso, además de que Davidson tendría que justificar la premisa de por qué se aferra a una idea de causalidad vinculada a leyes estrictas, se agregaría la complicación, -en la cual no me detendré- que Kim denomina "el problema de la exclusión".⁵²

Mi apreciación de este debate me lleva a darle la razón a Davidson en que la relación causal no convierte a la causación en un "hecho bruto" sólo por el hecho de que en su estructura lógica no se incluya la referencia a las propiedades responsables. Uno puede seguir fiel a la idea de que su forma es más básica que la forma que incorpora la referencia a las propiedades de los sucesos en cuestión y, en sí, es completa. Pero también

⁵¹ Cfr. Kim (1993), p. 22.

⁵² La tesis dice que los sucesos que interaccionan causalmente tendría dos propiedades, irreducibles, pero que en la medida en que la propiedad mental causa el suceso físico, en virtud de la relación causal que se da entre las propiedades físicas básicas de los sucesos, le robaría a la propiedad mental toda función causal. (pp. 25-26) En "Thinking Causes", en relación con el problema de la exclusión, separando la cuestión de la causalidad de la de explicación, Davidson se dirige a un trabajo anterior de Kim, "The Myth of Nonreductive Materialism" y sostiene que en el caso de las causas, una causa suficiente, no tiene por qué excluir otras causas independientes; más bien como señaló Mill, incluye a toda lo que precede en el universo al efecto y que tiene en él una injerencia causal. Un suceso-causa incluye, como partes propias, todos los sucesos mentales pertinentes. En tanto sucesos individuales, la existencia o inexistencia de leyes psico-físicas es pertinente cuando se trata de tipos de sucesos.

tiene razón al decir que no hay leyes que amparen los casos causales singulares, ya que se carece de la sistematicidad requerida. Por ello es que Davidson habla de relaciones entre instancias de propiedades y no de propiedades. Si hemos decidido aceptar con Davidson trasladarnos al dominio de las explicaciones cuando hablamos de lo mental, podemos dar un paso adicional que Davidson no da y tomemos a las explicaciones causales singulares -sean mentales o físicas. sin apoyarnos en leyes.

2.3 Causación racional

Cuando hablamos de propiedades mentales, nos hemos trasladado a la provincia de las explicaciones. La tesis de que una explicación física agote todas las explicaciones y, por tanto, que elimine a las otras, resulta una tesis insostenible. Ninguna explicación física, por más completa que sea, compite siquiera con las explicaciones que se dan en los dominios fuera de la física, y menos aún las excluye. Si aceptamos que la causación es una relación física puede haber muchas explicaciones aptas asociadas a un mismo suceso. Sólo las explicaciones "suficientes", completas, o sea las explicaciones de las ciencias físicas básicas, podrían proponer leyes aunque no participen como parte del modelo explicativo. En este sentido todos los discursos, incluyendo el físico, están en la misma situación que el discurso mental.

Paso ahora a la otra manera de atacar la postura de Davidson, que también incide en que las propiedades mentales no tienen ningún peso causal. En esta crítica se señala que Davidson finalmente independiza el problema causal de las razones dejando a lo físico todo el peso causal y a las razones les asigna otro tipo de criterios. Muchos críticos diagnostican que la dificultad la produce la combinación de que las explicaciones causales se divorcien de las relaciones causales, con el uso de una noción única de causación. En efecto, si sólo existe la causalidad física y los únicos sucesos que hay son físicos, en ese nivel no existiría la diferencia entre aconteceres y haceres. Para dar cabida a esta diferencia sólo queda que hablemos de propiedades más complejas (físicas también o

neurológicas) en el caso de lo mental. Si, alternativamente, no hay separación entre explicaciones causales y relaciones causales los críticos considerarían que las diferencias entre las explicaciones causales físicas y las mentales conlleva una diferencia entre tipos de causación. En la postura de Davidson las consideraciones causales y explicativas se asumen de manera complementaria, ya que la relación causal se encarga de restringir el número de posibilidades que las consideraciones explicativas dejan abiertas. En la propuesta de conjuntar los criterios en uno solo exigiría que también se renunciara a la separación entre relaciones causales y explicaciones causales proponiendo o un vínculo causal particular que no sea el simple vínculo causal físico, o proponiendo algún criterio adicional para las explicaciones psicológicas. En la primera vertiente se han propuesto nociones tales como "causalidad racional", "causalidad mental" o causalidad razonable". En la segunda, se habla de una explicación normativa, estableciendo una clara diferencia entre ellas y las explicaciones nómicas. Lo común a ambas vertientes es que se destaca de manera particular que en ellas desempeña el contenido en los respectivos estados y en las acciones. Discutiré más ampliamente esta propuesta en el apartado de las explicaciones intencionales.

II. MODELOS DE EXPLICACION INTENCIONAL

3. Posturas ontológicas y epistemológicas de la explicación

En el capítulo anterior señalé que iba a considerar separadamente la cuestión de las relaciones causales de la cuestión de las explicaciones. En este capítulo me referiré a los modelos de explicación y examinaré las propuestas más comunes para explicar las acciones: el modelo disposicional de Hempel, el modelo de explicación causal singular de Woodward, el modelo intencional de Davidson y el modelo biológico-funcional de Millikan.

Los dos primeros modelos -el de Hempel y el de Woodward- difieren básicamente en cuanto al papel que se les concede a las leyes -el modelo disposicional considera que las leyes son esenciales a la explicación mental y el modelo de explicación singular, prescinde de ellas. Los modelos intencional y biológico-funcional suelen entender los patrones de conducta intencional como patrones normativos y en los cuales se sostiene que los principios normativos funcionan de manera distinta de los principios nómicos. En el modelo biológico-funcional que aquí examinaré, que también se aplica a lo mental, se caracteriza por el modo en el que introduce los constreñimientos normativos, a saber, en relación con ciertas funciones que se constituyen evolutivamente. Y a partir de estas funciones específicamente para el modelo mental se alude a otra forma de normatividad -su sentido más usual- en la cual se alude a principios de inferencia y de racionalidad. En este capítulo defenderé una versión del modelo biológico-funcional para lo mental.

Entre los filósofos de la ciencia que se han interesado en el problema de la explicación se entabla un desacuerdo básico en cuanto a si la demanda de una explicación debe erguirse primordialmente como una cuestión ontológica o como

una cuestión epistemológica y/o pragmática. Esta diferencia de enfoques persiste aún en quienes no necesariamente los consideran excluyentes entre sí. Salmon, en sus escritos más recientes, concede que los desarrollos en cada una de estas posturas hace posible que quizá se puedan reconciliar, pero no al punto de permitir que el acercamiento entre las posturas se encauce en la dirección de que únicamente cuenten las consideraciones pragmáticas.¹ Friedman y Kitcher, han intentado llevar a cabo un proyecto de unificación entre la explicación ontológica y la epistemológica (o inferencial) como la que suscribe Hempel. En cualquier caso, lo importante es asumir que no son posturas necesariamente enfrentadas.

La noción de explicación -como noción epistemológica- guarda cierto vínculo con la noción -también epistemológica- de justificación de nuestras creencias. Aún en nuestra vida cotidiana decimos que no es lo mismo explicar que justificar. Una manera de trazar la diferencia, intuitivamente, es que podemos entender por qué sucedió algo, esto es, explicar por qué sucedió, sin estar dispuestos a aprobarlo o compartirlo. En algunas tesis filosóficas se conserva esta liga entre explicación y entendimiento y se distingue de la noción de justificación. Como bien señala Villoro, por ejemplo, con la explicación se intenta "dar razón de un hecho" y con la justificación se trata de suministrar "un fundamento de legitimidad"; el primero pertenece al discurso de lo que es y el segundo a lo que debería ser.² El contraste apunta hacia una diferencia entre una función descriptiva y una función normativa (epistémica), respectivamente. Con la justificación se presenta un patrón argumentativo en el cual se exhiben las razones en las cuales se apoya la verdad de cierta creencia o hipótesis y que entre sus propósitos incluye los de valorar

¹ Cfr. Salmon (1990a) pp. 180-86 y (1990b).

² Cfr. Villoro (1993).

positivamente esa creencia y pretender que otros también lo hagan. En una explicación, en cambio, partimos de la proposición y buscamos explicarla, ofreciendo un patrón en el cual encajarla. Las explicaciones tienen entre sus tareas la de organizar y sistematizar nuestros conocimientos.

¿Qué podemos decir, más en específico, de las explicaciones intencionales? En primera instancia, hay más consenso que en el caso de las explicaciones científicas en aceptar que en las explicaciones intencionales intervienen factores pragmáticos y, por lo general, se consideran como explicaciones causales. Como veremos más adelante, muchas discusiones en torno a las explicaciones intencionales han derivado en el problema de cómo interviene en ellas el rasgo causal. La postura de Davidson ha estado particularmente expuesta a este tipo de críticas. Ya aludimos a algunas de ellas en el capítulo anterior.³

La presencia de las consideraciones pragmáticas en las explicaciones nos obliga a considerar su fuerza ilocucionaria. La referencia a los actos de explicación resolvería dos problemas claves a los cuales se enfrentan las posturas que los excluyen, a saber, el problema de no distinguir los productos de las explicaciones de otros productos ilocucionarios y el problema del énfasis -al cual aludimos al hablar de la extensionalidad de la relación causal singular-, el cual puede alterar el significado de lo que se expresa en la explicación.⁴ En el capítulo final me detendré en el valor de la pragmática en relación con la discusión que lleva a cabo Wittgenstein en torno a los enunciados de creencia. En esos casos la voz desde la cual se enuncia la creencia puede cambiar considerablemente su sentido, al igual que las circunstancias en las que se hace el enunciado. Lewis hace la interesante

³ Cfr. pp. 51-8.

⁴ Cfr. Achinstein (1989), cap. III.

observación de que el hecho de que las explicaciones se nutran de consideraciones contextuales no es específico de los actos explicativos sino que es característico de cualquier acto que maneja información en general: todos ellos varían según su acceso, selección, transmisión y uso y dependen por igual de los informantes y los receptores. Los actos informativos y explicativos son satisfactorios o insatisfactorios por las mismas razones. A veces están al alcance y comprensión de unos, a veces de otros. Hay variaciones considerables tanto en los comunicadores como en los receptores.⁵ En una frase, la explicación no sería más que un modo de manejar o manipular la información.

3.1 Modelo nomológico deductivo

Los autores que suscriben la tesis de que todas nuestras explicaciones - independientemente del tipo de explicación de que se trate (científica, del orden común, en el ámbito social, económica, tecnológica, etc.), valen para las explicaciones de acontecimientos singulares al igual que para las generalizaciones. Hempel es uno de los más connotados autores que ha defendido la idea de un modelo general para todas las explicaciones "por qué". De los trabajos de Hempel, sin lugar a dudas, emanan casi todas las discusiones en torno al problema de la explicación, sobre todo de la explicación científica. De todos sus trabajos hay que destacar el trabajo seminal que elaboraron Hempel y Oppenheim en 1948, "Studies in the Logic of Explanation". En él, sus autores proponen dos modelos de explicación científica: el modelo nomológico deductivo (N-D) y el probabilista o inductivo (I-S) en sus dos variantes: el estadístico-inductivo (I-S) y el estadístico-deductivo (D-S). Tanto en los modelos N-D y I-S se trata de explicaciones de

⁵ Cfr. Lewis (1993).

fenómenos singulares o leyes, mientras que en el modelo D-S se explican regularidades. Pocos son los filósofos de la ciencia que, al menos de manera explícita, se hayan preocupado por hacer extensivo el tipo de modelo(s) de explicación científica, a los ámbitos humanos y sociales. En ese sentido Hempel ha sido una excepción que no deja de tener méritos.

Para Hempel las explicaciones, en su aspecto formal, son argumentos - deductivos o inductivos- que tienen poder predictivo y que hacen que los fenómenos que se describen en el *explanandum* sean esperables ("*expectability-criterion*"). Las explicaciones son argumentos en los cuales se subsumen los fenómenos en leyes generales y como tales, fungen como razones suficientes para la obtención del fenómeno a explicar. El punto nodal de su propuesta es que las leyes cargan con el peso explicativo. Mediante el modelo N-D se mostraría que el fenómeno que se quiere explicar (*explanandum*) era de esperarse, dadas ciertas circunstancias particulares, en conjunción con alguna ley general. Así, dice Hempel:

La explicación singular implícitamente afirma que hay leyes generales -digamos L_1, L_2, \dots, L_n en virtud de las cuales la ocurrencia del antecedente causal mencionado en C_1, C_2, \dots, C_k es una condición suficiente para que ocurra el suceso del *explanandum*.⁶

Hempel considera que su modelo cubre no sólo a las explicaciones causales singulares, sino también a las explicaciones que más frecuentemente se encuentran en la ciencia, a saber, aquéllas en las que sus leyes establecen relaciones entre propiedades. Para Hempel, ésta es la forma más común que toman las leyes científicas, que no excluye que a veces las leyes o explicaciones sean causales. En

⁶ Hempel (1965), p. 365.

cuanto a las explicaciones intencionales, Hempel considera que se ajustan al modelo de explicación inductivo, a través de la aplicación de la teoría de las decisiones. Más aún, Hempel considera que el modelo disposicional que es el más conviene para explicar las decisiones y acciones humanas. Atribuirle a un sujeto propósitos, creencias, rasgos de carácter, etc., equivale a atribuirle un haz complejo de disposiciones. Hempel esquematiza el argumento de la siguiente forma:

(C₁) C estaba en una situación del tipo S

(C₂) *i* tiene la propiedad M

(L) Cualquier *x* que tiene la propiedad M, en una situación del tipo S, se comportará de manera R

(E) *i* se comportó de manera R

M se refiere, en parte, a la totalidad de disposiciones específicas que la teoría asigna al término.⁷ Las creencias y fines vislumbrados en respuesta a los cuales el agente actúa de manera típica son, a su vez, rasgos disposicionales, por ello Hempel las llama "disposiciones de jerarquía superior".

En el caso de las explicaciones intencionales me parece que Davidson tiene razón al señalarle a Hempel que su modelo de explicación intencional es equivocado. Davidson ofrece dos razones: la primera, que no recurrimos a las generalizaciones empíricas porque aunque existieran, por lo general quienes ofrecen las explicaciones las desconocen. La segunda razón, es que las generalizaciones respecto a la manera en que la gente actúa en determinadas condiciones no intervienen directamente. Las generalizaciones no hacen que nuestras atribuciones

⁷ Hempel (1978), p. 141.

en un caso particular sean más fuertes, "pero las hacen más valiosas, permitiéndonos acomodarlas en patrones más generales."⁸ Para Davidson las leyes implícitas -en todo caso- serían leyes respecto a sujetos particulares, mientras que para Hempel serían generalizaciones respecto a sujetos semejantes en situaciones semejantes. Si bien discrepan las posiciones de Davidson y Hempel respecto de las explicaciones de las acciones, en el caso de las explicaciones científicas estarían básicamente de acuerdo. En el siguiente apartado, pasaré a la crítica de Woodward al modelo N-D de Hempel inclusive como modelo científico de explicación. Esta crítica también se aplicaría a la postura de Davidson.

3.2 Modelo de explicación causal singular

Woodward rechaza al modelo N-D de explicación y, en su lugar, defiende un modelo de explicación que da prioridad a los casos singulares, o sea un modelo que los toma como primitivos. Woodward condensa la crítica al modelo N-D en los siguientes puntos: a) Los explananda de algunas explicaciones causales singulares difieren en estructura de los explananda de las explicaciones N-D, por lo que ese modelo no puede dar cuenta de ellos; b) Si las explicaciones causales singulares se ajustaran al modelo N-D las leyes tendrían que tener una estructura que, por lo general, no tienen; c) Quienes ofrecen las explicaciones usualmente desconocen la información que les permitiría apoyar la explicación del suceso particular en la ley, lo que debilita al modelo N-D; d) El modelo N-D es inconsistente con la intuición respecto al papel que desempeñan las identidades de los sucesos particulares en las explicaciones causales singulares. Esta última va dirigida más específicamente a la posición de Davidson.

De estas cuatro tesis de Woodward, la segunda y la tercera se destinan contra

⁸ Davidson (1980), p. 275.

la idea de un criterio puramente formal para las explicaciones. Empezando por la tercera crítica, Woodward reclama que el criterio formal dejaría fuera cuestiones epistemológicas que es importante incorporar. La consideración epistemológica que señala Woodward es esencialmente la que propone Davidson, a saber, que si se asume que las explicaciones cotidianas son buenas explicaciones, y no habría razón para no suponerlo, quienes las ofrecen generalmente ignoran la estructura nomológica de las explicaciones de la cual la explicación singular supuestamente es una instancia.⁹ En la otra crítica al aspecto formal del modelo N-D -la segunda- cuestiona que su estructura obliga a Hempel a que las leyes hagan referencia a sucesos -sea como sucesos individuales o como una clase de sucesos- cuando las leyes, por lo general, no hacen referencia a sucesos sino a propiedades y a relaciones entre propiedades. No entraré aquí en la discusión de la forma que tienen las leyes básicas y si éstas prescinden de las propiedades causales o si las leyes pueden tomar una forma cuyo consecuente fuera existencial. Sólo mencionaré la dirección de la crítica. Como ejemplo típico de las leyes que se utilizan en la ciencia Woodward menciona el siguiente:

Las leyes científicas generalmente relacionan propiedades -nos dicen que si un objeto de un sistema posee ciertas propiedades, poseerá (o, tal vez, habrá una probabilidad definida de que posea) ciertas otras propiedades. Las leyes científicas nos dicen, por ejemplo, de la relación entre el largo de un péndulo simple, la fuerza del campo gravitacional en el que se localiza, y su periodo.¹⁰

Woodward no da argumentos para sostener que las leyes no

⁹ Cfr. Woodward (1986), pp. 233 y 235 y (1984).

¹⁰ Ibid. p. 262.

puedan tener una estructura cuyo consecuente sea existencial, excepto unas indicaciones generales. Una dificultad es que tanto las leyes existenciales como las leyes que establecen relaciones entre propiedades satisfacen los mismos criterios: el que sean generalizaciones y que se apoyen condicionales contrafácticos. Woodward sugiere, aunque no lo desarrolla, que las afirmaciones existenciales pudieran tal vez entrar como "afirmaciones acerca de condiciones iniciales". Pero entonces se precisarían requisitos diferentes que desempeñarían los diferentes papeles que juegan las leyes y las condiciones iniciales. Woodward sugiere que las leyes satisficieran los requisitos de cierta simetría e invariabilidad que no tendrían por qué cumplir las condiciones iniciales.¹¹ Tal vez una dificultad en esta línea que propone Woodward sería que la distinción que se necesitaría para sustentar la diferencia entre leyes y condiciones iniciales es difícil de trazar ya que la decisión entre qué es lo que entra como condición inicial y qué como parte de la ley es en buena medida una cuestión heurística.

En cuanto a la primera crítica -de que el modelo N-D no cubre a todas las explicaciones-, Woodward se ampara en el hecho de que en algunas explicaciones singulares queremos saber por qué ocurrió un determinado suceso, mientras que en otras explicaciones nos interesa saber por qué se dio con ciertas características específicas. Sostiene que la limitante en el modelo N-D es que todas las explicaciones se tratan como si fuesen del segundo tipo, i.e. como si fuesen explicaciones de por qué tienen las propiedades que tienen y a los sucesos como instancias de un tipo de suceso. Woodward dirá que en el modelo hempeliano el *explanandum* de "El corto circuito causó el fuego violeta", que se simbolizaría como " $(!x) (Fx \cdot Vx)$ " y que de igual manera se simbolizaría el *explanandum* de la

¹¹ Woodward (1986) nota de pp. 265-6.

oración "La presencia de sales de potasio causó que el fuego fuera violeta".¹² Esto es, que nos interese saber qué suceso causó qué otro suceso o que nos interesen sus propiedades se simbolizarían de la misma manera. Ya me referí a esta cuestión en la sección 1.2¹³

En síntesis, allí expliqué que Woodward defiende que debemos distinguir el caso de por qué ocurrió un suceso en lugar de ocurrir algún otro, en cuyo caso nos concentramos en el suceso *tout court* y el caso en el cual nos referimos a las propiedades del suceso. Ya dijimos que si bien es correcto distinguir que a veces nos interesa saber qué suceso provocó qué otro, no era correcto el uso pretendidamente semántico que Woodward empleaba para la distinción referencial/atributivo de la descripción definida. Woodward sustenta su propuesta de los dos tipos de explicaciones en la distinción semántica entre usos referenciales y atributivos. Y, la referencia a la clase de contraste puede justificarse o bien con el uso semántico o con el uso pragmático de la distinción. Woodward ilustra los dos tipos de explicación de fenómenos singulares con el siguiente ejemplo:

b) El corto circuito explica causalmente por qué se dio el fuego violeta

y

b') La presencia de sales de potasio explica causalmente por qué se dio el fuego violeta

Al ocurrir el fuego, podría interesarnos por qué ocurrió el fuego o por qué

¹² Cfr. *Ibid.* p. 261.

¹³ pp. 27-31.

se dio el fuego con cierta propiedad. En el primer caso al interesarnos en por qué ocurrió el fuego, presuponemos que el fuego es violeta, por ejemplo. En el segundo caso nos interesaría saber por qué el fuego es violeta y presupondríamos que se trata de un fuego. La presencia de las sales de potasio explica por qué el fuego tiene el color que tiene pero no explica por qué el fuego mismo ocurre. Y a la inversa, el corto circuito explica causalmente por qué se dio el fuego y presupone que es violeta. En realidad la idea de la presuposición aplicado a un enunciado explicativo no es más que la idea que mencionamos en el capítulo anterior, de que ponemos un énfasis en algún aspecto del suceso como guía para saber el tipo de explicación que se persigue. Como tal, sólo reforzaría el uso pragmático de la distinción referencial/atributivo. En el primer caso -continúa Woodward- "El corto circuito" y "El fuego", en la oración "El corto circuito causó el fuego" son descripciones definidas, cuyo correcto uso exige que en los alrededores sólo haya habido uno de cada uno y se quiere explicar la ocurrencia de ese fuego y no de algún otro. Sin embargo, no se pretende explicar por qué hay un y sólo un fuego, sino que se presupone. Cuando decimos que la ocurrencia del corto circuito produjo o causó el fuego, queremos explicar por qué un cierto fuego, el que se recogió mediante la descripción definida "el fuego" ocurrió y no por qué no hubo otros fuegos en el contexto de interés. Para Woodward, el uso correcto de la descripción definida "el fuego" presupone que sólo un fuego ocurrió en el contexto que nos concierne, pero que no todo lo que sigue de ello o se presupone en el uso de la descripción definida, se explica.¹⁴ Woodward podría haberlo dicho más sencillamente, diciendo por qué se dio ese suceso (en lugar de decir "el fuego").

Vimos que los enunciados causales singulares son extensionales y que la

¹⁴ Woodward (1986), p. 259.

diferencia entre los usos referencial y atributivo en conexión con las descripciones definidas, más que marcar como Woodward quiere, una diferencia semántica, es sólo una diferencia pragmática que no justificaría el que se hablara de dos tipos distintos de enunciados explicativos (o de dos tipos de enunciados causales). Para marcar la diferencia pragmática, bastaría con la alusión a la clase de contraste. Y respecto de la clase de contraste, Woodward está en su derecho de introducir la cláusula contrastiva con el objeto de restringir las sustituciones de términos en los enunciados explicativos, garantizando la extensionalidad de los contextos explicativos que él quiere suscribir. Pero como también sugerí en la sección 1.2 no veo ninguna motivación para decir que algunos enunciados explicativos se comportan como los contextos causales y convertir a ambos en extensionales. Para Woodward las explicaciones de sucesos singulares simplemente consistirían en el reporte de una relación causal. Contrariamente a lo que sostiene Woodward, podemos señalar que la referencia a la clase de contraste es indicativo de que estamos esencialmente interesados en la descripción del suceso y esa dependencia es lo que hace que las explicaciones sean intensionales.

Me parece que si Woodward no logra sustanciar la diferencia entre los dos tipos de explicación, tampoco sería correcto que simbolice a ambas como (!e) (Fe & Ge). Pero me parece que con el dispositivo de la clase de contraste puede muy bien dar cuenta de la diferencia entre pedir una explicación respecto de qué suceso provocó el efecto y el caso en el que se busque la explicación de por qué despliega cierta propiedad. No veo tampoco por qué Woodward no acepta, como lo hace Davidson, que en un enunciado causal como "El corto circuito causó el fuego violeta", "El fuego violeta" podría muy bien simbolizarse "El fuego violeta" como (E!e) Fe. En el enunciado causal sí queremos que el suceso se simbolice como un suceso singular o único. No tenemos por qué aceptar que hay dos formas de

enunciados explicativos singulares.

Paso ahora a la crítica de Woodward en cuanto al apoyo que reciben los casos explicativos singulares de las leyes, según los modelos de Hempel. Esta crítica, a diferencia de la anterior, me parece válida. Me parece que si combinamos esta crítica con la idea de que las leyes científicas no suelen referirse a sucesos singulares ni a clases de sucesos, desbrozamos el camino para dar cabida a que los casos singulares se tomen como primitivos. Optar por el modelo de las explicaciones causales singulares -como en Woodward- equivale a sostener que la invocación de las leyes correspondientes no confieren fuerza explicativa a los fenómenos. Con ello no se quiere decir que no haya regularidades en el mundo o que no haya leyes estrictas que regulen ciertos fenómenos: es posible, y tal vez necesario, pensar que las hay. Lo único que se niega es que en virtud de las leyes se consiga la explicación de un fenómeno particular. Más aún, podríamos agregar a la consideración de Woodward que, como sostiene Lewis, aun cuando concedamos que en la práctica la búsqueda de explicaciones no está desvinculada de la investigación acerca de las leyes "podemos preguntarnos si la información acerca de las leyes es en sí misma parte de la información explicativa."¹⁵ A esta interrogante Lewis da un respuesta negativa. Si explicamos causalmente la ocurrencia de b invocando la ocurrencia de a, el conocimiento de la ley pudo haber facilitado la elección de la causa de la ocurrencia de b, pero el desconocimiento de la ley, sea porque no la hay o porque la ignoramos, no invalida la explicación. Ciertamente en las explicaciones se echa mano de la información antecedente que poseemos y las leyes podrían formar parte de esta información antecedente, pero la explicación no pende ni de que conozcamos esas leyes, ni de que existan.

¹⁵ Cfr. *ibid.*, p. 206.

Woodward considera que la característica que da la fuerza explicativa a los enunciados causales singulares es la misma para todas las explicaciones. Considera que aun cuando fuera cierto que las explicaciones causales singulares se apoyan en leyes, éstas formarían parte de sus condiciones de verdad pero, que de allí no se sigue que las afirmaciones causales singulares sean explicativas proporcionando información contenida en esas leyes.¹⁶

Para Woodward, la fuerza explicativa de cualquier explicación radica en "la identificación de condiciones tales que cambios en estas condiciones harían una diferencia respecto a si se diera el fenómeno a explicar o una alternativa que se especificara." Se afirma que la presencia del corto circuito constituye una diferencia para la ocurrencia del fuego, manteniendo inalteradas las otras condiciones iniciales en las que ocurre el fenómeno. Con el corto circuito, la presencia del fuego, se hace al menos más factible de lo que hubiera sido sin el corto circuito. Esta intuición la recoge con el contrafáctico: Si no se hubiera dado el corto circuito, en las circunstancias en las que ocurrió, no habría ocurrido el fuego, siempre dejando de lado las dificultades que ofrece la sobredeterminación.¹⁷ Esta propuesta de Woodward es correcta si sostenemos que en el caso indeterminista, las circunstancias no hacen más factible que se diera la situación *vis a vis* que se diera otro suceso.

La conclusión en esta sección es que, compartiendo el punto de vista de Woodward, es más adecuado el modelo que da prioridad a las explicaciones causales singulares, de que su fuerza explicativa no deriva de las leyes, aunque las haya. Y de manera más radical que la que propone Woodward, podríamos

¹⁶ Ibid., p. 268.

¹⁷ Cfr. *ibid.* p. 236.

considerar que debe incluirse la dimensión pragmática en este tipo de modelo más allá del expediente de la clase de contraste, a la manera como sugiere Lewis, que ya mencionamos con anterioridad.¹⁸ La clase de contraste se introduce no porque exista una ambigüedad en el enunciado, sino por razones pragmáticas. Si damos prioridad a las explicaciones causales singulares y agregamos los *desiderata* que Woodward propone, contrariamente a lo que él sostiene, no hemos quitado todos los obstáculos que impiden cubrir a todas las explicaciones bajo un mismo modelo. En particular, no se ha dado cuenta del sello normativo que caracteriza a lo mental. La propuesta a la que quiero llegar diferencia las explicaciones causales singulares físicas y mentales introduciendo la normatividad de lo mental a través de la normatividad que se encuentra en la especificación de ciertas funciones biológicas evolutivas. Pasemos entonces a los diferentes modelos de explicación para lo mental que se distinguen entre sí y con los modelos que hemos venido discutiendo por los diferentes maneras en que ubican el requisito normativo.

3.3 Modelo de explicación intencional: normas de racionalidad

Los diferentes modelos de explicación intencional caen en dos grupos básicos, que reflejan diferentes convicciones respecto a cómo entender los contenidos mentales y en cuanto al papel que éstos desempeñan en las explicaciones. Hasta ahora, nos hemos referido únicamente a las posturas que no han reparado específicamente en la noción de contenido y simplemente hablan de las propiedades mentales preocupándose únicamente de garantizar que desempeñen un papel explicativo-causal genuino o, al menos, autónomo. En este tipo de posturas caben la de Kim (para quien la diferencia discursiva se expresa en

¹⁸ p. 62.

términos de diferentes niveles y de leyes correspondientes a cada nivel), la de funcionalistas como Fodor y la de Hempel. Hay otras posturas que hacen descansar la especificidad del modelo explicativo intencional precisamente en su referencia a los contenidos de las actitudes mentales o proposicionales y, a través de ellos, introducir cierta forma de normatividad. Las dos tendencias dominantes confrontan dos nociones de contenido: una, de corte lingüístico o semántico y la otra una noción de contenido no lingüístico (como por ejemplo las percepciones), que tiene que ver con la manera en que los contenidos internos a un sistema se constituyen relacionándose externamente con aspectos del medio ambiente. Esta idea se contrapone a la idea de contenido -asociada a las representaciones internas en los sistemas- constituida de manera interna como la que defienden los funcionalistas clásicos.

Como resultado del balance de las posiciones en pugna, me inclino por la idea de contenido no semántico, en particular, en la forma promovida por algunos defensores de la teoría de la evolución. Esta postura se defenderá en la siguiente sección. En esta línea se defiende la idea de que con el desarrollo de las capacidades lingüísticas emerge la idea de contenido que desplegará características semánticas. La noción de contenido, en tanto que tiene una función representacional, se identifica relacionadamente conectándose con aspectos del entorno. Las representaciones lingüísticas serían, según algunas formas de funcionalismo biológico sólo un tipo de representación -la más importante para los seres que tenemos las capacidades para formárnoslas-, que por expresar sus contenidos en forma lingüística tienen una estructura más compleja que las otras formas de representación. Los contenidos, tanto lingüísticos como no lingüísticos forman parte de ciertos estados mentales, de modo que hay que hacer alusión a ellos al identificar o caracterizar esos estados. Los contenidos, además de

desempeñar un papel representacional tienen la función de intervenir, en combinación con otros estados mentales, en la producción de las acciones. De ser correcta esta postura, incorporaría en la noción de función biológica una noción de normatividad que vendría a jugar un papel crucial en conexión con las explicaciones de lo mental. En esta sección me referiré a la primera solución, protagonizada por Davidson, y en la siguiente sección me referiré a la idea de contenido no lingüístico, en la forma en la que la introducen quienes la asocian a las funciones biológicas evolutivas.

Dentro de las posturas en las que la referencia al contenido de los estados mentales es ineliminable, algunos consideran que la dimensión normativa que entrañan las explicaciones en las que aparecen puede naturalizarse, sin cometer algún tipo de falacia naturalista. Para entender el debate conviene que advirtamos que la noción de normatividad puede entenderse de diferentes maneras, o introducir sus constreñimientos en diferentes sitios. El sentido de normatividad más comúnmente asociado a las explicaciones intencionales deriva del patrón lógico que corresponde a la articulación racional entre las actitudes mentales y las acciones. Thalberg, Brand y Harman, más que introducir otro sentido de normatividad, hablan de constreñimientos en las cadenas causales, que se requieren para evitar que se generen cadenas causales degeneradas o desviadas. Las cadenas causales intencionales desviadas son importantes ya que pueden aparecer pese a que se cumplan con los requisitos normativos racionales. Una cadena causal desviada es aquella en la que un agente tiene creencias, deseos e intenciones apropiadamente descritos que hacen razonable que haga X, éstas lo llevan causalmente a hacer X, sin embargo el agente no hace X intencionalmente. Así, los constreñimientos normativos en las cadenas causales mismas pretenden garantizar que las cadenas causales sean correctas. Otros autores, más que ubicar los constreñimientos

normativos en la cadena causal lo hacen en el sujeto. Para estos autores sólo entonces sus acciones podrían propiamente llamarse intencionales. El sujeto tiene la capacidad de planear o proponerse ciertos fines y diseñar ciertas estrategias para conseguirlos. La normatividad emana de sus intenciones, del compromiso que contrae de llevar a cabo ciertas acciones: su intención de hacer p supone que necesariamente actúe en los términos establecidos en su compromiso.¹⁹ A la normatividad racional se suma otra normatividad heredada de la noción de función biológica, a la manera como la entienden autores como Millikan, Dennett, McGinn y Papineau, entre otros. En esta postura se defiende una forma de normatividad asociada a procesos naturales. En estos procesos se detecta cierto patrón teleológico o intencional, en cuya constitución se introducen contenidos no lingüísticos. Se establece una cierta semejanza entre los diseños o patrones que son producidos por la "Madre Naturaleza" y los patrones de conducta intencional. Como haré ver, se puede hablar de una conducta intencional en los organismos más simples que los nuestros, que tiene sus raíces en los patrones conformados evolutivamente, por selección natural. En organismos más complejos como nosotros, en donde evolutivamente desarrollamos capacidades lingüísticas, aparece otra forma de normatividad, la denominada normatividad racional. Con esta nueva normatividad, los estados mentales adquieren contenidos muy específicos y puntuales, que alcanzan los márgenes de precisión que nuestro lenguaje permite.

La crítica que se hará al modelo intencional racional, desde el cuartel de los evolucionistas, es una crítica a la dependencia de lo mental en lo lingüístico; uno de los propósitos de los propugnadores del modelo biológico funcional para lo intencional es el de forzar ese vínculo, trazando una línea de continuidad evolutiva

¹⁹ Cfr. por ejemplo, Moya (1990).

entre las dos formas de contenido. Empecemos entonces con la posición que defiende Davidson, en donde se suscribe una noción de contenido lingüístico y en la cual la normatividad se asocia únicamente a los principios de inferencia y de racionalidad.

i) Davidson

Vimos que Davidson defiende un modelo de explicación psicológica distinto del modelo(s) de explicación física. La diferencia consiste no sólo en que las explicaciones utilizan descripciones que corresponden a sus respectivas disciplinas, en que sólo en las ciencias físicas las explicaciones singulares son respaldadas por leyes estrictas. Ya hablamos de esto en la sección anterior. En esta sección me referiré a la línea crítica en la cual se acepta el contraste entre los modelos de explicación física y mental, pero se rechaza la forma en que la defiende Davidson. La crítica atribuye la falla al hecho de que se admita sólo un tipo de causalidad, la física, para hablar de lo mental. En las explicaciones físicas, la explicación causal física procede tomando al suceso en cuestión como una instancia de una ley, en donde la ley expresa las condiciones suficientes para que se dé el suceso. Por el contrario, en la explicación causal psicológica se recurre a principios de racionalidad que no nos proporcionan ni las condiciones necesarias ni las condiciones suficientes que expliquen por qué se dio el suceso. Esta situación hace imperioso que la explicación se complementen recurriendo a los sucesos singulares que participan en la relación causal. Davidson considera que buena parte de la fuerza explicativa de las explicaciones por razones consiste en especificar cuál par, de los muchos posibles pares de creencias y deseos que podrían haber sido adecuados a la acción, en los hechos intervino causalmente. Si hubiera leyes estrictas, como en el caso de las explicaciones físicas básicas -dice Davidson-, la

los enunciados causales singulares se inferiría de ellas. Para Davidson se necesita la conexión con las causas físicas por la falta de leyes adecuadas. Escribe Davidson: "La ausencia de una ley del tipo adecuado, hace esencial que reparamos en la relación causal, ya que la creencia y el deseo pudieran estar presentes y la acción desarrollarse y sin embargo que la creencia y el deseo no expliquen la acción."²⁰ Por otra parte, agrega Davidson, que en toda explicación además de señalar cuáles son los sucesos físicos adecuados que se relacionan causalmente se requiere que se recojan descriptivamente de manera adecuada. La característica de las explicaciones -se trate de explicaciones físicas o mentales- es la dependencia de la especificación de propiedades o de descripciones para su verdad. Pero, si lo causal no interviniera, no estaríamos distinguiendo entre justificación y explicación. La justificación se daría si racionalizáramos cuáles serían las creencias y deseos apropiados que legitimaran la acción. Pero lo que legitima una acción y lo que de hecho la produce pueden no corresponder. De allí que en algunos de estos casos se hable de irracionalidad, como veremos en el siguiente capítulo.

El embate a la postura de Davidson pretende hacerse extensivo a todas las posturas que defiendan una teoría explicativa causal de la acción intencional, como lo serían por ejemplo la de Churchland y la de Goldman²¹, con base en que su falla sería estructural. La crítica corre en términos de que en las explicaciones por razones -las explicaciones en términos de deseos y creencias- se divorcian del aspecto causal, ya que las razones no causan las acciones en virtud de que las racionalizan o justifican, sino en virtud de que son una instancia de una ley; y las razones racionalizan las acciones en virtud de que admiten descripciones que son

²⁰ Davidson (1987), p.42.

²¹ Cfr. Churchland (1979) y Goldman (1971).

cóherentes con la manera en que el agente apreció su acción.²² Para que no hubiera tal divorcio -continúa la crítica- el contenido de los estados tendría que tomar parte causalmente, debido a que en esos términos se promovió la acción. Sin embargo, en las teorías en las que se admite sólo la causación física, el contenido es inoperante causalmente, en tanto que se convierte en un epifenómeno de los sucesos físicos. Si uno admite que la separación entre lo normativo y lo causal en las explicaciones psicológicas es fatal para el modelo, se podría buscar enmendar la situación de varios modos. Algunos eligen resolver la falla, proponiendo un modelo en el cual se fusionaran los aspectos causales y los normativos, dando lugar a un tipo de causación distinto al físico. Así, surgen posturas en las que se habla de una "causalidad mental", o de una "causalidad racional" o de una "causalidad razonable". Pero subsisten quienes no conformes con esta salida optan por modificar el modelo de explicación intencional. La justificación de la modificación se encuentra en que los constreñimientos que establece la normatividad racional es insuficiente para hablar de las explicaciones de las acciones intencionales. La pauta para la búsqueda parece encontrarse investigando lo que ocurre en las cadenas causales desviadas. En esta sección me referiré en particular a la forma en que Moya ofrece fusionar los dos aspectos. El problema, según Moya, es consecuencia de una característica constitutiva del modelo causalista, a saber, que el vínculo entre la causa y el efecto es "externo" o fáctico, cuando para dar cuenta de las acciones intencionales es menester recurrir a un tipo de vínculo interno o conceptual (y por tanto necesario) entre las razones y las acciones.

Lo que caracteriza a la postura de Moya es que conectará el divorcio de los requisitos causales y normativos con la presencia de las cadenas causales desviadas.

²² Cfr. Moya (1990), p. 112.

Esta liga, a la vez, lo llevará a buscar un criterio de normatividad que erradique la posibilidad de que se generen ese tipo de cadenas anómalas. Si interpreto a Moya correctamente, él sostiene que la posibilidad de las cadenas causales desviadas debe eliminarse, lo cual se lograría con la normatividad que propone. Una cadena causal intencional desviada es aquella en la que un agente tiene creencias, deseos e intenciones apropiadamente descritos que hacen razonable que haga X, éstas lo llevan causalmente a hacer X, pero en donde el agente no hace X intencionalmente. La manera que propone consiste, en primer lugar, que en el vínculo se incluyen, además de las creencias y deseos, las intenciones y, en segundo lugar, que se establece un vínculo interno (y por tanto necesario) entre las intenciones y las acciones. Cuando Moya habla de "intenciones" las entiende sobre todo en el sentido de "intenciones futuras". En su lectura, en estas intenciones futuras el agente tiene la intención ahora de realizar la acción en un tiempo futuro si, y sólo si, se compromete a realizarla, en los términos especificados en su intención. En su postura, lo intencional se erige como criterio general de lo que es ser un agente o actor.

Moya sostiene que lo que muestran las cadenas causales desviadas es que las condiciones de racionalidad y de causalidad no son independientes, sino que las razones han de causar la acción bajo ciertas descripciones si la acción ha de ser intencional.²³ Un presupuesto fuerte de su postura es que la posibilidad de las cadenas causales desviadas no sólo afecta a esos casos desviados, como si se tratara de casos atípicos, raros o periféricos sino que afecta incluso a las cadenas causales normales.²⁴ Si pensamos en lo que generalmente sucede cuando se producen

²³ Moya (en correspondencia).

²⁴ Cfr. cap. 11.

cadena causal desviada, en algún punto la cadena se rompe o interrumpe -por lo común debido a algo externo o independiente de ella- pero sin embargo se llega azarosamente a donde habría llegado de haber seguido su curso normal. En el caso específico de las acciones intencionales Moya supone que la independencia que genera la desviación es una cuestión de base, derivada de que los requerimientos para una cadena causal exitosa están desvinculados, de modo que su falla es estructural. Sin embargo, de ser correcto el diagnóstico de Moya en cuanto a que la falla es de origen, su solución tendría la consecuencia de que se excluirían muchas acciones que normalmente no dudaríamos en calificar de intencionales.

Para apreciar mejor la discusión veamos un caso de una cadena causal perceptual y comparémosla con el caso intencional. Una cadena causal perceptual desviada sería aquella en la cual un sujeto tiene una experiencia *X*, causada por *X*, pero que no es producida por la ruta perceptual normal. Como ejemplo podríamos poner el caso de una persona que bajo el efecto de algún alucinógeno, tiene la experiencia de un gato que se encuentra en su entorno, el cual a través de su olor le produce una imagen del gato que corresponde con la situación real.²⁵ En este caso, no se trataría de un caso de percepción aunque el olor lo haya desprendido el gato y aunque la imagen sea una réplica fiel de la escena real. Me parece que Moya describiría la asimetría entre la cadena causal desviada en el caso perceptual y en el caso intencional en términos de que en el caso perceptual el vínculo entre lo que se percibe y la experiencia que se produce es un vínculo externo, mientras que en el caso de la relación entre la intención y la acción es interna. Un ejemplo de cadena causal intencional desviada sería el siguiente. Carlos quiere matar a su abuelo y sabe que él invariablemente, con una puntualidad meridiana, sale a

²⁵ Este ejemplo es paralelo a los casos perceptuales desviados que ofrece Peacocke en Holistic Explanation, Clarendon Press, Oxford, 1979.

comprar el periódico a determinada hora. Un día, Carlos toma el auto con el objeto de llevar a cabo su plan. En el camino, un anciano penosamente va cruzando la calle y Carlos, nerviosamente, y con el propósito de no retrasar su fijado plan, arrolla al anciano. Sin percatarse, Carlos mata a su abuelo. En este caso, no podríamos decir que la acción de Carlos de matar al abuelo fue intencionada, a pesar de que eso es lo que se disponía a hacer. Así, estamos frente a una cadena causal desviada ya que la acción no es intencional, pese a que existen el deseo y la creencia correspondientes (i.e., con los adecuados contenidos proposicionales). Para Moya, en este caso, se genera la cadena causal desviada, ya que no basta con que los contenidos proposicionales de los deseos y las creencias sean adecuados, sino que además se necesita que la descripción adecuada forme parte del contenido de la intención; sólo entonces se convierte en una guía para la acción y sólo entonces nos comprometemos a hacer verdadero ese contenido realizando la acción adecuada.²⁶ Otro ejemplo de una cadena causal desviada sería la de alguien que está haciendo una suma que como resultado de dos errores llega al resultado correcto. En este caso tampoco diríamos que la persona llegó al resultado correcto intencionalmente.

Sin embargo, si es correcto lo que he venido sosteniendo, no es tan claro que la diferencia entre la cadena causal desviada y la cadena causal normal no tenga que ver con algo que el sujeto ignora y el error del sujeto consistir en no haber previsto una situación que iba a interferir en su plan y que, por tanto, la diferencia no pueda explicarse en términos de cadenas causales correctas e incorrectas. Uno podría mantener que tanto en el caso perceptual como en el intencional hay un paralelismo, independientemente de si el vínculo causal es interno o externo. En

²⁶ Cfr. p. 58.

ambos casos se viola un principio básico, de modo que si en el caso perceptual las rutas equivocadas no nos llevan a modificar lo que se entiende por percepción ¿por qué habría de requerirse la manera de entender las intenciones en el caso mental? Si en el caso perceptual, lo desviado trae como resultado que la experiencia no se pueda considerar una percepción, en el caso intencional una acción tampoco sería intencional descrita de cierta manera, pero sí lo sería descrita de otra manera. En síntesis, mis dudas tienen que ver con el supuesto escollo insuperable que generan las cadenas causales desviadas que nos empuja a buscar un modelo de normatividad que garantice que, de haberse dado la acción intencional, tuvo que darse de manera correcta, que nos hace pensar en una idea de normatividad rígida como si se trataran de "rieles mentales" impuestos en las intenciones.²⁷ Por otra parte, si se hiciese requisito de todas las explicaciones de las acciones intencionales que se introdujese el compromiso -sea implícito o explícito- me parece un modelo muy intelectualizado de las acciones intencionales. Exigir el elemento deliberativo o reflexivo para dar cuenta de las acciones intencionales, convertiría buena parte de nuestras conductas irreflexivas en no intencionales, contrariamente a nuestras intuiciones ordinarias.

Paso ahora a estrategia de Moya que consiste en tomar como punto de partida el paradigma de normatividad que según él se encuentra en las "actitudes intencionales puras". Estas acciones intencionales no inmediatas o futuras, según Moya, comparten el mismo rasgo que tienen los verbos como "prometer", "casarse", "saludar", "ordenar", etc., a saber, que de haberse realizado, *necesariamente*

²⁷ Respecto de esta idea de reglas como "rieles", que Wittgenstein considera equivocada, véanse *Las Investigaciones Filosóficas*, #218 y #219.

implican el cumplimiento de lo que describen. A estas acciones las considera "puras" ya que "la acción -bajo esa descripción- sólo pueden desempeñarse intencionalmente".²⁸ Las acciones "puras" no implican ningún suceso o acontecimiento físico de un tipo específico. Dice así: "Una intención debe entenderse como un tipo de compromiso a actuar y la acción intencional como una conducta que respeta ese compromiso y lo haga de manera correcta."²⁹ Moya propone que las intenciones de realizar una acción futura requiere entender la normatividad a la manera en que Wittgenstein entiende la relación entre "seguir una regla" y sus aplicaciones, a saber, que el entendimiento mismo de la regla conlleva su aplicación correcta. Así, para Moya, una acción es intencional (bajo una descripción D) "si y sólo si, al realizar esa actividad o conducta, bajo esa descripción, A (el agente) sigue correctamente su intención racional".³⁰

Además de estas "intenciones futuras" o a largo plazo, que Moya denomina intenciones en sentido fuerte, también habla de intenciones en un sentido débil o mínimo. En este último se da un patrón propositivo de conducta, empero inmediato. La intencionalidad inmediata se puede aplicar a ciertas conductas de ciertos seres del reino animal y a seres humanos que todavía no han alcanzado un cierto grado de desarrollo. Las intenciones futuras, nos dice Moya, dependen y surgen de las intenciones en sentido mínimo, pero revisten una forma propia que las hace cualitativamente distintas a ellas. La forma propia de las intenciones futuras sería -supuestamente- la normatividad que encarnan. La normatividad en las intenciones deriva del compromiso que el sujeto contrae al proponerse hacer algo.

²⁸ Cfr. p. 52.

²⁹ Ibid. pp. 4-5.

³⁰ Cfr. p. 130.

Con las acciones intencionales "puras", Moya pretende evitar caer en el argumento regresivo que amenaza a las posiciones en las cuales los sucesos físicos son condición necesaria para cualquier acción. Moya pretende así fincar una distinción irreductible entre los sucesos físicos y las acciones intencionales sin necesidad de recurrir a la postulación de "acciones básicas" como lo serían los actos volitivos. Moya conecta la superación de la causalidad física, con el repudio a la naturalización de lo intencional a lo físico. Me parece cuestionable esta liga entre la causalidad y el problema de la naturalización de lo intencional, pero dejaré pasar esta cuestión aquí.

Si entiendo bien la postura de Moya, podríamos distinguir dos tipos de normatividad que conectarían los contenidos proposicionales de los diferentes estados mentales y las acciones. Una normatividad -que acompaña a ciertas actitudes proposicionales- que diversos autores han denominado "causalidad racional", "causalidad mental" o "causalidad razonable", en la cual los contenidos proposicionales tienen una injerencia causal en el desarrollo de la acción. Y otra normatividad asociada específicamente a las intenciones en relación con las acciones, en sentido fuerte, que además de depender de las relaciones de contenido se requiere que el agente reconozca y asuma el compromiso de actuar con base en ellas. Me parece que la diferencia crucial entre la normatividad asociada a las intenciones y la normatividad que pudiera encontrarse en las otras actitudes como los deseos y las creencias estribaría en que con las intenciones el sujeto pueda convertir a las razones en objetos de reflexión y de evaluación y, como resultado de ello, perseguir su realización.³¹

Podríamos aceptar con Moya que es tan incorrecto dar cuenta de las acciones

³¹ Cfr. pp. 164-9.

intencionales refiriéndonos sólo a los deseos y las creencias del individuo, como entender las intenciones, como pro-actitudes o como formas de creencias. En ese sentido, podríamos estar de acuerdo con Moya en que la revisión que el propio Davidson hace de las intenciones respecto a la postura que suscribe en "Actions, Reasons and Causes" resultaría insuficiente. En su revisión Davidson acepta la inclusión de intenciones como razones -junto con las creencias y los deseos- pero las considera, al igual que a los deseos, como pro-actitudes: se toman finalmente como disposiciones a actuar de ciertas maneras bajo ciertas circunstancias. Para Davidson la intención consiste en "el juicio evaluativo incondicional de que cierta acción es más deseable que otra alternativa disponible."³² Según Moya, tomar a las intenciones como un tipo de pro-actitud supone seguir preso del modelo causalista físico. Pero podríamos conceder que las intenciones no deben asimilarse a otros tipos de actitudes y sin embargo resistirnos al desglose que Moya hace del aspecto normativo que conllevan las intenciones.

Considero que la distinción wittgensteineana entre "seguir una regla" y "estar de acuerdo con una regla", que en Wittgenstein tiene que ver con su discusión acerca de cómo ciertas reglas tienen que ver con su entendimiento o con su gramática, de modo que de aplicar la regla mal sistemáticamente equivaldría a no entenderla. Tomar las ideas de seguir una regla y la normatividad que conlleva no conducen a Wittgenstein a que se eliminen las equivocaciones o los errores, pese a que se conozca la regla y el sujeto se refiera a ella al aplicarla.

Es un hecho patente que a veces las acciones intenciones futuras revisten formas en las que se revela un compromiso explícito de llevar a buen puerto la acción con la cual nos comprometimos y, como bien señala Moya, el compromiso

³² Cfr. "Intending" en *Essays on Actions and Events*, Oxford University Press, 1980.

no es como un capricho que se tiene y que puede romperse con la misma facilidad, sin que a la vez otras de nuestras actitudes proposicionales sufran transformación. No es obvio que aún en los casos en los que existe un compromiso previo la acción intencional, modelarlo a la idea wittgensteiniana de "seguir una regla" resulta poco iluminadora para efectos explicativos. Si bien podemos aceptar que las intenciones futuras tienen una potencia eficiente propia, distinta a la de las otras actitudes proposicionales, el desacuerdo proviene de que se tenga que renunciar al modelo causal de explicación de las acciones y se les asocie en cambio un tipo de normatividad en el sentido interno que Moya sugiere. Me parece más promisorio la vía de buscar una normatividad en un sitio que Moya no explora, a saber, en el ámbito de las funciones biológicas seleccionadas evolutivamente.

Davidson a la vez que suscribe el modelo nomológico de explicación para lo físico, admite un modelo de explicación racional para lo mental. En esta última las explicaciones se hacen sin estar respaldadas en leyes estrictas y sirven a propósitos o intereses distintos a los de las explicaciones científicas. Davidson se ve forzado a este doble estándar de las explicaciones. En lugar de ello, podríamos prescindir del modelo nomológico de explicación y trazar la diferencia a partir del modelo de explicación causal singular.

Podemos aceptar con Davidson que la referencia a los sucesos mentales es indispensable para explicar las acciones humanas. Al hablar de acciones, dice Davidson, las razones pueden explicar la acción "cuando damos las razones por las que el agente hizo lo que hizo". Podemos estar de acuerdo con denominar a estas explicaciones "racionalizaciones" y que son una especie de explicación causal. Los sucesos mentales tienen la peculiaridad de no generar una ontología, ya que son los mismos sucesos físicos, empero descritos empleando un vocabulario psicológico. Esta dimensión nos permite ver a las personas como agentes racionales, con

**ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA**

propósitos y fines, y sujetos de evaluación moral. Si le restamos esta dimensión a la explicación, queda el suceso físico, como algo que pasa y en donde tampoco queda acción alguna. Escribe Davidson: "La conducta consiste en cosas que hacemos, sea con intención o no, pero donde hay conducta, la intención es relevante. En el caso de las acciones, la relevancia puede expresarse de esta manera: un suceso es una acción si y sólo si puede describirse de manera que lo haga intencional." Espero que en el modelo biológico funcional haya una compatibilidad entre aceptar un modelo explicativo-causal singular sin que se diluya la diferencia entre hacer y acontecer. No tenemos que renunciar a las ideas claves que están a la base de lo que denominamos acciones. Para ello tenemos que suponer que podemos hablar de una intencionalidad que abarque una región del mundo más extensa que la humana, sin que atente contra nuestro sitio humano privilegiado en el mundo y sin que se diluya en el mundo la manera tan decisiva y única en que nuestras acciones lo han transformado.

ii) Dennett: Instrumentalismo

Dennett defiende una postura muy cercana a la de Davidson respecto de las explicaciones intencionales, a la cual tiene el acierto de agregar la dimensión biológico-funcional. Pero la postura de Dennett también tiene un desacierto: el realismo instrumentalista con el que caracteriza a los estados mentales que se invocan en las explicaciones intencionales. En esta sección me referiré a su acierto. Uno de sus logros es haber introducido una estratificación en nuestras explicaciones que nos permiten ver a los objetos o sistemas desde muy diferentes perspectivas. Estas diferencias, sin embargo, no suponen perder de vista que las explicaciones están subordinadas a nuestro diseño biológico que, a su vez está subordinado a los mecanismos físicos del objeto o sistema.

En Dennett la noción de intencionalidad que caracteriza a la mayoría de los estados mentales cobra un sentido muy particular que básicamente recoge en lo que él llama "la perspectiva o estrategia intencional". Dicho brevemente, con esta estrategia Dennett esencialmente pretende preservar el uso de las nociones mentales de la psicología popular, que en su teoría revisten un valor meramente instrumental y legitimar la alusión a estados mentales, apoyándose en una teoría biológica de la evolución y en una idea ligada a ésta: la racionalidad ideal.

Dennett considera que su teoría emana de la aceptación de la tesis quineana de la indeterminación de la traducción, la cual lo compromete, por un lado, a rechazar la existencia de entidades mentales y, por el otro, a aceptar que en las explicaciones intencionales se legitima -heurística o instrumentalmente- la alusión a ellas. El empleo de las nociones mentales de la psicología popular en Dennett tiene la función de hacer postulaciones que nos permiten predecir exitosamente la conducta del objeto al cual le hacemos las atribuciones.

Dennett recurre a la idea de que existen diversas perspectivas que nos dan una dimensión distinta desde la cual apreciar los fenómenos y explicar su comportamiento. Así, distingue las perspectivas física, del diseño, intencional y personal³³, aunque en la mayoría de sus escritos se refiere a las tres primeras.

Cada perspectiva supone la adopción de la postura de la tercera persona, esto es, la del observador, el cual puede explicar desde diferentes perspectivas el comportamiento de un objeto o *sistema* -como prefiere llamarlo Dennett. Según los propósitos e intereses que animen nuestras explicaciones se asociarán éstas a una u otra de las diferentes posturas o estrategias y su validez dependerá, de manera

³³ Cfr. (1987a) Dennett introduce la estrategia "personal", que depende de la estrategia intencional y agrega a ella un elemento moral. La dependencia de la estrategia personal en la intencional contrasta, según Dennett, con las otras estrategias las cuales son independientes entre sí.

significativa, de su valor predictivo. Así, un sistema puede verse como un sistema físico, si nos preocupa su constitución; desde la perspectiva de su diseño si nos interesa su funcionamiento; y puede verse intencionalmente, atribuyéndole metas y considerando que su comportamiento es racional.

En la postura física se requiere que el observador de un sistema pueda "determinar su constitución física, la naturaleza física de los impactos (en el sistema), y el uso del conocimiento de las leyes físicas para predecir el producto ("output") de cualquier insumo ("input")".³⁴

Los objetos considerados desde la óptica de su diseño, se consideran como sistemas no sólo físicos, sino además como sistemas que bajo ciertas circunstancias, operan o funcionan de cierta manera. Así, nos referimos al mecanismo del objeto, el cual puede describirse con diferentes grados de abstracción y precisión.

Mediante la estrategia intencional atribuimos al sistema una serie de elementos que pretenden explicar, predecir e interpretar su conducta, i.e. la racionalizan. Consiste en "atribuir al sistema, la posesión de cierta información, suponer que está dirigido hacia ciertas metas para posteriormente elaborar la acción más razonable o apropiada sobre la base de estas atribuciones y suposiciones."³⁵ O como dice en "True Believers":

Primero decidimos considerar al objeto cuya conducta queremos predecir como un agente racional, después calculamos las creencias que debería tener, dados su lugar en el mundo y su propósito. Después calculamos los deseos que debería tener, según las mismas consideraciones y finalmente predecimos que este agente racional, a la luz de sus creencias, actuará para lograr sus metas.³⁶

³⁴ Dennett (1987b), p. 16.

³⁵ Ibid. p.6.

³⁶ Dennett (1987a) p. 17.

Dennett considera que a cada una de estas perspectivas le corresponden sus propias descripciones, no sólo en el sentido de que tienen un vocabulario propio, sino también en el sentido de que hay un aspecto real que se recoge, y que sólo puede recogerse, a través de su perspectiva correspondiente. La sugerencia de Dennett es que la misma realidad se puede "ver con diferentes ojos", lo que en los casos particulares de las perspectivas del diseño e intencional consiste en captar algún aspecto adicional o novedoso de la realidad. Ese "algo más", en el caso intencional, sería un patrón de conducta que obedece a la relación que guardan entre sí los elementos del sistema respecto a su entorno. Dennett reclama que su postura sería realista, realista moderada -acotaría él-, ya que los patrones de conducta que emergen en el punto de vista intencional son, para Dennett, objetivos y no dependen de cómo los vea o los quiera ver el observador: no serían aspectos "proyectados" por nosotros. Habría, sin embargo, una dependencia del observador en tanto que algunas personas tendrían más habilidad que otras para discernir un cierto patrón de conducta. Dennett pretende hacer convivir esta forma de realismo con su instrumentalismo. En el nivel intencional y aplicado particularmente a la creencia dice:

Así como la creencia es un fenómeno totalmente objetivo (lo cual me hace un realista), sólo puede discernirse desde el punto de vista de alguien que adopta una cierta estrategia predictiva y su existencia puede confirmarse sólo mediante la evaluación del éxito de la estrategia (que aparentemente me convierte en un interpretacionista).³⁷

Gran parte de las críticas que le han hecho a Dennett cuestionan el éxito de

³⁷ Dennett (1987a), p.15.

esta propuesta reconciliadora, que supone sostener armónicamente tesis que a primera vista son opuestas entre sí. Tocaré este problema cuando en el capítulo siguiente vuelva a la postura de Dennett y me refiera a la discusión del estatuto ontológico del estado mental de la creencia.

Dennett defiende que no todos los aspectos de la conducta puedan ser cubiertos por una sola perspectiva y, en función de los nuevos aspectos que surgen en las diferentes perspectivas se dará la relativa autonomía que Dennett establece entre las explicaciones intencionales con respecto de las físicas. Haciendo una comparación con el juego de la vida (John Horton Conway) para ilustrar la idea de que quizá haya algunos patrones objetivos que sólo pueden discernirse a través del punto de vista intencional, dice Dennett:

Esos patrones son objetivos -están allí para detectarse- pero desde nuestro punto de vista no están allí afuera de manera totalmente independiente de nosotros, ya que son patrones que se componen en parte de nuestras propias reacciones 'subjetivas'.³⁸

Dennett considera que el sistema procede de acuerdo a cierto patrón o diseño, aun cuando no tenga un diseño definido. El diseño específico dependerá de la función que se le otorgue al sistema, de cómo se emplee. Al respecto, Dennett pone el ejemplo de una máquina diseñada para surtir refrescos a cambio de una moneda, digamos, "un cuarto de dólar americano", que igualmente podría utilizarse para surtir refrescos a cambio de "un cuarto de balboas panameño", como "detector-K", como "arma mortal", como "tope de una puerta", etc.³⁹ Y así como la descripción de un diseño es compatible con un número indefinido de realizaciones físicas, una

³⁸ Dennett (1987b), p. 39 y (1991).

³⁹ Cfr. Dennett (1987c), p. 292.

descripción intencional es compatible con un sinnúmero de diferentes descripciones del diseño. En la perspectiva intencional, para Dennett, se podrían hacer las mismas atribuciones intencionales a sistemas que no compartieran ni el mismo diseño o programa, ni la misma constitución interna. Pero al hacer la atribución tampoco compartirían ningún "hecho" ("*fact of the matter*"), lo concebimos como proceso, estado o propiedad. Sería -nos dice Dennett- como si tuviéramos varias interpretaciones sin un texto subyacente a ellas. Dennett se resiste a la idea de construir algún "hecho" para que tengan sentido nuestras atribuciones. A pesar de ello, agrega Dennett, esto no constituye un obstáculo para decir que las atribuciones tienen un valor de verdad. Podemos reconocer en estas descripciones de Dennett la asimilación de la tesis de la indeterminación de la traducción de Quine.

Dennett sugiere defender que esta estrategia intencional igual puede adoptarse frente a cualquier objeto que a nuestros semejantes, siempre y cuando al hacerlo, como ya se dijo, se obtengan dividendos predictivos. Así, un sistema que se comporta intencionalmente se distingue del que no se comporta intencionalmente en que en el segundo, la predicción nunca resulta. Un sistema no intencional carecería de patrón o diseño, es decir, su proceder resultaría azaroso.⁴⁰ Pero este criterio, ¿acaso no enfrentaría a Dennett a la dificultad de no poder discernir los casos en los que la predicción no se cumple debido a que la interpretación del observador es falsa, de aquellos en los que la adopción de la estrategia intencional es equivocada? La salida viable para Dennett podría ser que en el caso de un sistema no intencional, el comportamiento del sistema se puede predecir adecuadamente, sin pérdida alguna, desde otra de las perspectivas. Es decir, que la adopción de la perspectiva intencional no agrega nada predictivamente. En el caso

⁴⁰ Cfr. Dennett (1991), p. 30.

de una interpretación falsa esto no sucedería, de allí que se continúe con la búsqueda de otra interpretación, tratando de detectar el lugar del error. Pero, por otra parte, son tan pocos los requisitos que Dennett impone, que se enfrenta a la dificultad de no poder discernir los casos en los cuales la predicción no se cumple debido a que la interpretación del observador es falsa de aquellos casos en los que la adopción de la estrategia intencional es equivocada. Muchos autores le han hecho esta crítica a Dennett, por ejemplo Stich⁴¹ y Bennett⁴². Bennett, por ejemplo, cuestiona que la inclusión de la máquina traga-monedas dentro de los sistemas intencionales, ya que no cumpliría con ciertas condiciones mínimas para contar como tal, empezando con que no cumpliría con los propios dictados evolutivos que Dennett mismo propone. En este respecto la observación de Bennett es atinada. Ciertamente la máquina puede compartir con con ciertas propiedades evolutivas la característica de identificar su contenido relacionalmente con el entorno. En el caso de la máquina nos remitiríamos al uso al cual lo someten sus dueños. Sin embargo, esta dependencia hace que lo que es un buen funcionamiento en un momento podría ser un mal funcionamiento en otro momento, según la variación de su uso. En este caso es difícil hablar de "errores" en la máquina. En el caso de un rasgo evolutivo, en cambio, esta fluctuación en valores no se da.⁴³ Aunque Dennett quiere escapar de las críticas diciendo que su estrategia en el caso de la máquina traga-monedas es llegar a la idea de que cuando se dice que la máquina "rechaza" una moneda falsa o que "cometió un error" o que su mecanismo no es intencional en realidad no se está discutiendo ninguna cuestión de fondo: no importa mucho cuál

⁴¹ Stich (1981).

⁴² Bennett (1993).

⁴³ Cfr. Bennett (1993).

descripción se elige y que se diga si es intencional o no. Su propósito es el de sugerir que lo mismo pasa en el caso de otros sistemas intencionales cuyo estatuto como sistemas intencionales es menos dudoso. Dennett quiere cercenar la postulación de la existencia de ciertos *hechos* mentales en todos los casos de intencionalidad.⁴⁴ Sin caer en los abusos de hacer postulaciones ontológicas robustas, que Dennett quiere evitar, de cualquier forma tendríamos que reforzar la estructura de lo intencional. Eso es lo que Bennett intenta hacer en su propia propuesta y es lo que queremos hacer aquí en la siguiente sección, aunque los reforzamientos que se sugieren toman diferentes cauces.

Lo que legitima que se adopte frente a un sistema una estrategia intencional, y por ende, que le atribuyamos estados mentales, depende de que a éstas se les asigne un papel en la regulación de su conducta. En palabras de Dennett:

No es que atribuyamos (o debamos atribuir) creencias y deseos sólo a las cosas en las cuales encontramos representaciones internas; más bien, cuando descubrimos algún objeto para el cual la estrategia intencional funciona, nos esforzamos por interpretar algunos de sus estados internos o procesos como representaciones internas. Lo que hace que algún rasgo interno de una cosa sea una representación sólo puede ser su papel en la regulación de la conducta de un sistema intencional.⁴⁵

Ciertamente tendríamos en la propuesta de Dennett los estados internos del sistema, que no son estados o procesos psicológicos sino neurofisiológicos, pero que interpretamos psicológicamente cuando queremos ver la conducta, no sólo

⁴⁴ Cfr. Dennett (1993), p. 390.

⁴⁵ Dennett (1987a), p. 32.

como parte de un mecanismo, sino como un producto regulado.

Dennett sostiene que hay una liga muy estrecha entre la organización interna de un sistema y su entorno, de manera que decimos que un sistema tiene una "representación" del entorno cuando lo refleja o cuando está implícito en su organización.⁴⁶ Si el vínculo de un sistema con su entorno es muy estrecho, ese hecho se refleja en su estructura interna y, sus respuestas, consiguientemente, serán muy limitadas. Pero en la medida en que se incrementa o enriquece la complejidad del vínculo, más versátiles serán las respuestas del sistema y más complejas nuestras atribuciones. De allí que las referencias al entorno tengan que aparecer en las atribuciones de contenido o las interpretaciones de los estados del sistema. Así, sostiene que en algún sentido muy pobre, un termostato por ejemplo, tiene una representación del mundo, empero muy restringida y, al aplicarle el punto de vista intencional dotándolo de creencias y de deseos (en el sentido que les da en su teoría), éstos van a ser muy limitados. Por ejemplo, podemos decir que el termostato de un calentador se apaga cuando "cree" que la temperatura ambiente es la adecuada, se vuelve a prender cuando "cree" que la temperatura descendió, etc. Un termostato por ejemplo no tiene una estructura interna muy compleja, y es claro que carece de conceptos, por lo que términos tales como "temperatura", "calor", "calentador", etc., dice Dennett, se deben "des-interpretar", o despojar de contenido semántico. Dennett no deja de reconocer que la representación de un termostato es tan restringida que se puede tomar como un caso degenerado de atribuciones intencionales. Su aplicación en términos predictivos no agrega mucho en comparación con la predicción que se obtiene en términos de su diseño. Se podría enriquecer el comportamiento del termostato, permitiendo un número mayor

⁴⁶ Cfr. Dennett *ibíd.*, p. 31.

de atribuciones sí, por ejemplo, le agregáramos maneras alternativas de capturar la temperatura, pongamos por caso que "oyera" cuando alguien dice cuál es la temperatura ambiente, o que tuviera acceso a más puntos geográficos. En un sentido, el ejemplo del termostato es más afortunado que el de la máquina de refrescos.

Dennett explica que su énfasis en la semejanza de los humanos con los objetos como el termostato, aunque como objeto intencional el último sea muy primitivo, es para destacar que no hay un momento "mágico" en el tránsito de objetos menos complejos a más complejos en el cual los objetos realmente empiecen a tener una representación interna del mundo. Con la ayuda del lenguaje podemos expresar creencias y deseos más precisos y sobre la base de la afirmación de sus poseedores podemos hacer atribuciones más específicas. Pero para Dennett es erróneo tomar el modelo de las creaturas que poseen un lenguaje como el paradigma de lo que es tener creencias y deseos.⁴⁷ Si la representación no tiene ningún efecto o no interviene en el curso que toma la conducta o, como dice Bogden, si la representación es inerte, no sería correcto llamarla una representación. Se requiere que los registros de alguna realidad desempeñen algún papel en la conformación de una respuesta ulterior.

Dennett "naturaliza" al hombre y en ese sentido no lo diferencia de otros seres vivos y artefactos o máquinas muy sofisticadas, respecto a la intencionalidad: a todos los ve por igual como "artefactos". Dennett comenta respecto al contraste que hacen los filósofos como Searle y otros, entre humanos y artefactos, según el cual atribuyen a los humanos una intencionalidad original o genuina mientras que a los artefactos sólo les conceden una intencionalidad derivada diciendo que, en su

⁴⁷ Cfr. Dennett *ibid.* p. 21.

teoría, también los humanos tendrían intencionalidad derivada. Supuestamente, un ser con intencionalidad original sería el que "decide, diseña, controla y dirige sus actos", mientras que en la intencionalidad derivada, sus actos intencionales estarían supeditados a la intencionalidad de sus diseñadores. Pero Dennett considera que, en ese respecto, las acciones de los seres humanos están igualmente sometidos a los movimientos de nuestros genes -los genes "egoístas"- y que serían éstos los únicos que estrictamente hablando tendrían una intencionalidad original.⁴⁸

Según Dennett, para que el calificativo a los humanos de "artefectos" no se queda en lo meramente metafórico, la idea de que no hay "hechos" o "significados" en el caso de los artefactos tendría que hacerse extensiva al caso de los humanos. Si, como sostienen Fodor, Searle y otros, la razón es que los artefactos no tiene significados o estados mentales porque no tienen intencionalidad original, entonces ceteris paribus tampoco tendríamos que suponer que la tienen los seres humanos. Si le concedemos a Dennett la naturalización que hace de la conducta humana, tendríamos que concederle también que la aparición de ocurrencias o pensamientos mentales específicos son productos que derivan de la complejidad de la estructura interna de los sistemas humanos. En el capítulo siguiente me detendré en cómo Dennett vincula los diferentes niveles para los cuales introduce diferentes teorías, que denomina "psicología cognitiva sub-personal" y "teoría psicológica personal". Podríamos sugerir que puede dársele sentido a la propuesta de Dennett de que es posible dar existencia objetiva a ciertos patrones de conducta que sólo son accesibles desde cierta perspectiva, considerando los patrones de conducta a la manera wittgensteiniana de "ver como". Como se verá, esta propuesta podría defenderse sin comprometernos con el instrumentalismo. Sólo tendríamos que

⁴⁸ Vide Dennett (1987c) p. 300.

aceptar la idea de Dennett de que existen diferentes niveles de explicación.

En síntesis, lo importante para Dennett, y el aspecto que lo distingue de otros intencionalistas, es que no circunscribe la terminología intencional a los sistemas más evolucionados y complejos como el de los humanos, es decir, no asocia la intencionalidad no hace su primera aparición con el desarrollo de las capacidades lingüísticas y de reflexión. Para Dennett sería equivocado suponer que la mayoría de nuestras atribuciones de estados mentales se hace con base en las afirmaciones verbales de los sujetos a los que se les adjudican. Dennett juzga que la limitación de las caracterizaciones intencionales a este nivel ha malencaminado la discusión de lo mental al campo de la semántica. En la semántica de actitudes proposicionales, por lo general, se supone que se trata de una relación uno de cuyos términos es el contenido de lo que se cree, desea, etc., el cual tiene un significado definido o determinado. Así, Dennett, considera que el problema se ha desplazado a un terreno inadecuado y propone enderezarlo considerando que pueden recogerse con las categorías intencionales una gran diversidad de patrones de conducta. En este respecto, Davidson adopta una posición excepcional ya que ni acepta la semántica de actitudes proposicionales ni tampoco acepta que a los seres que no son capaces de tener un lenguaje pueda atribuírseles con propiedad una intencionalidad (o que se explique su conducta aludiendo a creencias y a deseos).

Al inicio de la exposición de la postura de Dennett señalé que en tanto marco general para las explicaciones no resulta problemático aceptarla. Sin embargo, en la medida de lo posible, lo separé de la ontología que Dennett asocia a cada nivel explicativo y de la manera en que inserta en este marco general la distinción entre los niveles de teoría sub-personal y personal. En otro trabajo⁴⁹ me referí a la tesis

⁴⁹ Saab (1994).

ontológica de Dennett, en donde rescato su "realismo" en la forma de la idea wittgensteiniana de "ver como" -que en realidad sería un realismo muy débil. En el capítulo siguiente retomaré la idea de Dennett en cuanto a los niveles de teoría sub-personal y personal y su conexión con la distinción entre las diferentes estrategias explicativas, por lo que me reservaré las conclusiones para entonces. Aludiré en particular a la dificultad que le señala por ejemplo McDowell que surge de la manera en que Dennett yergue la teoría personal sobre la base de la teoría sub-personal.

3.4 Modelo de explicación biológico-funcional

En una sección anterior, me referí al modelo de explicación intencional que suscribe Davidson, en el cual la conformidad con los principios de racionalidad - con su correspondiente normatividad- funge como su rasgo distintivo. Me parece que el modelo biológico-funcional que expondré en esta sección constituye un mejor modelo. En el modelo funcional se aceptan los principios de racionalidad pero además se inscribe el modelo de lo mental en un modelo más amplio de explicación. Este modelo de explicación ha cobrado auge en la última década y agrupa a un número de posiciones que intenta inscribir a las explicaciones psicológicas dentro de las explicaciones biológico-funcionales, introduciendo una serie de principios que hace referencia a la teoría de la evolución. Las exploraciones de Millikan en este sentido son de singular importancia. Millikan se sitúa en la línea de Dennett, y contribuye en su desarrollo defendiendo y elaborando los detalles de la propuesta que en Dennett únicamente se delinearán de manera esquemática. Ambos autores comparten la idea de que hay explicaciones en las que el legado evolutivo es central para dar cuenta de la conducta intencional. Con ello no quiero sugerir que no haya diferencias importantes entre las posturas de Dennett

y Millikan, siendo el tipo de realismo que suscriben una de esas diferencias más notables.

En el modelo de explicación biológica-funcional de estos autores se hace hincapié en que el funcionamiento de ciertos rasgos de un organismo o sistema en particular se hereda del funcionamiento de esa *especie* de rasgo en ese *tipo* de organismo o sistema. Es decir, la estructura y disposiciones actuales del organismo y sus comportamientos en su *habitat* particular, se explican a la luz de la historia natural que reconstruye cómo estas estructuras y disposiciones han evolucionado y se encuentran consolidadas en su organismo o sistema. No se trata de dar una historia descriptiva o genética, sino de proporcionar los elementos funcionales constitutivos de ciertos rasgos que identifican y explican no sólo su proceder actual, sino cómo *debe proceder* dado el tipo de rasgo de que se trata. La estrategia es la de llegar a los estados mentales a través de la noción de función biológica, que en algunos autores se entiende etiológicamente o de manera causal (Wright, Millikan, Sober, Wimsatt), mientras que en otros se entiende de manera causal-normativa (McGinn, Papineau). Aquí buscaré defender una forma del modelo evolutivo del segundo tipo.

Las funciones comparten con las disposiciones la peculiaridad de identificarse por sus efectos o consecuencias. De todos los efectos que se asocian a un cierto rasgo, se eligen los que se considera que forman parte de su función, de su razón para existir. La función biológica de cierto rasgo o dispositivo agrega al mecanismo causal físico subyacente información acerca de aquello para lo que ha servido y sirve todavía ese rasgo. Así, los moldes que pertenecen a las funciones o diseños de los sistemas corresponden a categorías distintas a las físicas; sus taxonomías no son comparables. Una diferencia de taxonomías de este tipo está implícita en las diferentes perspectivas de explicación que propone Dennett. Así, por ejemplo, decir

que un órgano cae bajo la categoría de "corazón", es decir que es producto de mecanismos que han proliferado a lo largo de la historia evolutiva de ese órgano, hecho que a su vez se explica por su capacidad para bombear sangre en los sistemas que lo contienen, ayudando así a su perpetuación. Sober define el concepto de adaptación de la siguiente manera:

Una característica *c* es una adaptación para desempeñar la tarea *t* en una población si y sólo si los miembros de la población hoy en día tienen *c* porque, ancestralmente hubo una selección para tener *c* y *c* confirió un beneficio de adecuación ("*fitness*") debido a que desempeñó la tarea *t*.⁵⁰

En la terminología de algunos filósofos de la biología, los dispositivos seleccionados son "seleccionados *para*" ("*selected for*"), se distinguen de la selección *de* ("*selection of*") rasgos, en cuanto a que los últimos son parásitos o productos secundarios que la evolución ha traído consigo. Un ejemplo de estos últimos serían los latidos del corazón. No se considera que el latido del corazón sea una característica adaptativa ya que el corazón no evolucionó -hasta donde se sabe- y para producir latidos.

La manera en que las funciones a las que se aluden en las explicaciones se identifican y definen no es con base en leyes -sean deterministas o probabilistas. Si es que hubiera leyes, entrarían en el nivel de los mecanismo mecanismos físicos y no en el de sus funciones. Las narrativas histórico-biológicas son contingentes. Se narra una historia real de las funciones y no un proceso de selección que se aplicaría si ciertos contrafácticos fueran verdaderos.⁵¹

⁵⁰ Sober (1993), p. 83.

⁵¹ Fodor (1990), p. 76.

El que ciertos rasgos tengan funciones biológicas propias, es lo que los diferencia también de la noción de función que emplean los funcionalistas computacionales como Fodor. La noción clave de función de estos últimos se vincula únicamente a la estructura interna del sistema y donde se borra toda diferencia entre tipos de rasgos con base en su historia evolutiva. En algunos de los funcionalistas computacionales además de la estructura interna del sistema, cuando mucho, se hace la concesión de que el entorno contribuye causalmente. ¿Qué ventaja tiene hablar de una función biológica evolutiva frente a las funciones actuales del organismo que toma en cuenta al medio ambiente? ¿Qué trascendencia tiene para lo mental el que históricamente haya funciones que han demostrado tener un poder adaptativo exitoso en el medio ambiente en el que florecieron?

Si se catalogan ciertos estados mentales como un tipo de funciones que se caracterizan por hacer referencia a sus funciones primigenias, se sostiene que mediante ciertos rasgos o propiedades mentales que los estados mentales se han seleccionados para coadyuvar a la supervivencia y reproducción del organismo. Los estados mentales se incluirían entre los casos que tienen funciones propias y que han logrado sobrevivir y reproducirse en sus descendientes. La viabilidad de la propuesta biológico-funcional para las explicaciones psicológicas descansa en hacer plausible la integración de los estados mentales a las funciones, de modo que al estar frente a ciertos comportamientos se legitime la atribución de esos estados mentales. Se propone que estados mentales como las percepciones, las creencias, los deseos y las intenciones se vislumbren como estados que desempeñan dos funciones básicas: la primera, la de representarse el entorno y la segunda, la de cooperar con otros estados mentales para producir acciones. En ambas funciones de esos estados mentales se introducen ciertos contenidos no semánticos. Con la primera función, se dotan a los estados mentales de los contenidos que se extraen

de la función de representación. En la idea misma de representación figura ya la estructura básica de la intencionalidad: la de apuntar hacia algo, i.e. su direccionalidad. El valor de una representación depende de aquello que representa. En este tipo de teoría la forma de representación lingüística se entiende como un producto adaptativo adicional que capacita a quienes la han desarrollado, a incrementar su interacción con el medio ambiente. Se pretende explicar el tránsito a esas formas más desarrolladas de pensamiento a través de la idea de contenido - que en este caso sí sería un contenido semántico-, que se puede denominar "contenido representacional".

La segunda función de los estados mentales consiste en que al colaborar con otros estados mentales de manera que conduzca a la acción. Cada una de las funciones de ciertos estados mentales, representacional e inferencial, portaría su propia normatividad. Esta doble fuente normativa se exhibe sólo en el comportamiento de ciertos organismos como nosotros que hemos desarrollado capacidades lingüísticas. De las dos funciones que definen a ciertos estados mentales, ya que no todos los estados mentales tienen una función representacional, la menos discutible es la forma específica en que los diferentes estados mentales forman un patrón inferencial. Esta articulación entre estados mentales deriva de la función propia de su mecanismo físico subyacente. La creencia, por ejemplo, tiene como una de sus funciones propias la de participar en inferencias de tal modo que ayude a producir la satisfacción de deseos.⁵²

La otra función de ciertos estados mentales, su función representativa, que se establece en virtud de sus contenidos puede ser más controvertible. En las páginas siguientes me concentraré en entender esta fuente de normatividad que

⁵² Millikan (1986), p. 67.

compartimos con los organismos menos complejos que nosotros. La teoría tiene que resolver dos cuestiones básicas: que el contenido sea constitutivo de la función, i.e. que forme parte de la identificación de la función y no sólo de nuestra manera de interpretarla. Este tipo de dificultad la señalan autores como McDowell.⁵³ La otra dificultad estriba en hacer válido que cuando el sistema está en contacto con el entorno se obtiene la relación de representación. El primer paso consiste en aceptar que el contenido es una propiedad que se constituye relacionamente. Si se contara con una fuente de normatividad -implícita en la noción de función evolutiva-, entre sus ventajas estaría la de poder integrar los elementos normativos y los elementos causales que las explicaciones causales intencionales, que como vimos en el capítulo anterior, posturas como la de Davidson, no acertaban en integrar.

Muchos autores, antes de la inserción de la discusión de lo intencional en las redes de la teoría de la evolución, han destacado una u otra de las funciones que aquí de manera bifronte se asocian a lo mental. Por ejemplo Bennett y Burge suscriben un tipo de postura externalista en cuanto a la fijación del contenido de un estado mental. Al menos la postura que defiende Burge es un tipo de externalismo que se denomina *fuerte*.⁵⁴ El externalismo fuerte se caracterizará por el hecho de que el estado mental necesita la existencia en el entorno del sujeto de alguna instancia física de la cual depende la identificación del estado mental. Así, de la misma manera en que Putnam sostiene que los "significados" no están en el cerebro del sujeto, i.e. que la identificación de los significados depende del entorno, recurriendo a su afamado ejemplo de un mundo gemelo, se pretendería sostener que los estados mentales dependerían del entorno. Si hubiese dos estados mentales

⁵³ McDowell (1994).

⁵⁴ Cfr. McGinn (1989), p. 7.

indistinguibles internamente, ubicados en un *habitat* en donde hubiera una diferencia en cierta propiedad, entonces la referencia a esa propiedad también variaría. Si en nuestro mundo el líquido que denominamos "agua" tiene una composición química de H_2O , significaría para S -que habita en nuestro mundo- algo distinto de lo que significa en el mundo gemelo en el cual su composición química fuera distinta, por ejemplo XYZ. En estas posturas anticartesianas se defiende la tesis de que los contenidos de los estados mentales tienen un carácter relacional con respecto a aquello a lo que se refieren; la individuación de los estados mentales específicos requiere que se tomen sus contenidos como propiedades relacionales con el medio ambiente. Los ejemplos paradigmáticos que nos sirven para expresar contenidos son los deícticos y los conceptos de clases naturales (aquellos que se caracterizan por tener un estereotipo y una extensión). Habría otra forma de externalismo que se podría denominar *débil*, en la cual en el entorno inmediato del sujeto podría no encontrarse la propiedad o entidad a la que el estado mental hace referencia.

Defenderé, con McGinn, las siguientes tesis en cuanto a los diferentes estados mentales. No todos los estados mentales se identifican de la misma manera y, muy en particular, que no todos se identifican a través de contenidos. Supongo también que los estados que se identifican por su contenido, son extrínsecamente relacionales, en sentido débil. Los estados mentales como las creencias, los deseos y las percepciones serían de este tipo. Los contenidos de los estados mentales que se refieren a clases naturales⁵⁵ y los deícticos se asociarían con un externalismo

⁵⁵ De hecho, McGinn matiza la tesis con respecto a las clases naturales acotando que, en rigor, las clases naturales complejas, como por ejemplo " H_2O " y "carbón" no satisfacen la tesis del externalismo fuerte, ya que alguien podrían entender los conceptos sólo a través de los términos que lo constituyen. El caso de " H_2O " podría entenderse en términos de sus componentes moleculares y "carbón" en términos de "madera comprimida de cierta manera a lo largo de millones de años" (Este último a su vez contiene un término de clase natural, a saber, "madera"). En

fuerte. Finalmente, habría estados mentales que no se identificarían de manera externa como por ejemplo, las imágenes, las sensaciones y las emociones. Estos últimos estados mentales, en sentido estricto, ni siquiera tendrían una función representacional.

Uno de los retos de la teoría biológico-funcional es la de poder mostrar que el contenido fijada por la función biológica, que está emparentado con el contenido que arroja la función representacional, sea también adecuado para constituir el contenido del estado mental. Uno de los rostros del problema consiste en poder dar cuenta de los estados mentales que se identifican mediante contenidos muy específicos y que el filtro biológico-evolutivo no permita, o sirva para, decantarlos. Si bien en la relación de representación, R representa a la propiedad p y hay una relación de correspondencia entre sus *relata*, en el caso de las funciones biológicas sus contenidos podrían especificarse con diferentes propiedades coextensionales.

Empecemos con la noción de función en la cual cabrán los estados mentales. Millikan considera que las funciones biológicas en tanto funciones relacionales son una subclase de funciones que denomina propias, i.e. de cómo las desempeñan. Escribe Millikan:

...Cada órgano corporal o sistema cae dentro de las categorías biológicas o fisiológicas en las que cae debido a sus conexiones históricas con casos anteriores del tipo que han desempeñado ciertas funciones o, de manera típica, conjuntos de funciones. Así, sea que [el órgano o sistema] en sí mismo sea capaz o incapaz de desempeñar cualquiera de esas funciones, a cada órgano o sistema se asocia un conjunto de funciones que le son biológicamente "propias" -funciones que han ayudado a explicar la sobrevivencia y proliferación de sus

estos casos, el estereotipo no "revela la naturaleza esencial de la extensión". Cfr. *ibid.* p.35.

ancestros.⁵⁶

Como ya dijimos, la función propia va a estar dada por su historia y no por sus poderes o disposiciones internas. No parecería poder fijarse de manera *a priori* si una función es propia o no, sino que hay que observarlo diacrónicamente. Y, una función propia *relacional* en virtud de la direccionalidad que la caracteriza, engancha con los elementos ambientales. La noción clave de la intencionalidad, su rasgo de ser "acerca de", en esta teoría se derivaría -en algún sentido- de la noción de "función propia".(McGinn) Uno puede admitir, sin gran dificultad, que la introducción de la dimensión histórica en las explicaciones puede ayudarnos a entender más profundamente la presencia de ciertas conductas. Sin embargo, lo que aquí se quiere es algo más fuerte: que la historia nos dé una perspectiva de la constitución misma de esas funciones y su capacidad reproductora en la especie, que establezca la pauta de su funcionamiento Normal o su buen funcionamiento (i.e. una pauta normativa). Este es uno de los sostenes claves para la postura de Millikan, aunque ella califica a ese funcionamiento como "cuasi-normativo".⁵⁷ Nos dice cómo cierto dispositivo tiene que funcionar, al margen de que en los hechos en ciertos organismos el dispositivo tenga un comportamiento anormal. Lo que parece impedir que la capacidad reproductora se vea como la perpetuación de una función "anómala" en lugar de una función normal es que la anomalía sólo puede medirse comparándola con lo que se ha establecido como su función correcta: se tienen que especificar las circunstancias en las cuales la descripción de la función

⁵⁶ Millikan (1986), pp. 51-2.

⁵⁷ Millikan (1984), p. 5.

se fija.⁵⁸

En el caso de los estados mentales que incorporan contenidos, éstos tienen la función de representar ese aspecto en el mundo. Los estados mentales individuán sus propiedades externamente, en sentido débil, ya que pudiera darse el caso que el contenido representado no se encontrara en el entorno del sujeto. En este caso, el estado mental contendría una representación falsa. La función biológica establece el parámetro de lo que será el funcionamiento Normal del estado mental. Lo Normal desempeña un papel prescriptivo. Si una de las funciones de los estados mentales es la de representarse la propiedad que se encuentra en el entorno, al no encontrarse, su función sería anómala. Millikan desglosa esta idea de representación, no en cómo se produce la representación, sino en el uso que se hace de ella: "Si en verdad la función de una representación interna es la de indicar a lo que representa, claramente no es un signo natural, un signo que bastara con mirarlo para interpretarlo". Debe funcionar -continúa Millikan- como signo o representación para el sistema mismo.⁵⁹ Como diría Wittgenstein: "Todo signo parece por sí solo muerto. ¿Qué es lo que le da vida? -Vive en el uso."⁶⁰

En su crítica a un tipo de posición que habla de representaciones, McDowell considera que se confunden dos procesos: el proceso mediante el cual se desempeña la función de "captación" de información y el proceso que tiene la función de representar.⁶¹ La captación de información, continúa McDowell, es el modo en el que el dispositivo está abierto al medio ambiente o está en contacto con él. El

⁵⁸ Cfr. Millikan, (1986) pp. 284-5.

⁵⁹ Millikan (1989), pp. 283-4.

⁶⁰ Wittgenstein (1958) 4432.

⁶¹ Cfr. McDowell (1994).

medio ambiente es quien le proporciona la información. Una de las tesis centrales de McDowell es la de negar que en ese proceso se estén captando contenidos, cuando mucho el dispositivo reacciona a la presencia de cierto objeto que se encuentra en el entorno.⁶² McDowell piensa que los contenidos sólo son lingüísticos. En el segundo proceso McDowell introduce al receptor de la información y argumenta que sólo entonces aparecen los contenidos. Me parece que McDowell no muestra que en la captación de información se introducen contenidos. Su rechazo básico es: que los contenidos no son elementos "dados" en la experiencia, sino que se construyen. Sin embargo, se podría defender que portan contenidos de la misma manera que los portan las funciones, pensando que los contenidos no necesariamente se forman a través de la experiencia directa.

Las funciones y los estados mentales son relacionales con respecto al entorno de manera semejante. McGinn intenta dar cuenta de esa semejanza proponiendo que los estados mentales heredan esa propiedad relacional de las funciones biológicas. Así como las contenidos constituyen las funciones, igualmente los contenidos forman parte de la constitución de los estados mentales, a saber, por referencia a entidades, propiedades, etc. objetivas. El vínculo no sería meramente causal sino constitutivo. McGinn lo expresa diciendo que hay una relación sistemática entre la función y el contenido, a saber, que aquello que individúa el contenido del estado mental constituye también su función biológica: la doble aparición de aquello a lo que se refieren tanto el contenido como la función "debe reflejar una conexión sistemática conceptual o teórica -la función fija y es fijada por el contenido."⁶³ El puente entre el aspecto relacional de ciertos rasgos y el aspecto relacional de ciertas

⁶² *Ibid.* p. 196.

⁶³ McGinn (1989), p. 147.

funciones pretende tenderse por la interdependencia referencial de ambas. McGinn pone el ejemplo del deseo de beber agua. Este deseo es causado por la necesidad del organismo de agua. La función biológica del deseo es la de lograr introducir agua en el organismo o hacer que el organismo tome agua. Aunque el organismo no lo consiga, no deja de ser aquello por lo que tenemos el deseo (para lo que existe y sirve el deseo). La conexión radica en que al proporcionar el término del medio ambiente de la relación funcional propia, se menciona el mismo elemento del medio ambiente al que se refiere el deseo.⁶⁴ McDowell considera que -por ejemplo- en el caso del sistema visual de la rana, es equivocado suponer que el sistema reaccione a la presencia de un insecto -o que "insecto" forme parte de su contenido. Él piensa que sería más apropiado hablar de que la información es de "un punto minúsculo móvil" que nosotros podríamos redescubrir como información que tiene que ver con insectos, pero que no es, en sentido propio, lo que capta la rana. Pero independientemente de cómo discrimine la rana los elementos del entorno, sea en términos de insectos o de puntos minúsculos móviles, para que sirva a su vez para la satisfacción de su hambre, da lo mismo que se le asocie a insectos que a puntos móviles.

Fodor hace una objeción semejante a quienes quieren dar cuenta de lo intencional a través de la relación -ya no específicamente de la relación de representación- sino de cómo se identifica de la función de un cierto mecanismo. Su argumento es que los datos o los diferentes contenidos que serían extensionalmente equivalentes pero intensionalmente distintas subdeterminan tanto a las representaciones como a las funciones. Aplicado al caso de la rana, muchas representaciones podrían ser candidatos para el contenido de lo que la rana ve y que

⁶⁴ Ibid. p. 147.

explica su conducta en la presencia de insectos, siendo "insecto" sólo una de ellas.⁶⁵ Para Fodor esta forma teleológica no resuelve el problema de la disyunción, que es -para Fodor- el gran problema que se le presenta a cualquier teoría que intenta dar cuenta de la semántica en términos naturalistas. El problema de la disyunción, a grandes rasgos, consiste en que la teoría no pueda dar cuenta de los casos en los que se producen errores en la aplicación de un término. En una teoría semántica naturalizada se asume el principio de que si x causa y , entonces hay propiedades de y y de y en virtud de las cuales se establece el vínculo y hay una ley en la que se subsume, la cual relaciona las propiedades semánticamente. La ley es condición suficiente para que se de la instanciación de ciertas propiedades. Si "perro" significa perro, entonces se incluyen sólo aquellos objetos que entran en la extensión del término, en virtud de que existe una ley en la cual los perros causan "perro". Con este principio, se abren dos posibilidades: a) que sólo los perros causen "perro", en cuyo caso sólo los objetos que caen en la extensión de "perro" causan que se instancie. b) Que una disyunción -que un perro o un gato-en-una-noche-oscura- sea condición suficiente para causar una instancia de "perro". En este caso, por el principio antes citado, "perro" expresa la propiedad de ser un perro o un gato-en-una-noche-oscura, de modo que la expresión del término "perro" es la unión de ambos objetos. Las instancias de "perro" que son causadas por los gatos en las noches oscuras serían verdaderas, ocasionando que no hubiera lugar para que se produjeran errores o falsedades.⁶⁶ De manera similar, sostiene Fodor, las funciones interpretadas teleológicamente no resolverían el problema de la disyunción por la sencilla razón de que la función de un mecanismo tampoco puede

⁶⁵ Fodor (1990) y Hall (1990), p. 196.

⁶⁶ Cfr. Fodor (1990), p. 59.

fijar su contenido.⁶⁷ En ambos casos surge el mismo problema de la indeterminación.

Fodor objeta a Millikan, en específico, que en su posición ella incurra en el error de también atribuir funciones propias a los estados mentales específicos, cuando sería más plausible defender la tesis teleológica hablando de los mecanismos para fijar estados de ese tipo.⁶⁸ En otras, palabras, que no tendría por qué sostenerse que los estados intencionales mismos se individualizan en virtud de sus funciones. Así –continúa Fodor– Millikan no se tendría que comprometer –como de hecho lo hace– a decir que la creencia de que siete es un número primo, es una creencia *para*, comprometiéndola a que la descripción que identifica a ese estado sea también la descripción que identifica su función propia.

Sin embargo, contra Fodor, podría sostenerse que el análisis teleológico del contenido no tiene por qué dar cuenta de todos los aspectos que éste pueda adoptar. Donde se ve más claramente la conexión entre el contenido de los estados mentales y el de la función es en el caso en el que se relacionan las necesidades con las funciones más básicas del organismo. En defensa de Millikan diríamos que no tenemos que admitir que el contenido forma parte del estado mental tiene que reducirse al contenido funcional.

Otra crítica que presenta Fodor a un funcionalismo como el de Millikan estableciendo una semejanza la comparación entre el estado intencional y un estado funcional de órganos tales como el corazón, los ojos y los riñones, es que en el último caso hubiera una ley que conectara el funcionamiento del corazón con el bombeo de sangre (*ceteris paribus*), al igual que hubiera una ley que conectara el

⁶⁷ Ibid. p. 70.

⁶⁸ Fodor (1990).

deseo de comer con la satisfacción de comida. Sin embargo, en cuestión de las leyes, no existiría en el caso de deseos específicos como el deseo de ganar las elecciones para el puesto de presidente municipal. Fodor lo ilustra con el ejemplo de que: "ese deseo particular puede aislarse causalmente de la ocurrencia del estado de cosas que lo satisfiría, aun cuando las cosas fueran perfectamente Normales".⁶⁹ Al respecto podemos hacer dos comentarios. El primero es que, como sostienen Millikan y McGinn, la taxonomía causal es distinta a la taxonomía de las funciones, consistiendo una de sus diferencias justamente en que las funciones no tienen que operar bajo leyes. Por otra parte, si se destaca que una de las funciones propias de los estados mentales (intencionales) es el representacional, no tendría por qué aceptarse -como hace Fodor- que los contenidos de las representaciones se identifiquen extrínsecamente en sentido fuerte y por tanto que en el caso de un contenido específico como el que se menciona arriba, bien puede establecerse la relación representacional en términos referenciales en donde el contenido de lo que el deseo expresa no fuese aislable del estado de cosas al que refiere, contrariamente a lo que sugiere Fodor. El hecho de atarse a funciones propias no quiere decir que tenga que haber condiciones suficientes Normales que hagan que ese deseo se satisfaga.

Podemos considerar que la relación de representación una vez que se constituye como relación de referencia, i.e. como función semántica, si requeriríamos de capacidades lingüísticas y de auto-reflexión. Pero esto no impide que se desarrolle la función propia de un estado mental en seres complejos como nosotros sin que dé ninguna información acerca de sus contenidos específicos, ni tampoco de los tipos específicos de inferencias en las que participa.

⁶⁹ Fodor *ibid.*, p. 65 y pp. 67.

III. CARACTERIZACION DE LAS CREENCIAS

En el capítulo anterior me ocupé de diferentes modelos de explicación, con miras a elegir el modelo más adecuado para dar cuenta de las acciones. El modelo biológico funcional me resultó el más satisfactorio. Ese modelo nos permitía, además, conservar la manera natural y cotidiana en la que empleamos términos como creer, saber, desear, esperar, temer, etc., conformando un articulado universo interno que forma parte, *prima facie*, de las razones que explican lo que hacemos o decimos. Esta estructuración interna que supone un funcionamiento *holista* ha sido una consideración importante para la elección del modelo de explicación para lo mental. Los elementos holistas, los constreñimientos normativos y la aceptación de una forma de externalismo, como sostiene Davidson, son características que en un modelo suelen venir juntas. Si bien hemos estado de acuerdo en ello, hemos visto que el modo y el sitio en que se introduce la normatividad varía mucho según el modelo que se defiende.

En esta parte del trabajo, con el modelo explicativo como telón de fondo, me ceñiré específicamente a las creencias y algunos de los problemas que se desatan en torno a ellas. Las creencias en sí mismas no sólo constituyen una parte muy importante de nuestros estados mentales, sino también forman parte del análisis de algunos otros estados. Por ejemplo, forma parte del conocer, de las expectativas y como parte de cierto tipo de deseos o de cierto tipo de emociones. Una creencia además de admitir una gama muy heterogénea de contenidos, singulares y generales, puede reiterarse generando diferentes niveles: creencias, creencias de creencias, etc.. En el capítulo anterior señalamos que en el organismo humano se estructuran niveles de creencia y es una cuestión importante decidir en qué nivel emerge la conciencia y nuestras capacidades reflexivas. La conexión entre los

niveles de creencias puede entenderse como la manera en que nuestros diferentes sub-sistemas establecen canales y redes de información entre sí hasta que finalmente el sujeto desentraña la información que le llega a la conciencia. Con nuestras creencias, como dice Quine, se conforma una red ("*a web of belief*") que tiene una coherencia interna. Pero también se podrían admitir diferentes maneras de percatarse ("*awareness*"), de manera pre-reflexiva, asociada a diferentes sub-sistemas. Con toda esta información se integra la visión de nosotros mismos y la que nos formamos del mundo.

En medio de la búsqueda de esta coherencia que tratamos de establecer entre nuestras creencias no nos libramos de caer en ciertos formas de incoherencia. En esta parte del trabajo me centraré en ciertos tipos de incoherencia que se produce entre lo que se piensa, siente, etc., y lo que se hace y en donde entra en juego algo más que una cuestión de error o contradicción. Hay casos de incoherencia que se producen por inadvertencia -sea por omisión, error o ignorancia- en donde la falla se debe a que, pese a que aceptamos premisas que entran en conflicto, no acertamos en juntarlas. Esta forma de incoherencia se resuelve al caer en la cuenta del conflicto, eliminando alguna de ellas. Pero hay otras formas de incoherencia que constituyen casos de irracionalidad que son muy enigmáticas, como son el auto-engaño y la *akrasia*, que para unos inclusive constituyen patologías. Para explicar estas formas de incoherencia, a diferencia de las anteriores, tenemos que vivir con ellas. Como veremos más adelante, la admisión de diferentes maneras de percatarse de algo o de ser consciente permite dar una explicación de cómo se produce el auto-engaño. El fenómeno del auto-engaño se concibe como una falla que se refleja en nuestras actitudes cognitivas. En la *akrasia* o debilidad de la voluntad se concibe como una falla en la acción. (Davidson, Pears) No han faltado intentos de quitarles a estos fenómenos el carácter de irracionales o, cuando menos de disolver el aire

de paradoja que las encubre. Aquí no hablaré de la *akrasia* y me concentraré en el auto-engaño. En el caso del auto-engaño parte del enigma surge del hecho de que el engaño parece ser inducido y encubierto por el propio sujeto, sin que el sujeto cobre conciencia de ello. Aludiré a las posturas de Dennett y de Mellor como casos en los que mediante la referencia a diferentes niveles de creencias, un nivel inconsciente y el otro consciente, se pretende "explicar" el fenómeno del auto-engaño. Señalaré que la forma de ambos autores de entender esta forma de irracionalidad es equivocada por dos conceptos: por trazar la diferencia en términos de niveles de creencias resolviéndola o disolviéndola más que explicándola y por dejar fuera el papel deliberado y estratégico que juega el sujeto. Si bien en muchos autores se hace acompañar a estos fenómenos de la tesis del yo-dividido, en este trabajo preferiré hablar de diferentes sub-sistemas con funciones propias. No entraré en la discusión de la identidad personal, aunque para una presentación más completa de lo mental debería hacerse.

Además del auto-engaño, me referiré a otro caso que también genera alguna forma de incoherencia, recogido en un enunciado paradójico -conocido como la paradoja de Moore- en la que se afirma "Yo creo que p , pero p es falsa". Aquí también las opiniones oscilan en cuanto a si el enunciado es contradictorio o si se trata de algún otro tipo de incongruencia. Este tipo de caso -a diferencia del auto-engaño- se generaría aunque no aceptáramos un modelo de niveles de creencias. En este caso, siguiendo a Wittgenstein, diré que no se puede dar una sola respuesta y que pueden encontrarse circunstancias y usos que van desde los que lo convierten en un enunciado contradictorio hasta los que lo hacen legítimo o no-contradictorio. Me detengo en la discusión de Wittgenstein en torno a las creencias expresadas desde la primera persona y las diferencias que su uso genera respecto de las manifestaciones de las creencias vistas desde el observador o la tercera persona.

Aplico la discusión de los usos redundantes de la creencia en su aplicación en la primera persona y en ellas sustentan algunas de las respuestas que se pueden dar a la paradoja de Moore. En esa parte del trabajo se ve más clara la incidencia de las consideraciones contextuales y pragmáticas para el entendimiento y explicación de la conducta.

Pero antes, ¿qué podemos decir de la manufactura de las creencias? ¿Se pueden construir como entidades o propiedades reales de algún tipo? ¿En qué sentido se puede decir que causan -o forman parte de las causas- de las acciones? ¿Realmente la postulación de los diversos niveles de creencias -asociados a diferentes sub-sistemas- son una buena manera de acomodar los datos a partir de los cuales se construye la explicación? ¿Cómo se vinculan los niveles entre sí y con la conducta no verbal y verbal? ¿Dónde se inserta el asentimiento? Estas son algunas de las cuestiones que tocaré en los capítulos restantes. En el capítulo anterior, incluimos a las creencias dentro de los estados mentales que tienen contenidos, los cuales pueden representar elementos o rasgos objetivos. Una tesis que mantendré es que si bien los estados mentales y las acciones no se vinculan disposicionalmente, ese vínculo puede darse entre los niveles de creencias. También sostendré que la conciencia de nuestras creencias, si bien es un aspecto importante de algunas de las creencias de segundo orden, *pace* Mellor- no es su signo central.

Una característica de las creencias de primer orden (en adelante, *Bp*), extendida de la manera que se sugiere en el trabajo, nos permite recuperarla desde su función biológica. Esta tesis nos lleva a preguntarnos si esos organismos que no han desarrollado capacidades lingüísticas y una capacidad de conciencia y de auto-conciencia pueden formarse creencias de segundo orden, (en adelante, *BBp*). Sin embargo, en estos organismos biológicos simples se da una cierta forma de conciencia o de percatarse ("*awareness*") del entorno. Hay autores que ni siquiera

estarian dispuestos a conceder que los estados que caracterizan a esos seres menos desarrollados que nosotros puedan considerarse como estados de creencia, como es el caso de autores como Davidson, McDowell y Evans. En el caso de McDowell¹, como ya vimos, se establece la distinción entre la captación de información en tanto que existen canales por los cuales pasa la información y el procesamiento propiamente dicho de la información, del cual resultan las creencias. No me parece, sin embargo, que esta diferencia de funciones pueda separarse tan nítidamente como McDowell pretende.

Si partimos de la caracterización de las creencias, entendidas como estados psicológicos tal como se presentan en seres que han desarrollado capacidades lingüísticas y después la comparamos con los estados de otros seres u organismos, podemos aceptar la siguiente caracterización de las creencias que nos ofrece Williams²:

- 1) Apuntan a la verdad
 - a) Se puede predicar de ellas verdad o falsedad.
 - b) "Creer que p " es "Creer que p es verdadera".
 - c) "Creo que p " en general implica la afirmación de que p es verdadera.
- 2) La expresión natural de la creencia es la aserción.
- 3) La aserción puede ser insincera. Consecuentemente, la aserción no sería ni condición suficiente ni condición necesaria de la creencia.
- 4) Las creencias se pueden apoyar en evidencia, sea via otras creencias, o con base en el mundo circundante.
- 5) La noción de creencia es una noción explicativa.

¹ Cfr. McDowell (1994).

² Cfr. Williams (1973).

Si acaso Williams concede que se hable de creencias en conexión con otro tipo de organismos o en relación con cierto tipo de máquinas, lo hace con la reserva de que tendríamos una noción muy pobre de creencia. Por ello prefiere denominarlos "estados-C". Las atribuciones de creencia a los animales y las máquinas parecerían tener únicamente la característica 5), de ser explicativas y de que nuestras atribuciones están muy restringidas en cuanto al empleo de conceptos para explicar su comportamiento. La única característica a la que nos hemos referido hasta ahora es el empleo de las creencias en el contexto de las explicaciones. Vimos en el capítulo anterior cómo con base en que son explicativas -por ejemplo, en las posturas de Dennett y la de los funcionalistas- se defiende la atribución de creencias más allá de los humanos.

Si bien las creencias en sentido pleno tienen estas características que Williams recoge, que más bien las satisfacen las creencias de segundo orden o *BBp*, hablaré de creencias de primer orden o *Bp* para referirme tanto a las creencias inconscientes como a los estados-C. Para justificar la atribución de *Bp*, además de tener una función explicativa y predictiva, los seres a quienes se les atribuye deben ser capaces de apreciar rasgos de su entorno, que les permite adaptarse a él. Lo controvertible es si ese aspecto de la conducta del organismo al que se hace la atribución de alguna manera puede llamarse una conducta intencional.

Como complemento a las características que Williams atribuye a las creencias, introduzco los niveles de creencia, y discuto la manera en que interviene el asentimiento y la deliberación. Aplico el modelo disposicional para dar cuenta de algunas de las relaciones entre las creencias, a la vez que rechazo el análisis disposicional en cuanto relación entre creencias y acciones. En el momento en que se elige hablar de disposiciones surge nuevamente la cuestión de el tipo de relación que se establece entre la disposición y sus manifestaciones al igual que la cuestión

de si las disposiciones desempeñan algún papel causal en el desenvolvimiento de sus manifestaciones. Y, como sostuve en el capítulo segundo, se puede defender que las disposiciones desempeñen un papel causal.

Cuando nos referimos a las creencias de segundo orden (BBp), aparecen las divergencias entre diferentes autores en cuanto a si se identifican con el asentimiento (Mellor) o si se distinguen (Dennett, De Sousa y Hookway). Al igual que los últimos, aquí las distinguiré. La noción misma de asentimiento es una propiedad disposicional que, por ejemplo De Sousa, introduce para conectar los niveles de creencias Bp y BBp . La atribución al otro de una creencia de segundo grado puede hacerse sin que el sujeto al que se le atribuye asienta a ella. Del mismo modo, el sujeto puede tener una creencia BBp sin asentir a ella. Es posible que el sujeto mismo no repare o cobre conciencia de las creencias que detenta además de que puede cometer errores al estimarlas, lo cual pone límites a que el sujeto invariablemente se tome como una fuente confiable para decidir lo que cree. En autores como Dennett el asentimiento se conecta más bien con creencias de tercer orden ($BBBp$). También está la postura de Wittgenstein que identificará el asentimiento con Bp en algunos de sus usos en la primera persona. Me parece que el asentimiento puede construirse como condición necesaria y suficiente para Bp al igual que como condición necesaria para producir el estado de creencia de segundo orden.

Así como la tradición que nos viene de la modernidad nos legó un rico estudio de la creencia desde el punto de vista personal y destacando su aspecto de acto psicológico, en la actualidad hay una tendencia entre los filósofos de la mente de orientar la problemática de lo mental hacia el punto de vista impersonal o externo. La conexión entre las creencias y las acciones hace más difícil desvincular los actos mentales del entorno. Pero tampoco se trata de ver a las creencias sólo en

ese aspecto. Más bien se trata de articular las estrategias de la primera y la tercera persona. No queremos caer en la equivocación del enfoque tradicional de haberse aproximado a la problemática de la creencia sólo desde la perspectiva del conocimiento a partir de la primera persona, obsesionado por alcanzar un conocimiento cierto. La subjetividad no debe concebirse como un mundo cerrado, como un asunto que se dirime exclusivamente entre el sujeto y su interioridad de modo que su dimensión cognitiva puede satisfacerse de manera internalizada.

La integración de las voces en la primera y la tercera persona reflejan la convicción de que la vida interna no debe verse escindir de su función práctica y comunicativa. La conformación de los contenidos de nuestras creencias nos atan al mundo y, generalmente, el valor cognoscitivo de las creencias deriva de la información que encapsula y no a la inversa. El acceso que el propio sujeto pueda tener a sus estados no debe permitir que nos olvidemos que apuntan hacia la verdad y, en muchos casos, la verdad de lo que trasciende nuestra esfera interna. Las creencias tienen la función específica de dar dirección y sentido a nuestras acciones particulares. Esta concepción de las creencias la sugiere vívidamente Ramsey en su descripción de las creencias como un mapa que orienta nuestro navegar ("*A map of a neighbouring space by which we steer*")³. Los signos externos -los comportamientos- nos permiten inferir la existencia de las creencias y permiten nuestras atribuciones, sin dejar de ser inmensamente exiguos y propiciatorios de equívocos.

Se desprende que para que se dé el conocimiento subjetivo -las creencias BBp y el asentimiento- se requiere, más que tomarlo como la base sobre la cual se reconstruye el mundo. Por el contrario, la subjetividad misma requiere una

³ Ramsey (1931), p. 238.

estructuración más compleja que depende, por un lado, del mundo y, por otro lado, de la formación previa de una estructura de creencias de primer orden.

Pero tampoco es satisfactorio sólo tomar en cuenta el sesgo cognitivo-pragmático de las creencias, que en sus extremos llevó al conductismo o fenomenalismo. La creencia *para* es sólo uno de sus atributos.

4. LA CREENCIA COMO DISPOSICIÓN

Existen diversas maneras de entender los términos disposicionales. En una de ellas se entienden de manera realista, en el mismo sentido en que se entienden los términos teóricos que se postulan en las teorías científicas. Hay, a la vez, varias maneras de introducir los términos teóricos. Podrían introducirse como *abstracta*, según la distinción de Reichenbach, o de manera mixta, -a veces como *illata* y a veces como *abstracta* -, como sugiere Dennett.⁴ La manera como entenderé la disposición en conexión con las creencias es de manera realista del tipo de los *illata*, y también funcionarían como causa de sus manifestaciones. Las creencias *Ep* se construyen como disposiciones a tener creencias ocurrentes. Pero veamos antes, de qué otras maneras se ha pensado en las disposiciones, tanto en relación con las propiedades físicas como con las creencias.

Al igual que las funciones, como ya dijimos antes, las disposiciones se caracterizan por identificarse a través de sus efectos. Podemos mencionar como disposiciones físicas paradigmáticas la "solubilidad", "fragilidad", "elasticidad", "dureza". Las dos maneras más usuales de dilucidar las disposiciones son la reductivista y la realista. Si las disposiciones se toman como disposiciones a actuar,

⁴ Reichenbach (1961) considera que los científicos básicamente postulan dos tipos de entidades teóricas, ambos tipos de entidades son producto de una inferencia, esto es, no son directamente observables. Los *abstracta* son equivalentes a la existencia de concreta que son observables y su estatuto depende de una convención. Cfr. pp. 211-12.

en la posición reductivista las disposiciones se identifican con sus efectos conductuales. Las disposiciones en el sentido realista (Armstrong) se toman como propiedades que se encuentran en el objeto de manera latente. No tienen, sin embargo, una existencia independiente sino que dependen ontológicamente de las propiedades básicas que son de carácter categórico. El hecho de que tengan una existencia dependiente de otras propiedades, no impide que se introduzcan como términos teóricos, aunque de nivel muy primitivo. Otros autores, como Mellor, adoptan frente a las disposiciones una postura denominada racionalista que difiere de la postura realista en que se les da a las disposiciones un estatuto propio irreductible, sin referirlas a propiedades categóricas.

4.1 Fenomenalismo

Los fenomenalistas o los conductistas, como defensores de alguna forma de reductivismo, piensan que las disposiciones son ontológicamente prescindibles y eliminables de una teoría científica, toda vez que se descubran las propiedades categóricas de los objetos. Los enunciados que incorporan términos disposicionales se analizan en términos de enunciados condicionales y éstos no se referirían a ninguna situación existente: tendrían una existencia posible más no real. Dos connotados defensores de esta tesis son Ryle y Quine. Dice Ryle:

Los enunciados disposicionales no son ni reportes de estados de cosas observados u observables ni tampoco son reportes de estados de cosas inobservados o inobservables. No narran ningún incidente.⁵

⁵ En Ryle (1949) las disposiciones nos permiten hacer una inferencia ("Inference-ticker"). Decir que "a es soluble licencia la inferencia de "a se disuelve" a partir de "a se pone en agua". Cfr. p. 120.

De acuerdo a un fenomenalista, la atribución de una propiedad disposicional, nos indicaría que, si se dieran ciertas condiciones, el objeto tendría la propensión o la tendencia a entrar en un estado particular o a sufrir un cambio particular. Las disposiciones no engendrarían en el objeto más estados de los que ya tiene: son estados *posibles* en el que podría encontrarse un objeto y podría ser que ese cambio nunca ocurriera.

Quine traza un paralelismo entre las actitudes proposicionales y las disposiciones, que lo lleva a sospechar de ambas y, en última instancia, a eliminarlas. El primer paralelismo consiste en que las actitudes proposicionales y las disposiciones funcionan de manera muy similar respecto de los condicionales contrafácticos. Escribe Quine:

The subjunctive conditional depends, like indirect quotation and more so, on a dramatic projection: we feign belief in the antecedent and see how convincing we then find the consequent. What traits of the real world we suppose preserved in the feigned world of the contrary-to-fact antecedent can be guessed only from a sympathetic sense of the fabulist's likely purpose in spinning his fable.⁶

Otra semejanza, según Quine, conectada con la anterior, es que ni en las disposiciones ni en las actitudes proposicionales hay algún sustituto satisfactorio en términos realistas. Una disposición que se analiza en términos de condicionales contrafácticos, alega Quine, no tiene condiciones de verdad. Por ejemplo, que un terrón de azúcar sea *soluble* en agua en el tiempo *t*, quiere decir que si el terrón de azúcar se pusiese en agua en *t*, el terrón de azúcar se disolvería. No se puede afirmar ni la verdad ni la falsedad del enunciado. La única ventaja que concede

⁶ Quine (1960), p. 222.

Quine a los términos disposicionales frente a los condicionales subjuntivos es que los términos disposicionales se pueden volver "respetables" y recuperarse científicamente, si se logra dar con las propiedades físicas de los objetos. En este sentido, habría una semejanza entre las tesis de Quine y las de algunos realistas respecto a las disposiciones. Las creencias también pueden encontrar un sitio respetable similar al que se les puede encontrar a las disposiciones, a saber, haciendo una especie de reducción de lo psicológico a lo neurofisiológico.

4.2 Realismo

Un realista como Armstrong marca tres aspectos que distinguen a los estados latentes y los estados disposicionales: el primero, que un estado latente se construye de manera realista mientras que un estado disposicional, no necesariamente. Como ya mencionamos, una disposición puede construirse como un estado *posible*, como un estado en el cual un objeto podría encontrarse si se dieran ciertas condiciones específicas. Si se diera ese estado, se podría decir que el objeto sufrió un cambio o una modificación. Un estado latente, en cambio, es un estado real del objeto que está "quieto" o pasivo pero que, en ciertas condiciones, podría activarse. La segunda diferencia entre estados latentes y estados disposicionales es que a los estados latentes se les confieren poderes causales, mientras que las disposiciones podrían ser causalmente impotentes (dependiendo de si se les da existencia básica o no). Y, finalmente, una tercera diferencia sería que en el caso de las disposiciones se necesita de una condición propiciadora para desencadenar su manifestación, la cual forma parte del significado del concepto disposicional. Chomsky se refiere a esta característica de ciertos términos como "estímulo-dependientes", incluyendo en esta categoría no sólo a las disposiciones. Así, en el caso de un objeto soluble, por ejemplo, la condición de ser depositado en agua hace que el objeto se disuelva y

el ser depositado en un líquido constituye parte del significado del término "soluble". En síntesis, en la postura realista de Armstrong, la distinción entre propiedades disposicionales y propiedades no disposicionales en términos reales no recogería ninguna diferencia. Armstrong favorece la tesis de que las creencias son estados latentes, con la salvedad de las creencias generales que sí serían disposicionales.⁷ Y la manera como Armstrong entiende los estados latentes o subsistentes de un objeto es como *propiedades* no-relacionales estructuradas.⁸

Por lo general, se admite que las disposiciones tienen bases, i.e. que se apoyan o recargan, si no es que se identifican, en otras propiedades, o conjunto de propiedades, categóricas del objeto.⁹ Para un realista la diferencia entre una propiedad-base categórica y una disposición radica en el *modo* en el que hace referencia a las propiedades del objeto: la disposición lo hace indirectamente, a través de su efecto, y la propiedad-base se introduce intrínsecamente. La identificación del modo en el cual se introducen las disposiciones y, según se acepte o no una teoría de la identidad, o serían términos teóricos muy primitivos o serían términos temporalmente aceptables mientras la teoría no cuente con términos teóricos más básicos. Un objeto frágil es aquel que tiene una determinada estructura molecular y la propiedad disposicional de ser frágil puede identificarse con esa

⁷ En la idea de que las creencias generales son disposiciones Armstrong sigue a Ramsey, quien señala que las creencias generales son "hábitos de inferencia" que nos predisponen a pasar de un hecho particular a otro hecho particular. Cfr. Armstrong (1973), p. 5.

⁸ Armstrong señala que las creencias son propiedades que tienen una estructura pero que la complejidad no es característica de todos los estados latentes. Ibid. p. 11.

⁹ La afirmación básica del realista es que el discurso disposicional no se sostiene por sí mismo y que aun cuando los científicos todavía no hubieran descubierto las propiedades reales del objeto, hay el supuesto de que existen. Este supuesto para autores como Quine, Davidson y Armstrong es necesario, mientras que para otros es sólo una hipótesis plausible, por ejemplo para Mackie. Cfr. Mackie (1973).

estructura. Sin embargo, a quienes identifican la propiedad disposicional con la estructura básica física se les presenta la dificultad de poder dar cuenta de la posibilidad de que una misma disposición tenga bases diferentes en diferentes objetos. Por ejemplo, la fragilidad en unos objetos podría tener como base la conexión molecular y en otros objetos tener la estructura cristalina. Por ello, es más adecuado tomar a las disposiciones como propiedades que dependen causalmente de sus propiedades-base y se construyan como *propiedades de segundo orden* (propiedades de una propiedad o complejo de propiedades).¹⁰

4.3 Racionalismo

Finalmente, se encuentra la postura racionalista, uno de cuyos representantes es Mellor. En esta postura se sostiene que las disposiciones son propiedades intrínsecas de los objetos, esto es, son propiedades que existirían en los objetos independientemente de que se manifestaran. Una diferencia con el realista estriba en el hecho de que para el racionalista no tiene por qué haber bases categóricas para las disposiciones. Por ejemplo, Mellor, Popper y Goodman consideran que puede haber propiedades disposicionales irreductibles. De hecho, para los dos últimos autores *todas* las propiedades son disposicionales. En la propuesta racionalista de Mellor se identificarían las propiedades disposicionales con las propiedades latentes. Las disposiciones implicarían enunciados condicionales pero tendrían un valor de verdad. Mellor sostiene que dos objetos no serían iguales si compartieran todas las propiedades, salvo la disposición. Además de que, por el principio de conectividad que suscribe, los objetos no pueden diferir sólo en una

¹⁰ Cfr. Prior, E., Pargetter, Jackson, P. (1982).

propiedad.¹¹

Otra diferencia entre Mellor y los realistas es que para los realistas, por lo general, la relación entre disposiciones y sus manifestaciones es una relación lógica, mientras que para Mellor es una relación causal. Para una buena parte de los involucrados en el debate la relación causal y la relación lógica no son consideradas excluyentes entre sí. En contraposición, Mackie piensa que si es acertada la tesis de Hume según la cual la causación se establece entre dos sucesos o existencias distintas y la relación entre ellas es contingente entonces, las propiedades disposicionales, que establecen relaciones necesarias con sus manifestaciones, no pueden a la vez ser sus causas. Pero, como ya nos ha convencido Davidson, es un error pensar que son excluyentes. Pero Mellor, a diferencia de Davidson, complementa la tesis de que es compatible que un vínculo sea a la vez lógico y causal con una ontología mixta de sucesos y de propiedades. Davidson, como recordaremos, sólo admite una ontología de sucesos que pueden relacionarse causalmente. Las disposiciones serían una de las muchas maneras en que pueden describirse esos sucesos.

Refiriéndome específicamente a la postura de Mellor, él está en desacuerdo con quienes sostienen que con la atribución de una disposición no se apunta a una situación fáctica. Lejos de ello, defiende la tesis de que las disposiciones apuntan a una situación fáctica de la misma manera en que lo hace una propiedad categórica. Tampoco está de acuerdo con que la diferencia entre las disposiciones y las propiedades categóricas de los objetos se reduzca a una diferencia en el modo en el que se describen y se introducen las diferentes propiedades. Mellor enfatiza estas semejanzas para que en caso de que quede algo de la distinción entre

¹¹ Mellor (1974), p. 175.

propiedades categóricas y propiedades disposicionales -distinción que Mellor no extingue-, se trace en términos distintos a los de la existencia real y la existencia latente. También considera que el hecho de que las disposiciones impliquen un condicional subjuntivo no es privativo de los enunciados que aluden a propiedades disposicionales¹², además de que consideraría que dicha implicación sería incidental.¹³ Mellor sostiene que necesitamos algo más fuerte que un condicional material, que no conlleva que el contrafáctico sea la forma más indicada de recoger la propiedad disposicional. Por ejemplo, no satisface una de las características del contrafáctico, a saber, que el antecedente es falso: decir que un objeto es frágil no implica que no se ha tirado. Esta discusión de Mellor me parece importante, pero no es crucial para decidir la cuestión de si las disposiciones son propiedades reales. El problema que encuentro con la postura de Mellor es que la manera que él mismo utiliza para forzar la diferencia entre propiedades disposicionales y categóricas resulta poco satisfactoria. Veamos.

Mellor considera que las propiedades reales mismas, entre sí comportan variaciones muy significativas. Una de ellas es que los objetos adquieren o pierden propiedades como producto de ciertos sucesos pasados (o futuros), mientras que otras las adquieren o pierden como producto de un acontecimiento presente. Mellor incluirá a las disposiciones dentro de las últimas. Así, por ejemplo, la propiedad de tener cuarenta años es una propiedad del primer tipo, ya que se adquiere en virtud de haber nacido en x año y no en virtud de una situación presente: en este caso no cabría la pregunta de qué condiciones presentes ocasionaron que tuviera ahora

¹² Mellor no aceptaría la implicación conversas (i.e. que si se diera el efecto dadas ciertas condiciones iniciales se implicaría la disposición correspondiente) ya que -según Mellor-, "hay algo más en la fragilidad, por ejemplo, que el que se rompa al ser golpeado". (1974) p. 178.

¹³ Cfr. Mellor (1974), pp. 168-69.

cuarenta años, además de que todas las otras edades que va teniendo el individuo remiten al mismo hecho de haber nacido hace x años. Por contraste, la disolución de una sustancia cae dentro de las propiedades del segundo tipo, ya que se explica por un suceso presente, a saber, el ser depositado en agua y en virtud de tener la propiedad de ser soluble. La propiedad de ser triangular, de manera similar al caso disposicional, cae dentro de las propiedades del segundo tipo, ya que puede adquirirse o perderse según las condiciones a las que se someta el objeto. A esta capacidad de adquirir o perder propiedades la denomina Mellor *mutabilidad*.¹⁴ Mellor no podría valerse del rasgo de ser estímulo-dependientes, que atribuye Chomsky a las disposiciones, ya que este requisito excluiría a la triangularidad.

Armstrong disputa la forma del contraste entre propiedades que nos brinda Mellor diciendo que también la propiedad de tener cuarenta años cambia, ya que cambia su *propiedad relacional* respecto a su fecha de nacimiento.¹⁵ Adicionalmente, podríamos señalarle a Mellor que pese a que tanto ciertas propiedades categóricas como las propiedades disposicionales pueden ser causas que nos remiten a sucesos presentes, se soslaya el hecho de que el paralelismo se rompe en cuanto al papel que desempeñaría cada propiedad en la descripción de la explicación. Si bien la disposición explica su manifestación (es aquello en virtud de lo cual si se dan ciertas condiciones se desencadena cierto efecto) a la vez que guarda una relación conceptual con ella, no podemos decir lo mismo respecto de la triangularidad, por ejemplo. La triangularidad no sería la propiedad en virtud de

¹⁴ Cfr. Mellor (1974), p. 158. E. Prior en *Dispositions* niega que el paralelismo que ofrece Mellor mine la diferencia entre propiedades categóricas y propiedades disposicionales. Para ella la diferencia subsiste en la medida en que la relación con los condicionales subjuntivos es distinta en los dos casos. Para sustentar su crítica se vale de la semántica de mundos posibles.

¹⁵ Cfr. Armstrong (1973), p. 11.

la cual, o lo que explicara, la forma que manifiesta el objeto, ni tampoco sería su causa. Ciertamente la triangularidad guarda una relación conceptual con la forma que tiene el objeto pero no diríamos que si un objeto triangular de hule se calienta y cambia de forma, se explique en virtud de su triangularidad.

Mellor propone una prueba independiente para la existencia de las disposiciones. Su estrategia es distinta a la de algunos autores que se contentan con señalar que las disposiciones, a diferencia de otras propiedades latentes, sólo se introducen a través de sus efectos y por ello es que sus explicaciones suelen ser triviales. Son explicaciones tan poco informativas como la explicación Panglossiana de que un sedante tiene *virtus dormitiva* además de que se carece de una prueba independiente que establezca que el efecto resultó de ingerirlo. El argumento de Mellor es que la independencia se tendría de la misma manera que se tiene cuando la masa de un objeto explica la aceleración de un objeto bajo cierta fuerza y su prueba independiente consiste en su aceleración bajo otras fuerzas. La prueba independiente del sedante sería su composición química. No entraré en este argumento de Mellor ya que nos llevaría a la distinción entre sucesos y propiedades y la relación lógica que Mellor establece entre sucesos y propiedades. Sólo quisiera recalcar que las disposiciones van a tomarse como propiedades reales concebidas, como dije anteriormente, como propiedades de segundo orden.¹⁶ ¿Cómo se aplica esta tesis realista de las disposiciones a las creencias?

5. ONTOLOGIA DE LAS CREENCIAS

5.1 Análisis disposicional.

Admitamos que hay niveles de creencias en cuyo primer nivel, las llamadas

¹⁶ pp. 149-50.

creencias de primer orden (*Bp*), incluimos de manera generosa estados entre los que se encuentran los estados-C -como los llama Williams- de ciertas máquinas y los estados de ciertos organismos vivos. Argumentaré que las creencias *Bp* pueden producir disposicionalmente, creencias ocurrentes o presentes a la mente. En este nivel aparecen las creencias como representaciones mentales estructuradas relacionalmente. Se podría considerar que la atribución de *Bp* reportaría un valor cognoscitivo respecto de los sistemas u organismos y admitirían de predicción. En otras palabras, la atribución sería real y no sólo metafórica.

Ya hemos aludido a las razones más comunes que se esgrimen en contra de entender a las creencias como disposiciones a actuar. Tal vez la más decisiva de las consideraciones que llevan al rechazo de las creencias como disposiciones a actuar sea la presencia del carácter holista que caracteriza a lo mental que impide que a cada creencia se le asocie una acción de cierto tipo. Sobre todo, si se piensa, muy en particular, en las disposiciones físicas ordinarias, las cuales sí tendrían manifestaciones propias. Pero, si se pensara en las creencias como disposiciones "multi-vías", o de múltiples realizaciones, como propuso Ryle, tampoco haría factible la tesis de que las creencias fuesen disposiciones a actuar.¹⁷ La crítica sería en que cada una de las manifestaciones podría no darse, pese a que la creencia estuviera presente, si el creyente tuviera diferentes deseos, por ejemplo. Evans refuerza de manera enfática la desemejanza entre el tipo de estados-C que puede tener un animal y las creencias, en que los estados-C sí podrían considerarse disposicionalmente. Una rata, por ejemplo, manifiesta la "creencia" de que cierta comida está envenenada de sólo una manera, no comiéndola, mientras que no hay límite en las maneras en que puede manifestarse el que una persona crea que la

¹⁷ Cfr. Ryle (1963), p. 129.

comida está envenenada. Evans concede que en el nivel de entradas y salidas hay analogías entre las creencias y las disposiciones, analogía que se quiebra justamente por la infinidad de maneras en que puede manifestarse una creencia debido al funcionamiento holista de lo mental. Williams, en la misma dirección que Evans, considera que los estados-C son distintos de las creencias en sentido pleno en tanto que con ellos no se podrían hacer afirmaciones insinceras (esto es, carecerían de la característica 3). La posibilidad de hacer afirmaciones insinceras sería una de las múltiples maneras en que la aserción podría utilizarse para propósitos o proyectos muy distintos, entre los cuales pueden incluirse la simulación y el engaño al otro y en los cuales interviene de manera central la voluntad y la deliberación del sujeto. La característica 2) -la relación del estado con la conducta verbal- también se empobrecería en conexión con las estados-C. Podemos conceder que los señalamientos de Williams, Evans, Peacocke y de otros, en cuanto a la amplitud de posibilidades de intervenir, combinar, planear, etc., llevan al rechazo del vínculo disposicional entre las creencias y la acción. Sin embargo, no me parece que establezcan que el mecanismo que hay en los sistemas más simples sea un mecanismo distinto. ¿Por qué la ausencia de versatilidad lo convierte en otro mecanismo?

Otra consideración que agrega Evans en contra de considerar a las creencias como disposiciones es que hay una distancia muy grande entre el vínculo de la creencia y la acción y el vínculo entre la disposición de un sujeto a juzgar que *p* cuando se le pregunta, que Quine ofrece en defensa de la disposicionalidad de la creencia.¹⁸ Supuestamente esa misma distancia también existiría con la segunda característica que asocia Williams a las creencias, de que la expresión natural de la

¹⁸ Evans (1981).

creencia es la aserción, no entendida de manera conductista, como en Quine. Cuando Williams señala que la aserción constituye la expresión natural de la creencia, agrega que la aserción no constituye ni condición necesaria ni suficiente de la creencia (rasgo 3). La razón para decir que no es necesaria es que no todas las creencias llegan o llegarán alguna vez a expresarse. Y tampoco es suficiente ya que su expresión puede emplearse para engañar al otro y entonces ser insincero al afirmarlo. Podría, sin embargo, entenderse el vínculo de manera semántica, como en Wittgenstein, al conectar la aserción con su uso expresivo en la primera persona, como veremos más adelante, que testimonia un uso redundante de "Creo que p " y la aserción de p . Williams no hace referencia a la gama amplia de acciones a las que puede llevar la creencia o un conglomerado de creencias, deseos, etc., y se centra sólo en la relación de la creencia con la aserción verbal. En todo caso, las críticas a la tesis de que las creencias se consideren como disposiciones a actuar parecerían obstaculizar la caracterización de creer p como una disposición a afirmar p , cuando fuese sincera. En cualquier caso, la solidez de estas críticas deja abierta la posibilidad de que las creencias Bp pudieran considerarse como disposiciones a asentir internamente. En particular, se deja abierta la posibilidad de que las creencias de primer orden pudieran considerarse como disposiciones a asentir internamente o a formarse creencias de segundo orden. Esto es, no sería un argumento contundente contra la inclusión de otros estados como estados de creencia. Por ejemplo, no se excluye que la característica evolutiva del diseño humano compartiera su mecanismo y función con diseños más simples evolutivamente. Como dice Dennett, que no hay un momento "mágico" en el cual aparece ese mecanismo *sui generis* y que no lo tuviera ningún otro ser vivo.

La condición 1) que señala Williams -que las creencias apuntan hacia la verdad- restringe que las creencias se impongan y, por tanto, establece un límite al

control que el sujeto pudiera ejercer sobre ellas. Las creencias no no deben exponerse a la voluntad, ya que si se hiciese podría producir creencias irracionales. En efecto, la voluntad no puede "manipular" y acomodar las creencias a su albedrío sin exponer su liga con la verdad. La expresión de la voluntad y del aspecto deliberado que acompaña a las creencias puede desglosarse como desear creer, y sabemos que la voluntad no se rige por los dictados de la verdad. Como dice Williams: "No podría, con total conciencia, verla como una de mis creencias, esto es, algo que doy por verdadero, y también saber que la adquirí a voluntad."¹⁹ Pero esto no impide que el aspecto deliberado o intencional a veces intervenga en la formación de ciertas creencias produciendo el auto-engaño. El aspecto deliberado interviene como forma de impedir que una creencia verdadera llegue a la conciencia del sujeto. Pero antes de pasar al caso del auto-engaño y la forma de explicarlo, veamos primero cómo se conecta la creencia con el asentimiento.

Vimos que Williams destacaba como primera característica de las creencias que cuando un sujeto dice "Yo creo que *p*" manifiesta su adhesión o compromiso con la verdad de *p*, excepto en el caso en el cual la emplea como una forma resguardada o cautelosa de expresar su actitud respecto a la verdad de *p*. La fuerza asertiva que lleva implícita la creencia se reflejaría en el hecho de que si *S* descubriese la falsedad de *p*, se retractaría de su creencia. Williams sostiene que la manera usual de expresar una creencia no es aseverando que cree *p* sino afirmando, simple y llanamente, *p* (rasgo 2).²⁰ En este mismo sentido se pueden interpretar los pasajes de Wittgenstein al suscribir el uso redundante de la creencia en la mayoría de sus uso en la primera persona. En esta lectura se apunta a una de las

¹⁹ Williams (1973), p. 148.

²⁰ Cfr. p. 138.

tesis centrales de Wittgenstein respecto de los usos "expresivos" de muchas de las afirmaciones que se hacen desde esta voz.

Alternativamente, podríamos interpretar la implicación a la que alude Williams entre "Yo creo que p " y la afirmación de la verdad de p en la misma tesitura que las "implicaturas conversacionales" de Grice. En forma sucinta, Grice distingue las implicaturas conversacionales de la implicación lógica, sosteniendo que aunque lo implicado no esté estrictamente contenido en el significado de lo expresado, queda sugerido por lo forma en que se da la información. Grice da el ejemplo de un sujeto que dice "S ha dejado de golpear a su esposa", en donde se implica conversacionalmente, aunque no lo dice literalmente, que S la había estado golpeando antes. En este caso también, aplicado a un caso paralelo al de la paradoja de Moore, si una persona dijera, "Yo creo que p , pero no me comprometo con la verdad de p " implicaría, griceanamente, que con la segunda parte de la afirmación de alguna manera cancela la primera parte de la afirmación. Al final del trabajo aludiré a otras formas de reacción frente a la paradoja, que más que ir contra Williams, abren un abanico de lecturas de la paradoja.

5.2 Niveles de creencias

Si aceptamos una noción de Bp podemos asociarle una noción de asentimiento implícito entendido como un impulso epistémico -que De Sousa denomina "lujuria epistémica"- que nos hace recolectores de oraciones verdaderas²¹, y que lleva tanto a animales como humanos a actuar como si sus creencias fuesen verdaderas. En cierto sentido, Bp se impone al organismo y puede construirse como una "disposición a asentir". Si se dieran ciertas circunstancias, se

²¹ Cfr. De Sousa (1971), pp. 52-3.

actuaría *ceteris paribus* como si la creencia fuese verdadera. En los seres que son capaces de formarse creencias ocurrentes, el asentimiento trae las creencias a la conciencia y las hace propias. Con esta función el asentimiento toma del acopio de información que tiene, las que se usarán como premisas de argumentos, i.e. las premisas que guiarán conscientemente las acciones del sujeto. En ese momento el sujeto es capaz de formarse opiniones o creencias de segundo orden (BBp); el sujeto decide asumir esa opinión.²² Dennett retoma estas ideas y las incorpora en su modelo de niveles de teoría: sub-personal y personal.

Dennett sugiere que las creencias y los deseos -considerados como ilata- se ubiquen en el nivel de la psicología sub-personal. Esto es, que se construyan como términos teóricos que se infieren a partir de lo observable. Dennett recurre al modelo disposicional para las creencias, y liga éstas a la conducta observable. Los dos niveles de la teoría -sub-personal y personal- se refieren a la realización específica de un sistema intencional. Dennett analiza el sistema cognitivo como un conjunto de homúnculos, cuyo funcionamiento en su nivel más general se describe empleando vocabulario intencional. El sistema, a su vez, se descompone en diferentes sub-sistemas, a los cuales se les asignan funciones específicas -que también se caracterizan intencionalmente-, hasta finalmente llegar al funcionamiento meramente mecánico del sistema.²³ Los sub-sistemas, que tienen funciones específicas, a la vez colaboran con otras actividades que se desarrollan en otros sub-sistemas. Así, por ejemplo un sub-sistema visual puede pasar información al sub-sistema motriz o a un sistema de otro nivel haciendo al sujeto consciente de lo que

²² En este sentido me parece que Hookway no aprecia adecuadamente la doble función que puede desempeñar el asentimiento, cuando acusa a De Sousa de dejar fuera la deliberación que guía las acciones y que es conscientemente controlada. Hookway (1981).

²³ Cfr. Dennett (1978a) pp. 123-4 y (1978b) pp. 81-82.

está viendo. No habría para Dennett nada misterioso en que un equipo de "relativamente ignorantes, estrechos y ciegos homúnculos puedan producir una conducta general inteligente". Lo importante para Dennett es que los diferentes homúnculos que se introducen eventualmente pueden eliminarse. Mientras no se eliminen, de cualquier forma la postulación particular de esas entidades teóricas no requieren de representaciones internas.

Este modelo explicativo que se emplea básicamente en Inteligencia Artificial, se puede aplicar también a los humanos, excepto que en el modelo cognitivo humano se agregaría el nivel personal en el cual Dennett ubica a las **BBp**. Este nivel de la psicología personal sólo lo tienen los sistemas que desarrollan un lenguaje y adquieren capacidades de auto-reflexión.

Las opiniones, nos dice Dennett, son estados mentales cognitivos en los cuales, a grandes rasgos, "apostamos a la verdad de una particular oración que formulamos"²⁴ Estas opiniones serían el resultado de lo que hemos denominado asentimiento. Si bien Dennett admite que en este nivel podemos formarnos representaciones internas o pensamientos particulares, este rasgo no es esencial a las opiniones; éstas se forjan o se seleccionan de las creencias que se encuentran en el nivel sub-personal, sin ser en sí mismas creencias, sino actos. Si entiendo correctamente, estos actos u ocurrencias mentales que están presentes a la conciencia, no son producidos por el sujeto, sino que son manifestaciones de **Bp**: la relación entre los dos niveles sería disposicional. Cuando se hacen conscientes, el sujeto puede establecer un compromiso frente a ellas generando otros estados, las opiniones.²⁵ Dicho de otra manera, las opiniones son estados en los que revela

²⁴ Cfr. Dennett (1987a), nota p. 19.

²⁵ Cfr. Dennett (1978c), p. 303.

el compromiso del sujeto de tomar como verdaderos esos pensamientos, que por lo general son oraciones. Lo distintivo de la posición de Dennett, al igual que en la de De Sousa²⁶, es que no es correcto considerar que la conciencia es condición necesaria del asentimiento. Las creencias de segundo orden (BB ρ), por lo general, son conscientes, pero no es necesario que lo sean. Esta consideración es importante ya que se trata de que el ser consciente no sea constitutivo de esas creencias y de que lo consciente no sea el eje de la racionalidad. Dice Dennett de las opiniones o creencias de segundo orden (BB ρ), lo siguiente:

En todo caso, hay muchas maneras de incrementar nuestra colección de opiniones... Uno puede heredarlas, poseerlas sin percatarnos de ello, no eliminarlas después de haber decidido hacerlo, tomarlas como préstamo provisional y olvidar que eso es lo que hemos hecho.²⁷

Señalé que para Dennett la distinción entre creencias disposicionales y recurrentes no corresponde a la distinción entre creencias y opiniones, ya que para Dennett las opiniones son estados del sistema, y pudiera ser que la manera en que se conserven las opiniones no fuera en forma de oraciones. Dennett concibe la intencionalidad como un vínculo que se establece directamente entre el nivel subpersonal y la conducta, que sólo se daría de forma derivada en cuanto relación entre el nivel personal y la conducta. También considera que nuestro entorno social lingüístico nos fuerza constantemente a dar expresión verbal precisa a convicciones y creencias que carecen de la rigidez que la verbalización imprime.²⁸ Este

²⁶ De Sousa (1971).

²⁷ *Ibid.* Dennett (1978c), p. 307.

²⁸ *Ibid.*, p. 21.

señalamiento de Dennett no tiene que excluir, sin embargo, que a veces expresamos creencias muy específicas. Si bien la creencia en la que descansa tiene que existir en el nivel sub-personal, no tiene que hacerlo con esa característica. Dijimos que este tipo de rasgo no tiene que recogerse en una función evolutiva ni en una creencia sub-personal (Bp). Su descripción adecuada puede admitir un contenido más general, pero no que carezca de contenido.

McDowell critica cualquier teoría que introduzca en el mero acopio de información alguna forma de conciencia y también critica que se les dé a las Bp algún tipo de realidad. Para McDowell lo que legitima el que se hable de los niveles sub-personal y personal, que se entiende como relación entre sub-partes del sistema y que esta relación sin mayor argumento no se puede hacer extensiva a la relación entre parte y todo. Para McDowell no es lo mismo, por ejemplo, lo que le dice el ojo de la rana a su sub-sistema motriz o a su cerebro, que lo que le dice el ojo de la rana a la rana.²⁹ Para él, la relación cognoscitiva, en la cual aparece la noción de contenido, sólo se da en la relación del sub-sistema al sistema completo o a la persona. La persona misma es el receptor de la información que debe considerarse como información proveniente, no de una sub-parte, sino del entorno. Me parece que McDowell tiene razón al señalarle a Dennett que la conformación del contenido tiene que venir del vínculo que se establece con algún rasgo o aspecto del entorno. Ya señalamos en el capítulo anterior que ésta es una modificación bienvenida que también se asume en los modelos biológico-funcionales como los que suscriben Millikan y McGinn. McDowell expresa la crítica diciendo que la información del entorno no se puede reconstruir a partir de los datos que se recogerían a nivel de un sub-sistema, que es básicamente la

²⁹ Cfr. McDowell (1994), p. 194.

respuesta de Kant frente al *impasse* que produce la postura de Hume. Sin embargo, la otra crítica de McDowell, en la cual sólo admite que el receptor de la información, que es la persona consciente, descansa en la distinción entre la captación de una información y la transmisión de la información que no es una distinción muy clara, de modo que justifique el rechazo de la idea de que el sujeto introspecciona la información que está contenida en el nivel sub-personal. Para McDowell, en el caso de la rana, por ejemplo, decir que la información que se encapsula en el nivel sub-personal no se da en los términos o con los contenidos con los que nosotros interpretamos la relación que establece el sub-sistema visual de la rana y el entorno, i.e. que esos contenidos son nuestra manera de hablar acerca de lo que sucede en la rana y que de allí no es legítimo atribuírselo a la rana: la rana meramente reacciona a un estímulo externo, recogido -digamos- en términos de luz-, que podría resultar ser un insecto, que es muy distinto a decir que la rana reaccione al insecto. Cuando hablé del funcionalismo biológico consideré que la rana genuinamente se vincularía con un rasgo externo -de manera externalista débil- que tiene conexión con el insecto, cuando despliega una conducta que puede interpretarse como un movimiento para comérselo, se fija a partir de la función visual y no como hace McDowell de conectarlo de una manera externa fuerte con la experiencia de la rana. La función biológica se fija empleando ciertos contenidos y no tiene por qué pedirse que la rana como tal tenga que reconocerla como tal para poder llevar a cabo su función.

5.3 El asentimiento

Veamos ahora la postura de Mellor, como un ejemplo en el que se aceptan los niveles de creencias en cuyo segundo nivel aparecen las creencias como representaciones mentales estructuradas relacionamente y que se identifican con el

asentimiento y se caracterizan por ser conscientes. Mellor inclusive sostiene como una de sus tesis principales la identificación de las creencias de segundo orden con el asentimiento. Mellor aborda el problema de las creencias desde la perspectiva de la teoría de la acción. Él no acepta el tratamiento de las creencias, en cualquiera de sus niveles, como disposiciones. Sin embargo, introduce las disposiciones con respecto del auto-conocimiento y el auto-engaño y remite a ambas a BBp . El auto-conocimiento se forma como una disposición a asentir y el auto-engaño como una disposición a disentir. La razón por la cual Mellor no identifica BBp con el auto-conocimiento se debe a que acepta que podemos o equivocarnos o auto-engañarnos con respecto a nuestras creencias. La disposición al auto-conocimiento (y al auto-engaño) es lo que explica que juzga que explica en un momento dado y en ciertas condiciones el sujeto asienta a p (o disienta de p). Mellor entiende por auto-engaño cuando alguien que (sub-conscientemente) cree p , mientras que (conscientemente) BBp , cree que cree la negación de p , y el observador es quien se encarga de hacer las atribuciones, con base en diferentes aspectos de su conducta.³⁰ Dada la naturaleza del auto-engaño, el sujeto no puede estar consciente en el momento en que está siendo víctima de él. Mellor identifica el nivel Bp con el nivel que escapa a la conciencia y el nivel BBp con el nivel consciente, en el cual el sujeto está consciente tanto de p como de Bp . Sin embargo, ni esta conexión es correcta, ni tampoco su identificación entre el asentimiento y BBp . Mellor construye el acto de asentimiento como un acto de reflexión y no como diríamos nosotros de asumir lo verdadero y en donde Bp se vincula disposicionalmente con su asentimiento. El asentimiento, en nuestro uso, se distingue de BBp .

Mellor supone que tanto Bp como BBp pueden darse separadamente en un

³⁰ Cfr. Mellor (1978), p. 91.

sujeto. Las situaciones que ilustrarían el caso de que el sujeto tuviera BBp donde faltaría la creencia correspondiente de primer orden serían justamente las del auto-engaño. Los observadores atribuirían al sujeto S la creencia en p , creencia que S negaría tener y en cambio S creería creer su negación: ya sea que cree que no cree p ($B-Bp$) o que no cree que cree p ($-BBp$). Pero, en el caso en el que Bp llegara a la conciencia, entonces BBp implicaría Bp al igual que la requiere causalmente. Esto no garantiza -dice Mellor- la creencia captada por el sujeto llegue sin distorsiones: una creencia puede percibirse inadecuadamente ("*mispercieved*").

La identificación entre la creencia de segundo orden con el asentimiento constituye la tesis principal (T_1) de Mellor, en la cual el asentimiento se refiere a un estado mental. Para explicar esa noción se vale de la noción de creencia de primer orden, y no a la inversa. La identificación entre BBp y el asentimiento la sustenta Mellor simplemente en que no puede imaginarse un caso en el que haya una BBp y que no se tenga conciencia o se asienta a p e igualmente la conversa: que se tenga esa conciencia pero que no se tenga la BBp correspondiente. Mellor concede que su argumentación no constituye una prueba, pero que sin embargo le da apoyo para conjeturar que la coincidencia se debe a su identidad.

Mellor sostiene que la primera dirección de la implicación, asentir implica creer que uno cree, es obvia. Nos dice:

Assent is the conscious belief in p that is required for the sincere affirmation of p 's truth. In coming to assent to p I have perceived (or in cases of self-deception, misperceived) my belief in p : if that does not involve believing one believes p , what does?³¹

³¹ Mellor (1978), p. 92.

Parece correcto que el sujeto al percibir p está siendo consciente de su creencia y esto supone BBp . Y para la implicación conversas -de que crea que cree p implica asentar a p - se preguntaría Mellor si es posible que hubiera una creencia de segundo orden sin que hubiera conciencia de p . En otras palabras, si pudiera tener la creencia de segundo orden en los periodos en los que no está siendo consciente, o no está percibiendo su creencia. Los casos en los que inmediatamente se pensaría -según Mellor- serían los casos de auto-conocimiento y de auto-engaño, pero entonces su maniobra, como ya señalé, es de negarles el estatuto de creencias. Sin embargo, podríamos negar que el auto-conocimiento sea una disposición a BBp ya que podría haber BBp sin auto-conocimiento. Más bien, el auto-conocimiento podría verse como una atribución que hace el observador cuando supone que el sujeto tiene Bp y BBp . Y en el caso del auto-engaño podríamos decir que no se dan conjuntamente Bp y BBp . ¿Podríamos decir que la distorsión o la negación de que creemos p debido justamente a la intervención de deseos y otras creencias que tenemos? ¿Por qué tiene que deberse, como Mellor supone, a que el sujeto tiene la *disposición a disentir* de las proposiciones que en realidad cree, o la disposición a no hacer conscientes las proposiciones que cree. ¿Por qué invocar una disposición adicional, cuando podríamos dar cuenta de la situación con la idea de que el sujeto cree p (Bp) en conjunción con otras creencias y deseos que interfieren para que el sujeto haga esa creencia consciente y que en cambio propicien que crea la creencia contraria? Cuando Mellor hace referencia al engaño nos dice que éste se produce en el nivel de BBp : es el engaño en cuanto a la reflexión que hace el sujeto pero que el sujeto no se percata del engaño: el engaño se lo atribuye otro sujeto.

Además del error antes señalado, de que Mellor identifica la creencia de segundo orden o el asentimiento con el ser consciente, tiende también a quitarle

perplejidad al fenómeno del auto-engaño, dejando fuera la intervención del sujeto para sabotear Bp. En la versión de Williams que vimos anteriormente, el auto-engaño tiene que ver más bien con lo que sucede en el fuero interno del sujeto, los mecanismos y encubrimientos que tienen la función de poner un velo a lo que cree. O como dice Pears, el sujeto trata de aislar funcionalmente a la creencia. A pesar de los empeños del sujeto, el observador logra descubrirlo en aspectos de sus acciones. Mellor, en cambio, dirige las creencias de primer orden y las creencias de segundo orden a diferentes esferas o compartimentos de la mente (esfera inconsciente y esfera consciente respectivamente), aislándolas, restándole perplejidad al fenómeno mismo del auto-engaño, perplejidad que justamente tendría que explicarse. Para algunos, por ejemplo, la idea del auto-engaño supone la referencia a algo irracional o inconsistente, en la medida en que el sujeto intencionalmente se propone desconocer la creencia que tiene.³²

¿Qué pasaría con la dirección de la implicación de BBp al asentimiento a la luz de una manera distinta a la de Mellor de considerar el auto-engaño? ¿Depende crucialmente el éxito de la implicación de la idea de que el auto-engaño sea una disposición? Para Mellor es determinante que el auto-engaño (y el auto-conocimiento) pueda darse sin que el sujeto tenga la creencia en la mente en determinado momento; diría lo mismo respecto a BBp. Es difícil entender la tesis de que el auto-conocimiento y el auto-engaño en tanto *disposiciones*. No me parece necesario recurrir a su postulación, como un elemento adicional, cuando uno podría decir que el hecho mismo de habernos formado una creencia, llevaría a su asentimiento, *ceteris paribus*.

En síntesis, en Mellor el uso de los niveles de creencias le permite que en el

³² Cfr. por ejemplo, Davidson (1985).

auto-engaño se asignen las creencias en conflicto a diferentes niveles de creencias, independizando un nivel del otro. Lo que he intentado cuestionar es esa supuesta independencia que da como resultado que, más que explicar el fenómeno del auto-engaño, lo deja sin resolver.

5.4 El auto-engaño

Paso ahora a la solución particular que Dennett da del auto-engaño. La manera en que Dennett utiliza el modelo de niveles de creencias para dar cuenta de los fenómenos de la *akrasia* y el auto-engaño, al igual que en el caso de Mellor, tampoco es satisfactoria. Veamos.

Por una parte, Dennett considera que tanto los fenómenos de *akrasia* y los del auto-engaño son formas en las que "uno se *comporta* de cierta manera mientras que *juzga* de otra". La semejanza entre los dos fenómenos se recoge en términos de que hay un desajuste entre ellos, dan como resultado el mismo tipo de incoherencia. La incoherencia se explica en términos de que esto puede suceder porque tenemos niveles de creencias: las creencias de primer nivel producen el comportamiento, a pesar de que el sujeto juzgue en una dirección contraria. Las creencias de primer orden son inconscientes y los juicios, independientemente de que sean conscientes o no, se reflejan en otros aspectos de su conducta. La inconsistencia en las conductas se explica por el hecho de que las creencias asignadas a diferentes niveles están funcionando independientemente. Sin embargo, se le puede señalar a Dennett que este funcionamiento independiente requiere, a su vez de explicación. No es satisfactorio decir que algún deseo se interpuso para que la creencia que se tiene (*Bp*) formara la opinión correspondiente. Debe haber habido algún mecanismo que explique que el sujeto se forme la opinión contraria, que supone que el propio sujeto de algún modo solapa intencionalmente *Bp*.

Contrariamente a lo que dice Dennett, la *akrasia* y el auto-engaño no son sólo casos de "*wishful thinking*" o de desear creer. Tal vez sea más claro que no es una cuestión de desear creer en el caso de la debilidad de la voluntad. Pero aunque sí puede interpretarse como una forma de desear creer en el caso en el fenómeno del auto-engaño, el sujeto debe de alguna manera percatarse de la creencia que quiere negar. Dennett convierte la irracionalidad no es un mero caso de inconsistencia de creencias; el sujeto tiene que valerse de algún mecanismo para impedir que su deseo de creer algo compita y se enfrente con su deseo epistémico.

Hay autores que asocian al nivel más básico de creencias la característica de ser inconscientes y al nivel BBP la de ser consciente- y al hacerlo comparten con Mellor el error de conjurar el enigma que produce el auto-engaño. Más bien, se tendría que sostener que en los dos niveles existe alguna manera de advertir lo que está sucediendo. Así como el sujeto activamente tiene que ser consciente, también tiene que proponerse dejar de serlo. Fingarette sostiene, con una cierta fuerza metafórica, que decir que el sujeto llega a ser explícitamente consciente debe tomarse como "la habilidad de decir (o descifrar) lo que está haciendo o experimentando."³³

Williams marca dos sentidos de la noción de "desear creer". Uno tiene que ver con "motivos centrados en la verdad" y el otro con "motivos no centrados en la verdad".³⁴ En el caso en el que el desear creer se conecte con los motivos que tienen que ver con la verdad, lo que el sujeto desea es la *verdad* de la creencia, e intervenirla sería tan inefectiva como querer que la realidad fuera distinta de como es. El auto-engaño que se querría lograr sería una forma de "*wishful thinking*". Un

³³ Fingarette (1969), pp. 38-9.

³⁴ cfr. pp. 149-151.

ejemplo sería el de una persona que tuviera fuertes bases para creer que su hijo ha muerto y quiere convencerse de que no lo está. Recurrir al hipnotista con el fin de inducirse creencias, es buscar algún tipo de auto-engaño; engañarse en creer lo que sabe que es falso. Es un proyecto incoherente e imposible, ya que es claro que por esos medios el sujeto no puede cambiar los hechos. Lo que el sujeto desea creer no puede ir en sentido contrario a la verdad, además de que sabe que su recurso es ilusorio. Para lograr el propósito de infundirse creencias, el sujeto tendría que saber lo que es verdadero, de modo que pudiera evitar toparse todo el tiempo con las pruebas que le reafirmaran lo que quiere negar. Williams considera que no es posible esta negación sistemática, ya que se tendría que hacer una atribución de creencias contradictorias al sujeto. Él o ella tiene que saber qué es lo que no quiere creer para luego negarlo. Este proyecto compartiría la incoherencia que también habría en el auto-engaño ya que el sujeto tiene que tener una forma de acceder a lo que cree y a la vez enterrarla u ocultarla: a la vez que cree p , cree su contraria (o sencillamente no se forma ninguna creencia respecto de p).

En el sentido en el que el desear creer no estuviera centrado en la verdad, el sujeto sabría y asumiría que el recurso buscado no va a modificar los hechos y sólo se propondría recurrir al engaño con el objeto de hacerse la vida más soportable. En este caso, Williams sostiene que el querer creer lo falso sería *irracional*. Este es el proyecto de querer olvidar con el objeto de aliviar lo que es doloroso.

Cuando uno quiere negar una creencia que le cuesta trabajo aceptar, uno no puede buscar inducirse creencias mediante procedimientos más indirectos como lo serían el uso de drogas o recurrir a medios hipnóticos, como sucedáneos de la voluntad para moldear nuestras creencias: todo el plan estaría al descubierto para el sujeto *ergo*, el auto-engaño, pese a que es auto-inducido debe encubrirse para conservar su efectividad como engaño. Si inducirse creencias es incoherente,

inducirse creencias por otros medios es irracional.

Pero yendo en una dirección distinta a la de Williams podríamos decir que el proyecto de alguien de infundirse creencias falla en tanto que en ese caso no diríamos que se trata de *sus* creencias. Habría una especie de "asentimiento" gestual, una impostura, sin que hubiera una creencia que la respaldara. Sería un asentimiento que no expresara ninguna creencia del sujeto. Sería tan artificial y ficticia como el asentimiento de una persona que dice creer algo con el objeto de engañar al otro. Un asentimiento genuino requiere la presencia de la creencia (Bp) que está expresando. Si la creencia se construyera como una disposición a asentir su generación anómala destruiría su carácter de asentimiento.

Si bien no objeto la línea general que desarrolla Dennett al hablar de los diferentes niveles de creencias, cuando lo aplica a los casos de irracionalidad cae víctima de querer dar cuenta de un fenómeno como el auto-engaño en términos de asignar las creencias en conflicto a los diferentes niveles de creencia, quitándole su aspecto perturbador.

Si apelamos a los diferentes niveles de creencias, podemos sostener que el sujeto cree p (Bp) y cree que no cree p ($B-Bp$). La falla del sujeto se atribuiría a una forma de inconciencia, a un error de conciencia. Es decir, la falla se localizaría en que el sujeto no acierta a hacer conscientes ciertos datos que sin embargo forman parte de sus creencias de primer orden (Bp). Si en una época pre-freudiana se pensaba que la conciencia permeaba a todos los fenómenos mentales, en la época post-freudiana hay autores que simplemente restringirían la esfera de lo consciente sosteniendo que un sujeto es consciente sólo de las creencias del segundo nivel. El ser consciente se convierte en una forma de asentimiento interno. Pero también puede introducirse a la manera freudiana, por un lado, la tesis de un yo dividido y, por otro lado, la división entre consciente, pre-consciente e inconsciente. Como dice

Elster, la primera es más una distinción territorial y la segunda una distinción entre agentes.³⁵ Sin embargo, en la actualidad hay una tendencia a no aceptar la partición entre agentes y hablar en cambio en términos de hemisferios cerebrales con diferentes sistemas funcionales y en donde no es claro si cada uno de ellos tiene sus formas de conciencia. Me parece que la solución del auto-engaño conectando las *BBp* con lo consciente y *Bp* con lo inconsciente no es insatisfactoria. Al menos se tendría que explicar por qué tener una creencia de segundo orden implica la conciencia y por qué *Bp* no implica alguna forma de conciencia o de percatarse ("*awareness*") de *p*. Uno podría tomar la creencia de segundo orden, *BBp* implica ser consciente de *Bp* y de *p*, como una instancia de sustitución de *Bp* y en donde no se vería claro que se les asocien diferentes atributos. Pero uno también podría pensar que en este tipo de solución el auto-engaño se toma como un simple error -sea de conciencia u otro- sino dar cuenta de que hay alguna forma de intención escondida. Se omite que hay un conflicto de creencias en donde el sujeto lejos de "ignorar" que tiene cierta creencia evita tenerla, para lo cual se diseña una estrategia o coartada. Sartre en el Ser y la nada expresa la idea paradójica que subyace al auto-engaño en la forma de que el sujeto se miente a sí mismo, utilizando el mismo modelo que cuando una persona miente de mala fe ("*mauvaise foi*") al otro. Dice Sartre que cuando la persona que miente y la persona a la que se miente es la misma, esto quiere decir que "yo debo saber en mi capacidad de engañador la verdad que se me oculta en mi capacidad de engañado. Mejor aún, yo debo saber la verdad de manera exacta *para* ocultarla más cuidadosamente."³⁶ Esta situación nos obliga a buscar el motivo en los otros

³⁵ Elster (1985), Introducción, p. 20.

³⁶ Sartre (1943), p. 89.

elementos que conforman la psicología del sujeto. Pero también nos obliga a decir algo más respecto a la relación que se establece entre las creencias de los diferentes órdenes.

En un modelo biológico-funcional, si se acepta una forma de intencionalidad que no supone una capacidad auto-reflexiva o de introspección, lo enigmático del fenómeno del auto-engaño pierde parte de su misterio. El proceso de encubrimiento cumpliría la función de aminorar la ansiedad que ciertas creencias le pueden producir a la persona, sin necesitar que medie una forma de conciencia plena, pero que tampoco le quita al proceso un propósito o finalidad.

IV. ADSCRIPCIONES Y AUTO-ADSCRIPCIONES DE CREENCIAS

6.1 Usos expresivos y descriptivos de los enunciados de creencia: Wittgenstein

En el capítulo de las explicaciones hicimos referencia a que las consideraciones pragmáticas jugaban un papel importante y los comentarios que a continuación señalaremos en relación con Wittgenstein son una buena ilustración de ello.

La postura de Wittgenstein se ubica dentro de las teorías disposicionales no reductivistas de la creencia, al igual que de otros estados mentales: ni reduce los fenómenos mentales a la conducta, ni los identifica con procesos neurofisiológicos. Wittgenstein se vale del contraste entre la perspectiva de la primera persona y la de la tercera persona para decir que en las afirmaciones en la primera persona respecto a las creencias propias, más que ser descriptivas o informativas respecto a lo que el sujeto cree son expresivas. Wittgenstein también negará que esos actos lingüísticos sean efecto o resultado de la creencia en cuestión; i.e., niega que su vínculo sea causal. Wittgenstein defiende que el vínculo entre el estado mental y su manifestación externa no es independiente. La manifestación externa como "expresión" de lo que sucede internamente, nos permite reconocerla como su "expresión natural", y así justificar nuestra atribución. Wittgenstein no cree posible que uno capture el estado mental, y que una vez capturado, pudiera asociársele una expresión. La tesis de la expresividad de la conducta verbal y no verbal respecto a los estados internos, constituye uno de los aportes más originales de Wittgenstein, a la vez que uno de los más controvertidos. Pero veamos qué entiende Wittgenstein por "expresión". Una característica de la expresión, que en el caso de las

sensaciones es más patente, es que no es posible tener la sensación sin su expresión correspondiente.³⁷ Una expresión natural del dolor, por ejemplo, es que la persona grite, se retuerza, haga ciertas gesticulaciones, etc.. Estas expresiones naturales, una vez que se adquieren las capacidades lingüísticas, son reemplazadas por expresiones verbales, y en cuanto sustituciones también serán expresión de la sensación. Dice Wittgenstein: "La expresión natural del dolor reemplaza al grito y no lo describe."³⁸ Sin embargo, Wittgenstein concede que no contamos con manifestaciones naturales para cada uno de nuestros estados mentales, ya que una misma expresión podría estar asociada a diferentes conceptos, y pone como ejemplo la expectativa.³⁹

Wittgenstein considera que la creencia cuenta con una expresión natural, que es una expresión verbal: la expresión "yo creo que...". Wittgenstein sostiene que cuando una persona se auto-adscribe una creencia, esta afirmación nos arroja luz sobre el estado de la persona que lo dice, y nos permite extraer conclusiones a partir de lo que dice. La expresión de nuestras creencias en la primera persona no es informativa o descriptiva. Wittgenstein entiende la relación descriptiva de manera semejante a la relación referencial entre nombre y cosa nombrada. En una situación descriptiva, el sujeto observa la situación objetiva y como resultado de la observación describe la situación. Otra característica de las descripciones es que cuentan con criterios que nos permiten que juzguemos si la descripción es verdadera o falsa.

Wittgenstein aplica su uso de la noción de expresión, en contraste con la

³⁷ Wittgenstein (1953) & 308-10.

³⁸ Wittgenstein (1953) & 244 y (1956) & 313.

³⁹ Cfr. Wittgenstein (1953) & 577.

noción de descripción, a la paradoja de Moore, para mostrar que no siempre arroja una contradicción, contrariamente a la opinión de muchos autores respecto de la paradoja, entre ellas el propio Moore.

En el contexto de la discusión de las sensaciones, Wittgenstein nos dice cómo es que las palabras se refieren a las sensaciones, que no debe confundirse con la manera como se refieren las palabras a objetos externos, a saber, como si se tratara de una relación entre nombres y cosas nombradas. Las palabras que se refieren a estados internos no describen porque supondría que se refieren a ciertos objetos, a saber, sensaciones o estados y para Wittgenstein es incorrecto ya que supondría que el sujeto "observa" lo que sucede internamente y como resultado de esa observación le asocia la palabra. Pero para Wittgenstein esa supuesta observación no tiene por qué ser reglamentaria: a veces se daría y a veces no. Me parece que aquí radica la razón central de Wittgenstein para negarles a las expresiones verbales el estatuto de descripciones: si lo fuesen nos comprometería a pensar en un referente - que en el caso de la creencia sería un pensamiento- al cual aludieran esos actos verbales. No se requiere que el pensamiento esté presente para luego extraer o derivar de él su descripción.

Algunos consideran que la tesis de Wittgenstein es la tesis fuerte de que las auto-adcripciones que hacen referencia a estados mentales nunca tendrían carácter descriptivo. Consideran que si bien es un acierto de Wittgenstein hablar de usos expresivos asociados a los actos lingüísticos, sin embargo es erróneo no adjudicarles también un uso descriptivo. Me parece que Wittgenstein no niega que haya usos descriptivos, aunque tal vez no conceda que estos usos sean muy frecuentes. Pero aquí debemos tener cuidado de separar la tesis expresiva cuando se refiere a las sensaciones y a las emociones, y cuando se refiere a los casos como la creencia. En los primeros, Wittgenstein es explícito en que el uso expresivo -tanto verbal como

no-verbal- es el predominante, pero no es claro que lo sostenga con esa generalidad para el caso de las creencias. No obstante, me parece correcto el comentario que hace Strawson en el sentido de que Wittgenstein exagera el énfasis en los usos expresivos, ya que nos hacen suponer que los actos verbales se nos imponen, cuando muy frecuentemente son actos intencionales que tienen el propósito de comunicar y de informar a los otros lo que creemos.⁴⁰ Volveré más adelante a este punto.

Wittgenstein no explora el concepto de la creencia con la profusión y el interés con los que, al hablar del lenguaje privado, explora otros conceptos mentales. Wittgenstein encuentra más semejanza entre el concepto de la creencia y los conceptos de tener la expectativa, desear y entender.⁴¹ La razón pareciera ser que una misma proposición puede ser manifestación de cualquiera de ellas, además de que para que se considere que se trata de cualquiera de esos estados, no es necesario que el sujeto en ese momento esté pensando en ello. Lo importante para Wittgenstein no es lo que pasa en ese momento, sino la circunstancia que rodea al fenómeno: lo que pasa antes y después. Sus comentarios más importantes al respecto se hallan en la parte segunda de las Investigaciones Filosóficas, en los comentarios preparatorios a esta obra, recogidos en los Remarks on the Philosophy of Psychology y en Zettel.⁴² En las Investigaciones Filosóficas Wittgenstein escribe:

Pienso esto: creer es un estado mental. Tiene duración; y es independiente de la duración de su expresión en una oración, por

⁴⁰ Cfr. Strawson (1968), pp. 55-6.

⁴¹ Cfr. Wittgenstein (1953) § 573.

⁴² Wittgenstein, Zettel.

ejemplo. Es por tanto una especie de disposición del que cree. Se me revela, en el otro, por su conducta; por sus palabras. Tanto por la expresión "Yo creo que..." lo mismo que por la simple aserción.- ¿Qué pasa en mi propio caso: cómo reconozco yo mi propia disposición?- ¡Yo tendría que hacer como el otro, ponerme atención, escuchar mis palabras, ser capaz de sacar de ellas conclusiones!

Yo tengo hacia mis propias palabras una relación completamente distinta de la de los demás. Si yo pudiera decir "Me parece que creo", ese otro desarrollo del verbo hubiera sido posible.⁴³

Wittgenstein señala claramente el contraste que priva entre el acceso que el sujeto tiene respecto a sus propias creencias y la manera en que acceden a ellas los demás. A la vez cuestiona que la diferencia entre la perspectiva en la primera persona y la perspectiva en la tercera persona vaya acompañada del contraste entre el auto-conocimiento y el conocimiento de los demás. Wittgenstein le arrebató al sujeto el privilegio de tener un conocimiento directo y certero que a la vez convertía el conocimiento de los demás a un conocimiento indirecto, hipotético y expuesto a la duda. Wittgenstein intenta resistir prácticamente todas las características epistemológicas que tradicionalmente se han trazado para establecer el contraste entre la primera y la tercera persona. El juego del lenguaje de la auto-atribución supone, para Wittgenstein, un aprendizaje previo en las que el sujeto en las situaciones semejantes a la suya concuerda con los otros al hacer atribuciones de estados mentales.

Wittgenstein introduce la discusión de que la expresión "Yo creo que p" tiene el mismo significado que la aserción de "p", en combinación de esta tesis con la tesis expresiva del lenguaje, y con ellas ilustra los diferentes usos que pueden

⁴³ *Ibid.*, x p.192.

asociarse con la paradoja de Moore. No en todos se obtendría una contradicción.

Para Wittgenstein hay casos -los más usuales- en los que la expresión "Yo creo que p" puede identificarse con la aserción de "p". La sustitución se hace con base en que "reaccionamos más o menos de la misma manera cuando alguien dice la primera que cuando dice la segunda"⁴⁴ o, a que tienen el mismo uso. Recordemos la discusión de esta característica en la discusión de Williams y la conexión que hicimos de ésta con las implicaturas griceanas. En la terminología de Rosenthal⁴⁵ equivaldría a decir que tienen las mismas condiciones de asertabilidad. La sustitución de las dos expresiones pone a Wittgenstein en situación de poder restarle a la expresión "Yo creo que p" el carácter auto-referencial que le han otorgado muchos autores, poniendo a la cabeza de la lista a Descartes. Ya vimos que el modo referencial - según Wittgenstein- nos lleva erróneamente a suponer que hay un sujeto, el hablante, el cual se refiere a su estado mental de creencia. En Wittgenstein el acento está en que el sujeto no se refiere a su estado mental descriptivamente, empero no niega que el estado del sujeto se expresa o se pone de manifiesto en su afirmación.

La situación del sujeto que hace la aseveración no es análoga a la de los observadores externos en tanto que para hacer la aseveración ni observa su conducta (y en consecuencia no establece ninguna hipótesis), y no podría decir, dándole algún matiz de incertidumbre, "Me parece que creo". Pero de igual manera que el sujeto no observa su propia conducta, tampoco acepta Wittgenstein que el sujeto -por lo general- "vea internamente" su estado mental. Esta es la imagen que

⁴⁴ Wittgenstein, Remarks on the Philosophy of Psychology, p. 477.

⁴⁵ Cfr. Rosenthal (1986). Las condiciones de asertabilidad se distinguen de las condiciones de verdad en el sentido de que dos proposiciones pueden tener las mismas condiciones de asertabilidad y sin embargo diferentes condiciones de verdad, como sucede justamente con el caso de "Yo creo que p" y la aserción de "p" o el caso en el que yo diga "Gracias" o que diga "Yo estoy agradecido".

muchas veces nos viene a la cabeza cuando autoreflexionamos y que tenemos frente a nosotros una representación. La mayoría de las veces la autoreflexión no se apega a este modelo, de allí que la idea de asentimiento a esa representación tampoco resulte feliz. En otras palabras, este no es un buen modelo para el autoconocimiento.

Ahora bien, en dado caso que una persona llegara a formarse la hipótesis de que cree que p, según Wittgenstein, sería ilegítimo sustituir "yo creo que p" por "p": la hipótesis de que yo crea que p no es equivalente a la hipótesis de que p sea el caso.⁴⁶ La razón por la que la sustitución es incorrecta parece ser que la suposición de que yo creyera p no tiene la función de describir, como tampoco la de expresar el estado mental en el que se encuentra el sujeto, sino la de suponer hipotéticamente que él o ella tuviera la creencia en cuestión. Vemos que esta situación hipotética claramente nos remite al sujeto y no a la hipótesis de que p sea el caso. En esta situación la característica general que invoca Williams de que tienen un uso centrado en la verdad, con referencia a los hechos, no se sostendría. Aquí lo importante es el estado en el que se encuentra el sujeto, como estado psicológico, que en ciertos casos pasa a segundo plano su función de no ir contra la verdad.

6.2 La paradoja de Moore

Wittgenstein aplica la distinción entre el uso redundante y el uso no redundante de "yo creo" a la llamada paradoja de Moore, (la aserción "Yo creo que p, pero p es falsa"):

La paradoja de Moore se puede expresar así: la aserción
"Yo creo que esto es así" se usa de manera similar a la

⁴⁶ Cfr. Wittgenstein (1963). II-X pag. 190a.

afirmación "Esto es así";y, no obstante, la suposición de que yo creo que esto es así no se usa de manera similar a la suposición de que esto es así.⁴⁷

En la tradición cartesiana -e inclusive para Moore- lo enigmático de afirmaciones como las de la paradoja radica en que a pesar de que las aserciones son independientes entre sí, -ya que una se refiere a algo interno y psicológico mientras que la otra a algo externo, i.e. a los hechos- sin embargo la unión de ambas produce una afirmación que en lugar de ser coherente y normal, como cualesquiera otras afirmaciones independientes entre sí, es absurda. Wittgenstein difiere de esos autores en dos puntos: 1) que a pesar de hablar de cosas distintas, la aserción "yo creo que p" sea independiente de la aserción de "p" y 2) que su aserción conjunta produzca una aserción absurda. Wittgenstein alude a la no independencia de las dos proposiciones por ejemplo en el \S 490 de los Remarks...:

Esta es la paradoja: la suposición se puede expresar de la siguiente manera: "Supóngase que esto sucediera dentro de mí y aquello afuera" -pero la aserción de que esto está sucediendo dentro de mí afirma: esto está sucediendo fuera de mí. En tanto suposiciones las dos proposiciones acerca de lo de adentro y lo de afuera son totalmente independientes, pero no en tanto aserciones.

La idea de la independencia de las proposiciones hace plausible la idea - compartida por dualistas e idealistas- de que habría un discurso referido a nuestra interioridad, cuya significatividad se mantendría aun cuando no hubiese mundo

⁴⁷ Wittgenstein. *ibid.*, x p.190e

externo. La no independencia de las dos aserciones también le permite a Wittgenstein explicar lo que sería inexplicable para alguien que sostuviese la independencia entre las dos proposiciones, a saber, que cuando dos personas tienen creencias contrarias, el hecho de que se refieran a algo interno, no quiere decir que no entren en conflicto y que para resolverlo recurramos al valor de verdad de p.

Wittgenstein distingue los usos hipotéticos de los usos asertivos de "yo creo que..." dándole primacía a los primeros. Como aserción, "yo creo que p" expresa o manifiesta una actitud del sujeto respecto a p. Wittgenstein hace alusión a otro uso de "yo creo que ...", que no es ni asertivo (o enfático) ni hipotético, y que se usa con el propósito de restringir o menguar el compromiso del sujeto con respecto a p. En este uso podría entenderse como "tal vez p", que en conjunción con la negación contundente de p produce una afirmación absurda.

En las Investigaciones Filosóficas Wittgenstein nos invita a que imaginemos un juego del lenguaje en el que la expresión "Yo creo que p" sólo se usara para informarnos acerca de lo que la persona cree, esto es, usándose en forma descriptiva. En este caso las dos proposiciones de la afirmación compleja de Moore serían aserciones que producirían una contradicción. Cito:

Supongamos que yo introdujera una expresión -por ejemplo "Yo creo"- de la siguiente manera: debe anteponerse a los informes que sirven para decir algo acerca del que lo comunica. (La expresión, pues, no tiene por qué ir acompañada de ninguna incertidumbre. Tengamos en cuenta que la incertidumbre de la aserción también se puede expresar de modo impersonal; "Puede que hoy venga") -"Yo creo que..., y no es así" sería una contradicción.⁴⁸

Por otra parte, Wittgenstein nos conmina a pensar en un lenguaje en el que la expresión "Yo creo que p" se identificara con la aserción "p", o sea que se

⁴⁸ Wittgenstein, *ibid.*, II-X pag. 191e.

refiriera al contenido en lugar de referirse a la persona. En este caso habría una asimetría entre la persona que lo expresa y los otros, de manera que si los demás quisieran atribuirle a alguien una creencia de manera tentativa, no dirían "El o ella cree" sino "El o ella tiende o se inclina a decir...". La persona misma no usaría en el modo indicativo la expresión "Yo creo que p", pero sí tendría un uso hipotético o subjuntivo: "Supóngase que yo dijera"; "faltaría una inflexión del verbo" en el sentido de que la persona no podría decir "Yo me inclino a decir", sólo podría aseverarla (ibid).

La paradoja de Moore en una situación en la que se leyera en el modo subjuntivo, en el sentido de "supóngase que yo dijera", en donde el sujeto pudiera aseverar "me inclino a decir" y no sólo el uso asertivo, no sería contradictoria, ya que el primer conyunto tendría que leerse hipotéticamente y no como una aserción. Pero entonces no se referiría hipotéticamente a la situación a la que se alude en el otro conyunto: la expresión no querría decir "Supóngase que p". (ibid. pag. 192e)

Finalmente en el uso tentativo de "yo creo que ..." la proposición "Podría ser que estuviera lloviendo, pero no lo está" sería ilegítima -más no contradictoria. Wittgenstein señala que "debemos cuidarnos de decir que "podría ser que estuviera lloviendo" significa: "yo creo que estará lloviendo", en donde lo incorrecto radica en considerar "a una aserción vacilante como una aserción de vacilación."⁴⁹

Entre los pocos comentaradores que se detienen en los pasajes en los que Wittgenstein considera el problema de las creencias se encuentran Linville ⁵⁰ y Rosenthal ⁵¹. Linville toma los comentarios de Wittgenstein en torno a la paradoja

⁴⁹ Wittgenstein, ibid.

⁵⁰ Linville (1979).

⁵¹ Rosenthal, (1988).

de Moore como un ejemplo más de perplejidad filosófica ocasionada por una especie de "ceguera de algún aspecto", o a no "ver como". La ceguera en el caso de la creencia radicaría en no percatarse de las diferencias de uso entre la primera y la tercera persona. Me parece interesante enfocarlo de esa manera. Más aún, me parece que las consideraciones que aquí se señalaron podrían enriquecerse con esa manera de apreciar las tesis de Wittgenstein. Pero señalaré algunas diferencias. Linville piensa que si se admite que un enunciado de la forma de la paradoja de Moore es absurdo -y no lógicamente contradictorio- nos obliga a que el primer conyunto se refiera a la persona, y que sólo en caso de que se refiriera al contenido se daría la contradicción. Cito:

"...in terms of language games, in normal circumstances (wherein asserting M [un ejemplo del tipo de Moore] would be "absurd") one does not conjecture or hypothesize that one believes such-and-such. If one is hypothesizing when one says "I believe it is raining, but it is not raining", then this is a hypothesis about the weather, not oneself. But when inclined to claim that the sentence "I believe it is raining, but it is not raining" asserts something logically possible, one must take the former conjunct to be about the speaker. If regarded as a truth claim about the weather, clearly such a sentence is a contradiction. Since what Moore means to be talking about in connection with such sentences is not logically impossible, what he understands or means by M must be the hypothesis that I believe..., and it isn't so. But the expression "I believe" is not used to formulate the hypothesis that I believe. However, this expression can be seen as a hypothesis; and it is the "dawning" of that aspect, I submit, which inclines one to claim that M is logically possible.³²

Me parece que el punto crucial en Wittgenstein, el cual Linville no enfatiza

³²Linville (1979). p.296.

suficientemente, es que aun cuando se hiciese un giro en el uso de, por ejemplo, "yo creo que está lloviendo" de manera que nos refiriésemos a la persona, de igual forma ese uso no sería independiente de lo que "está sucediendo afuera", por lo que no nos salvaría de la contradicción.⁵³ No caer en contradicción supone un tour de force mental, en el cual imagináramos circunstancias en donde tuviera sentido que yo derivara que tengo una creencia a partir de lo que digo, o que tuviera un yo dividido con lo cual pudiera darle sentido a las proposiciones del tipo de Moore. Tampoco estaría de acuerdo en lo que dice Linville al inicio del párrafo citado, ya que las circunstancias normales, para Wittgenstein, no serían aquellas en las que la afirmación de Moore fuera "absurda".

Considero que para Wittgenstein en un uso normal de los ejemplos de Moore no habría contradicción y que -por lo general- nuestro uso de la primera persona del indicativo, nuestro interés está dirigido a lo que se dice y no a quien lo dice. Por otra parte, nos dice Wittgenstein que para comprender el uso hipotético de "Supóngase que yo crea...", dependemos de una comprensión previa del uso asertivo de la creencia.⁵⁴

Rosenthal, por su parte, defiende en sus escritos una explicación causal de los actos lingüísticos que expresan pensamientos. Cuando alguien dice algo con sinceridad, por lo general el acto lingüístico y el pensamiento correspondiente tienen el mismo contenido proposicional y la misma fuerza.⁵⁵ Sin embargo en relación con el tema que nos ocupa, sostiene Rosenthal, 1) que la manera como transmitimos nuestro pensamiento influye en que el pensamiento y el acto

⁵³ Véase la cita de Wittgenstein (1953) p. 191e antes mencionada.

⁵⁴ Wittgenstein, *ibid.*, p. 192e.

⁵⁵ Cfr. Rosenthal (1988), p. 9.

lingüístico tengan o no la misma fuerza y el mismo contenido proposicional. En el caso de que lo transmitamos empleando la aserción simple, tendrían la misma fuerza y contenido proposicional, mientras que si digo "Yo creo que p", son distintos.⁵⁶ 2) en el primer caso, se estaría expresando la creencia y en el segundo se haría un reporte de que se cree p. 3) ambas formas tienen las mismas condiciones de asertabilidad, empero diferentes condiciones de verdad. Para Rosenthal, estas tesis se ilustran claramente en los ejemplos del tipo de Moore, en los cuales no habría manera de que tuvieran condiciones coherentes de asertabilidad, ya que "el pensamiento que expreso al decir que p es justamente el que niego tener al agregar que no creo que p."⁵⁷ Si no distinguésemos los usos expresivos de los usos descriptivos de nuestros pensamientos, continúa Rosenthal, la paradoja de Moore sería contradictoria y no solamente absurda. Si bien es cierto que Rosenthal recurre a la noción wittgensteiniana de "expresión", sin embargo, los fines para los cuales utilizan dicha noción son distintos de los de Wittgenstein.

La identificación que hace Wittgenstein entre "Yo creo que p" y la aserción de "p", debido a sus tesis respecto al significado de las expresiones y su uso, no le dejaría mucho margen para establecer una distinción tan clara entre condiciones de verdad y condiciones de asertabilidad, y dado que el uso es crucial para la semántica de los conceptos, dice Rosenthal, sería difícil no asimilar los usos expresivos y los descriptivos. (cfr. "Intentionality", p.156) Pero como hemos repetido en las páginas anteriores, Wittgenstein no confunde esos usos, sino hace a los últimos menos comunes de lo que se piensa.

En mi interpretación, Wittgenstein estaría diciendo que en un uso expresivo,

⁵⁶ Ibid. pp. 7-8.

⁵⁷ Ibid, p.9.

las dos aserciones "p" y "yo creo que p" dicen lo mismo, las dos manifiestan el estado mental del sujeto. En cambio, si admitimos el uso descriptivo o informativo, en donde la descripción nos remite al estado mental del sujeto, a que está siendo consciente de lo que cree, de manera similar a la que estaría en el caso de las sensaciones, las emociones etc., y entonces no podría identificarse con la aserción simple de p. En este caso no habría posibilidad de error, que querría decir meramente que para reconocer su estado mental, el sujeto no efectúa ninguna inferencia ni recurre a justificaciones, ya que su relación con su propia interioridad es distinta a la relación que tienen los demás respecto a mi interioridad. En el caso descriptivo, la paradoja de Moore, según dijimos arriba, sería una afirmación contradictoria, mientras que en el uso expresivo no lo sería.

¿Cómo difiere esta interpretación de Wittgenstein de la paradoja con las de las interpretaciones de Williams y de Mellor?

Mellor, con base en la conexión que establece entre BB₁ y la aserción, que bien puede ser la expresión sincera de lo que el sujeto cree que cree, o bien su expresión insincera. Escribe Mellor: "La afirmación sincera es decir lo que uno asiente, aquello que uno cree que cree, y no sólo aquello que uno cree."⁵⁸ Con ello, Mellor procede a mostrar cómo se puede dar cuenta de la paradoja de Moore. Cuando un sujeto dice "Creo que p, pero p es falsa" o "p, pero no lo creo" no puede decirlo sinceramente. El enunciado paradójico podría ser verdadero si el sujeto pudiera decir sinceramente que p, esto es, decir lo que conscientemente no quiere reconocer.

Recordemos que la manera en que Williams diagnostica la paradoja de Moore, alternativamente a la manera en que lo hace Mellor, es diciendo que al

⁵⁸ Mellor (1978), p. 97.

decir el sujeto "p, pero no lo creo" está diciendo algo paradójico porque por un lado se mantiene un grado de compromiso o adhesión por parte del sujeto respecto a lo que expresa (i.e. al creer p, cree que p es verdadera), y por otro lado, niega o cancela lo que dice creer. En otras palabras, para Williams lo absurdo de la afirmación deriva de que se afirma p a la vez que se niega. No hay independencia entre la creencia y el hecho, en la medida en que nuestro interés va dirigido a los contenidos. Si entendemos la afirmación del primer conyunto, como sugerí antes implicando, griceanamente, la afirmación de la verdad de p, no puede evitarse la cancelación del segundo conyunto. Sin embargo, al igual que Wittgenstein, es sólo una de las lecturas de la paradoja de Moore. En las otras lecturas, el matiz de la paradoja, que va desde lo absurdo hasta lo contradictorio, va cambiando hasta llegar a los casos en los que la paradoja puede disolverse.

Comparando las dos maneras -la de Mellor y la de Williams- de dar cuenta de la paradoja de Moore, podemos señalar que en Williams aunque no traza la distinción -ciertamente útil- que hace Mellor entre Bp y BBp, remitiría al igual que Mellor los actos lingüísticos a BBp, ya que se requiere que el sujeto tenga algún grado de conciencia para poder afirmar o negar algo, pero no resuelve la paradoja ya que ésta se reproduce en conductas que chocan entre sí aunque al sujeto se le escapa el conflicto. Mellor, en cambio, señala que el sujeto es insincero ya que niega lo que en realidad cree. De hecho el sujeto no tiene BBp; no está asintiendo a p. El sujeto no está sinceramente asintiendo a la creencia: él erróneamente cree que está asintiendo pero en realidad no lo está. En términos de la paradoja de Moore la falta de sinceridad impide, según Mellor, que se tome como un enunciado contradictorio.

Tal vez una manera en que podríamos entender la paradoja de Moore podría expresarse -como dijimos en conexión con la discusión de Williams- en el caso en

que la primera parte de la aserción "Yo creo que p " fuera un asentimiento desvinculado de lo que el sujeto cree, semejante al caso en el que se le induce la aserción, no la creencia. Y el segundo conyunto haría referencia a los hechos. Esta sería una manera de establecer la incoherencia de embarcarnos en un proyecto de inducción o inyección de creencias. Es como si pudiéramos en términos de la paradoja de Moore el ejemplo de la persona que no quiere admitir que su hijo ha muerto y dijera: "Yo creo que Juan (mi hijo) ha muerto pero es falso que Juan esté muerto", expresando lo que yo creo y que no quiero admitir y el segundo conyunto lo que yo querría que fuera el caso. Al buscar inyectarme la creencia de que no está muerto, modificar el primer conyunto y producir el auto-engaño, reemplazándolo por lo que yo querría creer. La creencia Bp (inconsciente) que tengo y quiero modificar deliberadamente de modo que cuando llegue a BBp se niegue ese hecho.

V. CONCLUSIONES

En este trabajo discutí el problema de las explicaciones de las acciones intencionales y el papel que en ellas desempeñan en particular las creencias. El modelo para las explicaciones intencionales, para ser satisfactorio, debe acomodar adecuadamente la característica de referirse a un tipo de conducta guiada o dirigida hacia un fin. De las varias propuestas presentadas defendí la tesis de que el modelo de las explicaciones intencionales es un modelo de explicación singular en el cual el contexto en el cual surge la indagación, al igual que los sujetos que participan en ella, desempeñan un papel central. Considero también que las explicaciones intencionales son explicaciones funcionales-causales, cuyos *explanans* como *explanandum* son enunciados. Los *relata* de esos enunciados son las razones (creencias, deseos, etc.) y la acción en la que desemboca. La forma de funcionalismo que se defiende es un funcionalismo biológico que considera necesario introducir en sus explicaciones algunos conceptos de la teoría de la evolución.

En la primera parte separé la cuestión ontológica o metafísica respecto a las entidades y relaciones que las vinculan de las cuestiones de la explicación propiamente dicha, sin con ello suponer -como hace van Fraassen- que las explicaciones no tienen implicaciones ontológicas; i.e. sin suponer que la problemática de los modelos de explicación sea neutral frente a los compromisos ontológicos que se asuman. Mi separación obedeció únicamente a que consideré que los enunciados de relaciones causales y los enunciados causales-funcionales de explicación tienen diferentes estructuras y apuntan hacia diferentes requisitos.

En primer término me ocupé de la relación causal aplicada al mundo físico

y centré la discusión en dos aspectos: la cuestión de si los enunciados causales son extensionales o intensionales y el supuesto vínculo semántico que guarda la noción de causa con la noción de necesidad.

En cuanto al problema de la extensionalidad de la relación causal, i.e. si la sustitución de términos correferenciales en los enunciados causales altera o no la verdad del enunciado causal, sostuve que los contextos causales son extensionales y que son relaciones binarias entre sucesos (básicamente la postura de Davidson). La defensa de esta tesis supuso responder a una serie de contraejemplos que han presentado en su contra. Los contraejemplos más comunes los agrupé en aquellos en los que la referencia a los sucesos incluye términos complejos y los que contienen adverbios. La presencia de esos términos o su énfasis en los enunciados causales producía oscilaciones en cuanto a la verdad de sus enunciados, amenazando así la supuesta extensionalidad de los contextos causales y a que -con base en la identidad de sucesos-, las inferencias válidas se pudieran convertir en inválidas.

Mi posición al respecto fue que si tomamos la diferencia entre usos referenciales y usos atributivos y consideramos que en los enunciados causales las descripciones que utilizamos tienen un uso eminentemente referencial, habilitando la tesis de la substitutividad amplia que permite que se inserte cualquier descripción singular siempre y cuando se trate de términos correferenciales. Los contraejemplos que amenazan la extensionalidad de la relación causal parecen sustentarse en el paso de un uso referencial a un uso atributivo que, si no se hiciese, no alteraría el valor de verdad del enunciado.

Si no se aceptara la distinción entre los usos atributivos y referenciales y queremos preservar la extensionalidad de los contextos causales podríamos optar, en las líneas en que sugiere Kim, por hacer más compleja la estructura de los

enunciados causales, i.e. renunciar a considerarla como relación binaria. Kim sugiere introducir a las propiedades, en tanto responsables de que se produzca el efecto, como parte de la estructura del enunciado causal. El argumento de fondo que utilicé para defender la extensionalidad de los enunciados causales, destacando el uso referencial de los sucesos que intervienen en la relación, tiene como consecuencia que los enunciados causales no se consideren ni ambiguos ni incompletos. No se requiere de la especificación la propiedad en virtud de la cual se induce el efecto, sea para desambiguarlo o completarlo, como sugieren autores como Kim.

En cuanto al análisis semántico de la noción de causa, llegué a la conclusión de que dicha noción no tiene que analizarse en términos de la noción de necesidad. Al deslindar la noción de necesidad de la noción de causa, me referí a la necesidad que suele acompañar a la idea de un mundo determinista y que hace de la "causalidad probabilista" un sinsentido. Me remití a la discusión entre Hume y sus antecesores como un ejemplo en el cual se da por sentado que las nociones de causalidad y necesidad están vinculadas y, en ese marco, se daba el enfrentamiento entre dos modelos de causación: el modelo humeano sostenido en leyes y el modelo de causación singular en el cual se dotaba a los objetos de poderes inherentes.

El rechazo del vínculo semántico entre causalidad y necesidad se apoya en la posibilidad de que el mundo fuera indeterminista, dando así cabida a la idea de causalidad probabilista. En un mundo indeterminista, se puede replantear la discusión en cuanto a si la modalidad que caracteriza a la causalidad depende de las leyes -irreductiblemente probabilistas- o si se toma a la causalidad singular como primitiva. Ofrecí las ventajas por las que se le daba primacía a la causalidad singular. En esta postura se deja abierta la posibilidad de que se empleen

contrafácticos en relación con los enunciados causales, pero no como parte de su análisis sino como una glosa informal que se presupone al usar la terminología causal. En cualquier caso, lo que importa es dejar fuera a las leyes como parte del análisis de la causación. En algunas de las posiciones, como la de Lewis, se trata de analizar la noción de causa en términos de contrafácticos y éstos a su vez en términos de "mundos posibles". Para Lewis, las leyes intervienen en la pragmática, más no en el análisis lógico-semántico de la causa. Tanto en la posición de Lewis como en la de Woodward, se toma a la causación como una relación entre sucesos, i.e. como una relación singular, y ésta es la que se adoptó en este trabajo. Sin embargo, preferí la idea de Woodward según la cual un contrafáctico se expresa del siguiente modo: "si no hubiese sucedido p no habría sucedido q ", o que " p hace la diferencia respecto a la ocurrencia de q ". Woodward, a diferencia de Lewis, no analiza la relación causal en términos de contrafácticos sino en términos de presuposición. La ventaja de no tomarla como parte de su análisis es que permite aceptar la idea de que la causación es más básica que la relación contrafáctica.

En la siguiente sección, una vez mostrado de qué manera puede defenderse la problematización de los enunciados causales que vinculan sucesos físicos, pasé al problema de la relación entre razones y acciones. En virtud de que consideré a los contextos causales como extensionales y mi reconocimiento de que, en algún sentido, la relación causal entre razones y acciones es intensional, me llevó a admitir -con Davidson- que cuando hablamos de acciones, en realidad nos hemos trasladado al dominio de las explicaciones. Esta maniobra hace innecesario que se recurra a una noción de causación mental distinta de la causación física.

Las razones son causas de las acciones, pero el vínculo entre ellas no puede establecerse de manera directa, como de hecho se establece en el caso de la

relación entre sucesos físicos. En el contexto mental el vínculo causal depende no sólo de los sucesos a los cuales nos referimos sino también de cómo describimos tanto las razones como las acciones. Las razones en tanto tales, si se expresan en términos de creencias, deseos e intenciones tienen que describirse de ciertos modos, lo mismo que las acciones que se desprenden de ellas. Si además admitimos que las creencias, deseos e intenciones se consideran como relaciones entre ciertas actitudes y sus contenidos, éstos pueden considerarse como elementos que introducen proposiciones u oraciones. Pero una de las tesis centrales es que podemos hablar de estados mentales que se identifican en virtud de su contenido, pero que este contenido no es lingüístico o semántico. La manera usual de entender el contenido es en términos de proposiciones, i.e. con descripciones muy diferenciadas de los sucesos, es porque se está tomando como paradigmático el modelo de los seres que han desarrollado capacidades lingüísticas. Necesitaríamos a la vez referirnos a la propiedad bajo cierto modo de presentación, como sugiere Peacocke en la línea de Frege.

Vimos que la cuestión más difícil de resolver surge en cuanto a cómo acomodar en un modelo causalista mental la idea de contenido, de modo que el contenido tenga una injerencia causal. Defendí la postura en la cual se aceptan niveles o estratos de propiedades, en donde se establece una relación de superveniencia de las propiedades específicas de cada nivel respecto de las propiedades físicas del sistema. En tanto propiedades supervenientes (que podrían ser una forma de relación mereológica), podrían adquirir una conformación o manera de operar independientemente del nivel subveniente. Este modelo que proponen autores como Kim, pretende ofrecernos una manera de combinar cierta autonomía entre las relaciones causales que se dan en los diferentes niveles respecto de los inferiores, sin perder la idea de que la existencia de los niveles

superiores depende de los niveles inferiores. Este modelo de estratos es una alternativa poderosa en la cual, a pesar de las dependencias entre los estratos, se evita caer en un modelo reductivista. Mis reservas con este modelo se condensan en dos cuestiones: que sus defensores la proponen como un modelo en el cual las leyes siguen siendo determinantes. Las leyes establecerían relaciones entre tipos de propiedades. Algunos autores (Sober y Segal) proponen modificar el requisito de Kim haciendo que la relación se dé entre instancias de propiedades y no entre tipos de propiedades. Sin embargo, considero que el modelo sigue siendo inadecuado en la medida en que sigue asimilando los patrones normativos a los patrones legales o nómicos, en donde las instancias siguen viéndose únicamente como instancias de leyes. Otra de las tesis centrales del trabajo es que las explicaciones teleológicas o dirigidas hacia un fin y su direccionalidad hacia ese fin es normativa en un sentido en el que las leyes no lo son.

En el segundo capítulo, me concentré básicamente en cuatro modelos de explicación, que a través de sus diferencias y sus diferentes acentos, me permitiera elegir el modelo más apto para dar cuenta de la intencionalidad de nuestras acciones. Me referí a las explicaciones disposicionales, singulares, racionales y biológico-funcionales. Ubiqué la diferencia entre ellas en cuanto al tipo de postulación que hacen. En los dos primeros modelos, al igual que en el caso de la causación, la discusión se centra en el papel que desempeñan las leyes y se postulan entidades o propiedades. En los dos últimos modelos se hace referencia a ciertos patrones de conducta -una manera de sistematizar y establecer normas para ciertas conductas.

Hempel afirma que las explicaciones disposicionales son un tipo de explicación teórica, de modo que no tendría dificultad en acomodarlas en su modelo nomológico deductivo. Considera también que el modelo disposicional es

el modelo que tradicionalmente se asocia a las explicaciones intencionales: es el modelo que explica las decisiones y las acciones en términos de creencias, propósitos, rasgos de carácter etc. La introducción de los términos disposicionales sigue la misma ruta que la de los términos teóricos en una teoría. Una diferencia crucial entre las disposiciones físicas y las disposiciones a la conducta es que las últimas se pueden manifestar de maneras muy diversas, según las circunstancias. Ya Ryle apuntaba a esa diferencia, distinguiendo entre disposiciones que tienen una manera de manifestarse y las que tienen múltiples maneras ("*multi-track*"). Tienen las características que Braithwaite denominó "variabilidad" y "plasticidad". Cada una de las diversas maneras en que se puede desplegar una disposición, según Hempel, constituye o una condición necesaria o suficiente para la presencia de la disposición general y a su vez implica otras propiedades de los objetos teóricamente caracterizadas. De este modo las explicaciones disposicionales pueden verse como una forma de explicación teórica y las disposiciones como términos que se constituyen a la manera de los términos teóricos de una teoría científica.

Tanto los modelos disposicionales de la explicación como ciertos modelos funcionalistas comparten con el modelo nomológico deductivo el hecho de se consideran que las leyes tienen un peso explicativo central. Considero que todos estos modelos explicativos, en tanto que concentran la fuerza explicativa en las leyes, no son correctos. Asumo que hay otros valores explicativos que cobran importancia una vez que se da el viraje hacia la primacía las explicaciones causales singulares y los contextos particulares en los que se invocan. Ya no es sólo cuestión de si las explicaciones intencionales pueden acomodarse en el modelo nomológico deductivo hempeliano, sino lo inadecuado de ese modelo en tanto modelo explicativo.

Destaqué la falta de unanimidad entre los filósofos de la mente respecto al

giro que debe tomar la discusión de sus explicaciones. La apelación a las leyes ha sido central para entender lo que sucede en el mundo y también como seña distintiva de algunos modelos de explicación. Y hay quienes no renuncian a la idea de que la plataforma central de las leyes es adecuada y debaten con otros autores en cuanto a si las generalizaciones psicológicas, por el hecho de no ser estrictas, pueden legítimamente llamarse leyes, o si deben considerarse como un tipo de leyes probabilistas. Quienes están convencidos de que el modelo de leyes probabilistas es el indicado para las explicaciones intencionales, recurren a la teoría de las decisiones. Para otros, como Davidson, no basta con que haya generalizaciones sino que tiene que haber leyes estrictas para que las explicaciones psicológicas tengan la misma estructura que las explicaciones físicas. Discutí y traté de defender la idea de que las explicaciones de las acciones intencionales, *p*ace Hempel, sigue sin poder asimilarse al modelo nomológico de explicación. Me parece inadecuado el modelo nomológico para las explicaciones intencionales, más allá de la cuestión de si hay o no leyes psicológicas o psico-físicas. Aunque las hubiese, no serían pertinentes para las explicaciones ya que, como dice Woodward, no forman parte de las condiciones de verdad del enunciado explicativo ni de su estructura.

A mi modo de ver, la discusión en torno al estatuto de las generalizaciones psicológicas o psico-físicas, cobra sentido sólo si antes se ha admitido que las leyes son el criterio preponderante. En el trabajo argumento que no lo es y que intervienen otros valores explicativos. La amplitud de valores explicativos deriva de la discusión de las explicaciones, viéndolos primeramente como actos ilocucionarios que ubican la discusión en el contexto de la pragmática, cuyo tratamiento satisfactorio hace que emerjan de manera central las consideraciones contextuales.

Una razón, como sostiene Davidson, es que las explicaciones nos remiten a las características, disposiciones, rasgos de carácter, etc. de un sujeto particular, que no tiene por qué ser referidas -para ser explicativas- a las maneras normales en que sujetos con características semejantes a las del sujeto en particular, en circunstancias similares, se desempeñan. Ha sorprendido a más de un crítico, que el mismo autor que hace afirmaciones como éstas, a la vez se mantiene fiel al modelo nomológico explicativo para lo físico. La adhesión se atribuye a que Davidson mantiene con la causalidad física como la única forma de causalidad, que a su vez es producto de su tesis de la identidad entre sucesos físicos y sucesos mentales. No hay que perder de vista que al referirnos a la noción de causalidad, nos volvemos a enfrentar al problema de las leyes, en la medida en que para muchos autores se diferencian las relaciones causales de las meras sucesiones temporales en el hecho de que las relaciones causales se apoyan en leyes.

Como sugerimos antes, algunos autores intentan aliviar esta tensión que existe en el pensamiento de Davidson proponiendo un tipo de causalidad particular que conviniera más a los fenómenos intencionales. Pero otros autores, intentan aliviar una tensión semejante a la que se aprecia en el pensamiento de Davidson, proponiendo un modelo de explicación alternativo. En las propuestas más extremas, en ese modelo de explicación racional o intencional tiende a desaparecer la referencia al vínculo causal. En ambas maneras de aliviar la tensión el denominador común es la insistencia de referirnos a patrones normativos, que variarán en cuanto al sitio en el que hacen intervenir esos principios normativos.

En el caso de las explicaciones intencionales partí de la forma cotidiana de entender las acciones intencionales (denominada en la literatura "psicología popular") con todo y su modelo causal. Considero que hablar de causalidad en

relación con la explicación es sustancialmente correcto y no sólo porque en nuestras prácticas sociales y de comunicación así lo hacemos sino porque hay una manera natural de entender que las razones fungen como causas de las acciones. En la postura de Davidson, postura clave en torno a la cual giran muchas de las discusiones contemporáneas en filosofía del lenguaje y de la mente, se considera que si bien las explicaciones intencionales como tales no se apoyan en leyes, dada su tesis de la identidad entre sucesos físicos y sucesos mentales, sí lo harían sus redescpciones en sus correspondientes explicaciones físicas. Así, la causalidad que aparece en las explicaciones causales intencionales no serían causales en forma propia sino de manera derivada, por su dependencia de las relaciones causales físicas. Al discutir la postura de Davidson y la evaluación de las muchas críticas que se le dirigen a su postura, defendí la tesis de que el modelo de explicación intencional es un modelo que busca explicar el hecho de que se hayan dado ciertas acciones particulares.

Sugerí, con autores como van Fraassen, Achinstein y otros, que el debate debe pensarse en términos de una multiplicidad de virtudes explicativas y en el contexto de la pragmática de las explicaciones. El señalamiento de estas virtudes apuntan hacia la incorporación de muchos elementos contextuales y pragmáticos ineliminables de las explicaciones, como lo son los de relevancia, accesibilidad y disponibilidad de información, predicción, vocabulario pertinente, precisión, verosimilitud etc. Se defiende también que la mejor manera de dar cuenta de estos elementos contextuales es dándoles prioridad explicativa a los casos particulares, en los que en su carácter de actos ilocucionarios, los interlocutores cobran una importancia capital. Lo satisfactorio o insatisfactorio de una explicación no puede evaluarse al margen de los propósitos que se persiguen y de las personas que las persiguen. Si las explicaciones -en general- son sensibles a los aspectos

contextuales, de modo que no sólo se modifican de acuerdo a las variables particulares que se dan en cada individuo, también su valor fluctuará en función de muchas otros factores contextuales que tienen un impacto en la situación del sujeto particular. Esta perspectiva es constitutiva de las explicaciones y no una dimensión más que opcionalmente pueda agregarse. No es separable la pragmática de las explicaciones de la búsqueda de las teorías que mejor las modelen. Por ese mismo motivo las explicaciones son irremediablemente parciales e incompletas.

Aludí antes a la idea de normatividad para sustanciar el contraste entre las explicaciones físicas y las explicaciones intencionales. Recordemos que la primera manera de establecer el contraste fue a partir de las leyes que asisten a las explicaciones físicas. Con la normatividad se pretende hacer referencia a una característica constitutiva de las explicaciones intencionales y que no aparece en las explicaciones físicas. La forma más usual de normatividad es la que se encuentra a la base de la racionalidad.

La idea de normatividad que opera en la idea misma de la racionalidad es la que está gobernada por principios lógicos y de inferencia. En un patrón explicativo se entrelazan diferentes estados y se da por supuesto que de manera cooperativa y concertada producen una acción. De este modo, intervienen los diferentes estados como los deseos, las intenciones, las creencias y las emociones. Sin embargo, la manera en que se produce esa cooperación y concertación se hace en virtud de los contenidos que identifican a los diferentes estados. Además de esta forma de normatividad se habla de una normatividad que ata internamente a ciertos actos con acciones particulares, como en el caso de los performativos. Finalmente, en algunas interpretaciones de las teorías de la evolución se habla de una normatividad asociada a ciertas funciones biológicas, según la cual la función fija y es fijada por su contenido: Estas funciones establecen los parámetros de lo que

se consideraría como un funcionamiento normal de algún dispositivo, con referencia al cual se miden las anomalías en su funcionamiento.

Para el caso de las explicaciones de las acciones intencionales defendí, en el marco del modelo biológico-funcional, dos formas de normatividad para el caso de los humanos: el que se asocia a la caracterización misma de la función biológica -que se da también en los organismos más simples- y la normatividad asociada a la racionalidad. Este modelo biológico-funcional permite que se expliquen ciertos patrones de conducta natural en términos teleológicos que trasciende la esfera humana de una manera no metafórica.

En varios puntos del trabajo me detengo en la postura de Dennett en la cual de manera esquemática se recogen muchas de las tesis que aquí trato de sustentar. Muchas de ellas me llevan a discutir el problema de lo mental en autores en los que una teoría de este tipo se encuentra más desarrollada, como por ejemplo en Millikan y McGinn. Expongo con un cierto detalle la postura de Dennett. De sus tesis, encuentro rescatable la manera en que habla de diferentes tipos de explicaciones, recogidas en lo que denomina diferentes "perspectivas" o "estrategias", que se introduce como una propuesta intermedia entre fisicalistas e intencionalistas, compartiendo en ello la postura de Davidson, emplea para las diferentes explicaciones diferentes perspectivas. Dennett mismo ha empleado el término "instrumentalismo" para describir su postura, pero a la luz de las críticas y malentendidos que ha suscitado el empleo del término, lo ha llevado a diferenciarla de otras teorías que llevan el mismo nombre. La postura de Dennett también nos ofrece una manera de ligar los problemas de la explicación sin desligarlo totalmente de los problemas ontológicos y, por el otro, la introducción de las consideraciones evolutivas en las discusiones de filosofía de la mente. La primera lo lleva a concebir a los estados de deseo y de creencia como

postulaciones teóricas a la manera como se postulan entidades en la práctica científica. Me parece que Dennett, al hablar de *abstracta* en relación a las explicaciones intencionales, tiende a identificarlas como un tipo de explicación disposicional y por ende teórica, aún cuando él insista que no se trata de la postulación de "entidades" hipotéticas e inobservables y, extensivamente, tampoco se trataría de la postulación de algún tipo de propiedad particular, sea categórica o disposicional. Me parece que la referencia a los *abstracta* oscurece y desvirtúa una intuición particular que tiene Dennett y que me parece correcta: la importancia de la idea de patrones de conducta y la manera en que éstos introducen la teoría de la evolución. En ese sentido es más apto hablar de funciones más que de disposiciones y con ciertos contenidos implícitos en su especificación. De esta manera también se establecería la manera de identificar los contenidos de manera externa a lo que sucede en el fuero interno del sistema u organismo. Traté de dar respuesta a objeciones como las de Fodor y McDowell en las cuales se impugna que los contenidos -y por tanto las funciones de las que forman parte- existan de manera genuina en los organismos o sistemas a los cuales se atribuyen. Los contenidos son parte de la manera en que nosotros -los interpretadores- resolvemos dar cuenta de la conducta.

En las dos partes finales del trabajo me concentro en la discusión del estado mental de la creencia y la manera en que da cuenta de él en el modelo biológico-funcional favorecido. En este modelo pueden señalarse niveles de creencias, que se dan como resultado de nuestra evolución humana. Con la formación de niveles superiores de creencias aparece la otra función cognoscitiva de las creencias, que en colaboración con las otras creencias conforman un patrón racional. Además del carácter explicativo de las creencias, señalo sus otras características. Entre ellas, la característica del asentimiento -tanto interno como verbal y los diferentes

proyectos e intereses en los que interviene.

Con la introducción de los niveles de creencia me pregunto por la manera en que se vinculan estos niveles. En Dennett los niveles de creencia se desdoblan como diferentes niveles de explicación psicológica y sus correspondientes postulaciones teóricas: el nivel sub-personal y el nivel personal. En el nivel sub-personal se ubicarían los estados disposicionales que se manifiestan directamente en el comportamiento del sistema -fuese un cierto tipo de animal, bebé humano e inclusive, si así se quisiera, cierto tipo de máquinas. En este nivel no se requiere que los sistemas sean capaces de formarse "representaciones internas" pero de alguna manera tiene que aceptarse que esos sistemas son capaces de captar de alguna manera aspectos que se encuentran en el entorno. Para el nivel personal, en cambio sí se necesitarían estas capacidades más complejas. En la teoría sub-personal Dennett introduce a las creencias entendidas como *illata* u "opiniones". Las opiniones son estados cognitivos, en los cuales el sujeto "apuesta a la verdad de una oración en particular", o asiente a ella. Una característica de las opiniones es que son estados y no, como uno pudiera estar inclinado a pensar, ocurrencias concretas presentes a la mente, o asentimientos conscientes. Así como sostengo que la relación entre los estados mentales y las acciones no puede recogerse satisfactoriamente en términos de disposiciones, sin embargo considero que los diferentes niveles sí podrían vincularse disposicionalmente.

Para finalizar, así como considero la manera en que las diferentes creencias se integran para formar patrones racionales y coherentes, también me detengo en la manera en que pueden producirse ciertos patrones que calificamos de irracionales. Para explicarlos solemos valernos de los diferentes niveles de creencias y la manera en que se manifiestan en diferentes aspectos de nuestra conducta verbal y no verbal. Me centré en particular la forma de incoherencia que

produce el auto-engaño. Me refiero también a otra forma de incoherencia, una incoherencia verbal, la llamada paradoja de Moore. En este último, valiéndome de la discusión que lleva a cabo Wittgenstein en torno a la creencia y las diferencias que se producen según su expresión en la primera o tercera persona y las circunstancias que la rodean, ilustro las diferentes maneras en que se puede entender el enunciado.

Mi discusión de la creencia intenta mostrar que en tanto actitud cognitiva debe estudiarse tanto en su aspecto psicológico como en su aspecto pragmático ligado a las acciones y que algunos de los errores en la tradición derivan de haber estudiado sobre todo su aspecto subjetivo, sin percatarse de que la subjetividad se construye más que ser el dato epistemológico inmediato del cual se parta.

BIBLIOGRAFIA

- Achinstein, P., The Nature of Explanation, Oxford University Press, 1983, La naturaleza de la explicación, Traducción Leticia García Urriza, Fondo de Cultura Económica, México, 1989.
- Armstrong, Belief, Truth and Knowledge, Cambridge University Press, 1973.
- Baker, R.A., "Metaphysical and Mental Causation", en Mental Causation, Heil & Mele, 1993.
- Beauchamp T.L. & Rosenberg, A., Hume and the Problem of Causation, Oxford University Press, 1981.
- Churchland, P.M.: "The Logical Character of Action Explanations", Philosophical Review 79, 1970.
- Cummins, R., "What Can We Learn from Brainstorms", Mind, Brain and Functionalism, J. Biro & J. Shahan (eds), University of Oklahoma Press, 1982.
- Davidson, D., "Actions, Reasons and Causes", Essays on Actions and Events, Clarendon Press, Oxford, 1980a.
- _____, "Intending" en Essays on Actions and Events, Clarendon Press, Oxford 1980b.
- _____, "Hempel on Explaining Actions", Essays on Actions and Events, Clarendon Press, Oxford, 1980c.
- _____, "Deception and Division", en J. Elster (ed.), The Multiple Self, en castellano: Quinto Simposio Internacional, UNAM, 1991.
- _____, "Problems in the Explanation of Action", en Metaphysics and Morality: Essays in Honour of Smart, Pettit, Ph., Sylvan R., Norman, J. (eds), Blackwell, 1987.

- _____ "Thinking Causes", en Mental Causation, Heil and Mele (eds), Clarendon Press, Oxford, 1993.
- Dennett, D., "Artificial intelligence as Philosophy and as Psychology" en Brainstorms, Harvester, Hassocks, 1978a.
- _____, "Why the Law of Effect will not go away" en Brainstorms, 1978b.
- _____, "How to change your mind", en Brainstorms, 1978c.
- _____, "True Believers", en The Intentional Stance, MIT, 1987a.
- De Sousa, R., "How to Give a Piece of your Mind: or, The Logic of Belief and Assent", Review of Metaphysics 25, 1971.
- Donnellan, K., "Reference and Definite Descriptions", Philosophical Review 75, 1966.
- Dretske, F., "Referring to Events", Midwest Studies in Philosophy, 2, 1977.
- Elster, J., The Multiple Self, Cambridge University Press, 1985.
- Evans, G., "Following a Rule: Objectivity and Meaning", en Wittgenstein: To Follow a Rule, Holtzman and Leich (eds), Routledge and Kegan Paul, 1981.
- Fodor, J.A., A Theory of Content and Other Essays, M.I.T. Press, 1990.
- Fingarette, H., Self-Deception, Routledge and Kegan Paul, Londres, 1969.
- Goldman A., "The Individuation of Action", Journal of Philosophy 68, 1971.
- Guariglia, O., "La definición y la explicación causal según Aristóteles", Revista Latinoamericana de Filosofía, vol. VIII, No 2, Buenos Aires, Argentina, 1982.
- Hempel, C., "Aspects of Scientific Explanation" en Aspects of Scientific Explanation, New York, The Free Press 1965.
- _____ "Dispositional Explanation", en Dispositions, R. Toumela (ed.), Reidel

Publishing Co., Holland, 1978.

Hookway, C., "Conscious Belief and Deliberation" en Supplement of the Aristotelian Society, LV, 1981.

_____, Quine, Stanford University Press, 1988.

Hume, D., A Treatise of Human Nature, being an attempt to introduce the experimental method of reasoning into moral subjects, L.A Selby-Bigge (ed), Clarendon Press, Oxford, Segunda Edición, 1978. Tratado de la naturaleza Humana Traducción al castellano Félix Duque (ed), Editora Nacional, Madrid, 1977.

Kim, J., "Causation, Nomic Subsumption and the Concept of an Event", Journal of Philosophy 70, 1973.

_____, "Causation, Emphasis and Events", Midwest Studies in Philosophy, 2, 1977.

_____, "The Non-Reductivist's Troubles with Mental Causation", en Mental Causation, Heil and Mele (eds), Clarendon Press, Oxford, 1993.

_____, "Can Supervenience Save Anomalous Monism?", en Mental Causation, Heil and Mele (eds), Clarendon Press, Oxford, 1993.

Kripke, S., "Speaker's Reference and Semantic Reference", Midwest Studies in Philosophy. Vol. II, French, Uehling and Wettstein (eds), 1977.

Lewis, D., "Causal Explanation", en Explanation, D. Ruben (ed), Oxford University Press, 1993.

Linville, K., "Wittgenstein on Moore's Paradox", en Wittgenstein: Sources and Perspectives, C. G. Luckhardt (ed), Cornell University Press, Ithaca, New York, 1979.

Mackie, J.L., Truth, Probability and Paradox, Oxford, Clarendon Press, 1973.

McDowell, J., "The Content of Perceptual Experience", The Philosophical Quarterly, Vol. 44, No. 175, Abril, 1994.

- McGinn, C., "Action and its Explanation" en Philosophical Problems in Psychology. N. Bolton (ed), Methuen, London, 1979, trad. castellana Problemas filosóficos en Psicología, Alhambra, España, 1982.
- _____, Wittgenstein on Meaning, Aristotelian Society Series, vol I, Blackwell, 1984.
- McLaughlin, B., "Davidson's Response to Epifenomenalism", en Mental Causation, Heil and Mele (eds), Clarendon Press, Oxford, 1993.
- Marcus, R., "Extensionality", reimpresso en Reference and Modality, Linsky, L., (ed.), Oxford University Press, 1971.
- Mellor, H., "In Defense of Dispositions", Philosophical Review, 1974.
- _____, "Conscious Belief", Proceedings of the Aristotelian Society, Vol. LXXVIII, 1978.
- Mosterfn, J., Racionalidad y acción humana, Alianza Universidad, Madrid, 1978.
- Moya, C., The Philosophy of Action: An Introduction, Polity Press, Cambridge, England, 1990.
- Platts, M., Ways of Meaning, Routledge and Kegan Paul, 1979, Sendas del significado, Traducción de Cecilia Hidalgo y Eduardo A. Rabossi, Fondo de Cultura Económica/Universidad Nacional Autónoma de México, 1992.
- Prior, E., Pargetter, Jackson, F., "Three Theses about Dispositions", American Philosophical Quarterly, Vol. 19, N° 3, Julio 1982.
- Prior, E., Dispositions, Aberdeen University, 1985.
- Quine, V.O., Word and Object, M.I.T., 1960.
- _____, "Reference and Modality", en From a Logical Point of View, Harper & Row Publishers, N.Y., 1961.
- _____, Pursuit of Truth, Harvard University Press, 1990.

- Railton, P., "Probability, Explanation and Information", en Explanation, D. Ruben (ed), Oxford University Press, 1993.
- Ramsey, F.P., "General Propositions and Causality", en Ramsey The Foundations of Mathematics, Routledge and Kegan Paul, 1931.
- Reichenbach, H., Experience and Prediction, The University of Chicago Press, 1938, Phoenix Books, 1961.
- Rosenthal, D., "Intentionality", en Midwest Studies in Philosophy, vol.X, editado por P.A. French, T.E. Uehling, Jr. y H. K. Wettstein, University of Minnesota Press, 1986.
- _____, "Thinking that one thinks" (manuscrito) 1988.
- Russell, B., "On the Notion of Cause", en Mysticism and Logic, Unwin Books, 1963.
- Ryle, G., The Concept of Mind, Penguin Book, 1949.
- Saab, S., "Una posición intermedia entre el fisicalismo y el intencionalismo: Dennett", Critica, Vol. XXVI, No. 76-77, México, Abril-Agosto, 1994.
- Sartre, J.P., El ser y la nada, París, 1943. Trad. Juan Valmar, Losada, Buenos Aires, 1989.
- Segal, G. & Sober, E., "The Causal Efficacy of Content", Philosophical Studies 1990.
- Salmon, W.C., Four Decades of Scientific Explanation, University of Minnesota Press, 1990a.
- _____, "Scientific Explanation: Causation and Unification", Critica, Vol. XXII, No. 66, México, Diciembre 1990b.
- Stich, S., "Belief and Subdoxastic States", Philosophy of Science, 45, 1978.
- _____, From Folk Psychology to Cognitive Science, M.I.T. Press, 1983.

- Strawson, P., "Review of Philosophical Investigations, en Wittgenstein, ed. G. Pitcher, MacMillan & Co. Ltd, London, 1968.
- Stroud, B., Hume, Routledge and Kegan Paul, London, 1977. Traducción al castellano de Antonio Ziri6n, Universidad Nacional Aut6noma de M6xico, 1986.
- _____, "Hume and the idea of causal necessity", Philosophical Studies 33, 1978.
- Villoro, L., "Hobbes y el modelo de convenio utilitario", Dianoia, XXXIX, Instituto de Investigaciones Filos6ficas/Fondo de Cultura Econ6mica, M6xico, 1993.
- Williams, B., "Deciding to Believe", en Problems of the Self, Cambridge University Press, 1973.
- Woodward, J., "A Theory of Singular Causal Explanation", Erkenntnis, 21, 1984.
- _____, "Are Singular Causal Explanations Implicit Covering-Law Explanations?", Canadian Journal of Philosophy, 16, 1986.
- Wittgenstein, L., Philosophical Investigations, Blackwell, 1953. Investigaciones Filos6ficas, Coedici6n Instituto de Investigaciones Filos6ficas de la Universidad Aut6noma de M6xico y la Editorial Cr6tica de Barcelona, M6xico-Barcelona, 1988.
- _____, Last Writings on The Philosophy of Psychology, vol.I, C.H. Von Wright & H. Nyman (eds), Traducci6n de Luckhardt & Maximilian, Blackwell, 1982.
- _____, Zettel, E. Anscombe & G.H. Von Wright (eds), Berkeley, University of California, 1970. Traducci6n, Instituto de Investigaciones Filos6ficas, Universidad Nacional Aut6noma de M6xico, M6xico, 1985.