

Gustavo Arroyo

Teoría de las definiciones
Una introducción crítica

EDICIONES **UNGS**



Universidad
Nacional de
General
Sarmiento

Arroyo, Gustavo

Teoría de las definiciones : una introducción crítica / Gustavo Arroyo. - 1a ed. -

Los Polvorines : Universidad Nacional de General Sarmiento, 2019.

192 p. ; 21 x 15 cm. - (Humanidades ; 41)

ISBN 978-987-630-438-2

1. Lógica. 2. Argumentación. I. Título.

CDD 160

EDICIONES **UNGS**

© Universidad Nacional de General Sarmiento, 2019

J. M. Gutiérrez 1150, Los Polvorines (B1613GSX)

Prov. de Buenos Aires, Argentina

Tel.: (54 11) 4469-7507

ediciones@campus.ungs.edu.ar

ediciones.ungs.edu.ar

Diseño gráfico de la colección: Andrés Espinosa

Diseño de tapa: Daniel Vidable

Imagen de tapa: Piet Mondrian. *Lozenge composition with yellow, black, blue, red and gray*, 1921. The Art Institute of Chicago. CC0 public domain designation.

Diagramación: Eleonora Silva

Corrección: Gustavo Castaño

Hecho el depósito que marca la Ley 11723

Prohibida su reproducción total o parcial

Derechos reservados

Impreso en Ediciones América

Abraham J. Luppi 1451, CABA, Argentina

en el mes de octubre.

Tirada: 200 ejemplares.



Libro
Universitario
Argentino

Índice

Prólogo.....	9
Introducción	11
Conceptos indefinibles	41
Sócrates y las definiciones	53
Análisis de un caso: la definición de <i>mentira</i>	73
Imaginación y categorización.....	85
Semejanzas de familia	97
El concepto de arte.....	113
El concepto de religión	127
Retorno al esencialismo: la definición de <i>juego</i> de Bernard Suits	147
Epílogo: filósofos como abogados intuitivos	167
Preguntas y ejercicios.....	183
Bibliografía.....	187

Prólogo

Cuando los manuales de lógica abordan el tema de las definiciones (si es que acaso lo hacen), rara vez van más allá del nivel más elemental de análisis. En los pocos textos que encontramos sobre el tema (y en los poquísimos que encontramos en lengua castellana) aparecen repetidas una y otra vez, de forma acrítica, las mismas cosas. Este libro responde al deseo de llevar la discusión un paso más allá.

Realizar esa tarea supone no solo investigar algunas cuestiones generales poco estudiadas, como puede ser el presupuesto de que las definiciones son proposiciones necesarias o el de que una buena definición debe enunciar condiciones suficientes y necesarias, sino además intentar conectar el problema de la definición con temas que interesen a los alumnos de filosofía, a quienes los textos de lógica, normalmente, están dirigidos. El segundo objetivo ha sido algo más fácil de alcanzar porque definir conceptos es algo que los filósofos han hecho desde siempre. Es incluso una de las tareas propias de la filosofía. Ciertos textos canónicos, como, por ejemplo, algunos diálogos platónicos, consisten esencialmente en buscar definiciones para determinados conceptos. El método que he seguido consiste en discutir primero los problemas definicionales desde una perspectiva general e ilustrarlos luego con ejemplos de definiciones de conceptos que interesan a la filosofía.

Dado que el libro podrá servir de manual en un curso introductorio de lógica, he incorporado un apéndice con ejercicios. Estos podrán ser realizados por los alumnos en el aula o por el lector individual que desee mejorar la comprensión de los temas que en él se discuten. A lo largo del libro, aparecen marcas en el texto con números romanos que remiten a las consignas correspondientes de la sección “Preguntas y ejercicios”, que se encuentra al final del libro.

Soy consciente de que muchos puntos merecían un análisis más pormenorizado del que he podido ofrecer; esta falencia se debe no solo a mis limitaciones personales, sino también al hecho de que muchos de los temas y problemas

tratados en el libro tienen pocos antecedentes en la literatura y tuvieron que ser abordados, por así decir, desde cero. Si mis planteos carecen de profundidad tendrán al menos el mérito de señalar posibles caminos para quienes deseen continuar investigando el tema.

He compartido resultados parciales de mi investigación en diversos grupos durante los últimos tres o cuatro años. En particular en mis clases de Lógica y Argumentación en la Universidad Nacional de General Sarmiento. Pero también en cursos que dicté en la Universidad Nacional de Rosario, en la Universidad Nacional a Distancia de Madrid, en la Universidad Popular del Estado de Puebla, en la Universidad Autónoma de la Ciudad de México y en la Universidad Católica Argentina. Muchas de las ideas de este libro han surgido o se han nutrido de los comentarios y las críticas que he recibido de los asistentes a esos cursos. A todos ellos, mi agradecimiento.

Los Polvorines, noviembre de 2018

Introducción

Clases y definiciones

Objetos que tienen uno o más rasgos en común conforman una clase. Dado que todo objeto tiene algún rasgo en común con algún otro, cualquier par de objetos conforman una clase. Algunas de esas clases son reconocidas socialmente. Por ejemplo, la clase de los árboles, la de los metales, la de los peces. En general, el simple hecho de que haya un nombre para designar una clase indica que ella es una a la que se le ha dado reconocimiento. Pero ¿cómo se llama el residuo húmedo que deja la nariz de un animal al tocar una ventana? No hay un nombre para eso. Esa clase no es socialmente reconocida. En un gesto de humor, el filósofo Gregory Murphy (también conocido por sus “leyes de Murphy”) bautizó ese residuo como *pupkus*.

¿Qué condiciones deben darse para que una clase sea socialmente reconocida? La primera condición es que existan objetos dentro de esa clase. Es decir, que la clase no esté vacía. Recientemente se adoptó el término *bullying* para designar el acoso de un niño o de un grupo de niños hacia otro en el contexto escolar, aunque antes ya existía ese tipo de conducta, pero tal vez no en una proporción relevante como para ser socialmente reconocida. La segunda condición es que la existencia de los objetos que componen la clase sea conocida. Esa es la razón por la cual, por ejemplo, solo desde hace poco tiempo tenemos un término para designar el uranio. La tercera condición es que la clase tenga relevancia práctica. Con la clase de los metales que denominamos “uranio” se puede generar energía y producir bombas. El *bullying* es un tipo de conducta que causa mucho sufrimiento a las víctimas y puede dejar secuelas psicológicas. El *pupkus*, en cambio, no tiene ninguna consecuencia relevante en nuestras vidas.

En “El idioma analítico de John Wilkins”, Borges nos habla de la clasificación natural propuesta por John Wilkins en el siglo xvii. La clase “piedra”, en su taxonomía, contempla las siguientes subclases: comunes (pedernal, cascajo, pizarra), mógicas (mármol, ámbar, coral), preciosas (perla, ópalo), transparentes (amatista, zafiro) e insolubles (hulla, greda y arsénico). Casi tan alarmante como la octava es la novena categoría. Esta nos revela que los metales pueden ser imperfectos (bermellón, azogue), artificiales (bronce, latón), recrementicios (limaduras, herrumbre) y naturales (oro, estaño, cobre). Más extraña aún es una clasificación de los animales que, en el mismo texto, Borges atribuye a una antigua enciclopedia china. En ella, los animales son clasificados en 14 clases:

- Pertenecientes al emperador
- Embalsamados
- Amaestrados
- Lechones
- Sirenas
- Fabulosos
- Perros sueltos
- Incluidos en esta clasificación
- Que se agitan como locos
- Innumerables
- Dibujados con un pincel finísimo de pelo de camello
- Etcétera
- Que acaban de romper el jarrón
- Que de lejos parecen moscas

En la sociedad imaginaria en la que está escrita esta enciclopedia, clases como la de los gatos o la de los pájaros no reciben ningún reconocimiento. Esto no quiere decir que allí no se pueda hablar sobre gatos o pájaros, pero solo en calidad de miembros de otras clases, por ejemplo, la de los animales que de lejos parecen moscas o la de los animales que acaban de romper un jarrón. En esta sociedad, gatos y pájaros son designados con el mismo término y pertenecen a la misma clase.

Pero ¿no es incorrecta esa clasificación? ¿No es incorrecto agrupar a los gatos y a los pájaros en la misma clase y no establecer ninguna diferencia entre ellos? ¿No resulta claro que los gatos y los pájaros no pertenecen a la misma clase de animales? Quien se incline por responder de manera afirmativa estas preguntas considerará seguramente que la perspectiva que expusimos en el

inicio es incorrecta. En lo que dijimos antes parece sugerirse que ninguna clase existe por sí misma, que *todas* las clases tienen el mismo *estatus ontológico* y que, si alguna de ellas recibe reconocimiento, esto se debe en parte a su relevancia práctica. El ejemplo de Borges nos puede llevar a pensar que, en realidad, hay ciertas clases que existen por sí mismas, mientras que otras son creaciones humanas. La clase de los gatos y la de los pájaros existen por sí mismas. La clase de los animales que acaban de romper un jarrón es una construcción humana. Las clases de las manzanas y las peras existen por sí mismas, pero la clase de los alimentos culinarios no. El oro, la plata, el zinc y todos los otros metales son clases naturales. Pero la clase de los metales caros no lo es. Se trata de una posición o tesis filosófica conocida como *realismo ontológico*. La tesis opuesta es la del *convencionalismo ontológico*, que sostiene que hay muy diversas maneras de clasificar el mundo, todas ellas en un mismo pie de igualdad, y que, si alguna de esas posibles maneras es adoptada, esto se debe a su utilidad o relevancia práctica. No hay clasificaciones más objetivas que otras, sino más útiles que otras.¹ Borges adopta esta posición cuando dice lo siguiente en el pasaje que viene a continuación del que acabamos de citar: “He registrado las arbitrariedades de Wilkins, del desconocido (o apócrifo) enciclopedista chino y del Instituto Bibliográfico de Bruselas; notoriamente, no hay clasificación del universo que no sea arbitraria y conjetural”.

Nuestro objetivo ahora es utilizar el concepto de clase para explicar qué es una definición, pues definir no es otra cosa que enunciar las condiciones que debe cumplir algo para ser aceptado o considerado como miembro de una clase. Al definir *bullying*, por ejemplo, indicamos el elemento (o conjunto de elementos) que debe tener una conducta para calificar como *bullying*: el acoso físico o psicológico prolongado por parte de una persona o grupo de personas hacia otra. Las definiciones suelen ser respuestas a preguntas del tipo: ¿qué es x? (donde x representa el concepto a ser definido), ¿en qué consiste x?, ¿cuál es la naturaleza de x?, ¿cuál es la esencia de x?, etcétera.

La búsqueda de definiciones persigue a veces únicamente un interés teórico. Para muchas personas, comprender la esencia de lo que llamamos causa, bueno, muerte, arte, etcétera, es algo lo suficientemente importante como para que necesite de una justificación ulterior. Pero a veces la búsqueda de la definición se hace con vistas a algún fin posterior. En ocasiones, responde a la necesidad de determinar si una instancia particular pertenece o no a una cierta clase. Por ejemplo, la búsqueda de la esencia de lo que llamamos ciencia ha tenido como finalidad, muchas veces, responder la cuestión de si el psicoanálisis es una ciencia; la búsqueda de una definición de arte, que tanto ha preocupado a los filósofos del

arte en el siglo xx, está motivada, en parte, por la pregunta de si los *readymade* de Marcel Duchamp o las esculturas de Wolf Vostell son objetos artísticos. A primera vista, esto podría parecer un asunto de poca monta. El psicoanálisis o las obras de Duchamp son lo que son, y no cambiarán por el hecho de que los coloquemos en una categoría u otra. Pero esa impresión es errónea. Si bien una cosa no se modifica por ser ubicada en una categoría u otra, puede modificarse la manera en que es percibida y tratada socialmente. Por ejemplo, las disciplinas científicas gozan de prestigio en nuestra sociedad. La cuestión de si el psicoanálisis es o no una ciencia es, en parte, la pregunta de si el psicoanálisis es merecedor de esa estima. Pero no es solo una cuestión de prestigio. Las personas y las cosas adquieren derechos o los pierden en función de la clase a la que pertenecen. Decir que un objeto es artístico equivale a decir que ese objeto tiene derecho a ser exhibido en una sala de arte o en un museo, que debe ser cuidado y preservado, etcétera. Esto se ve claramente en relación con la definición de muerte, tema al que nos referiremos en breve en esta introducción y que ha sido muy debatido en la filosofía de la medicina. No se trata de una mera cuestión clasificatoria el hecho de decidir si una persona está viva o muerta. Si una persona está muerta, sus órganos pueden ser removidos del cuerpo para ser trasplantados en otro, el cuerpo puede ser sepultado y sus bienes pueden ser entregados a sus familiares directos. Lo mismo ocurre con la definición de enfermedad, de la que también hablaremos en breve. Si algo es una enfermedad, un médico está obligado a tratarla, las obras sociales deben cubrir su tratamiento, etcétera. Finalmente, las definiciones suelen estar motivadas por el deseo de permitir la contrastación de las afirmaciones generales que podamos hacer acerca de una clase. Por ejemplo, que todas las enfermedades tienen un componente psicológico podría depender de lo que entendamos por enfermedad. Para determinar si es verdad que el arte ha estado presente en todas las fases de la historia humana, necesitamos no solo recabar evidencia empírica sino, en primer lugar, determinar qué debemos entender por arte.

La estructura de las definiciones

Las definiciones constan de dos partes, el *definiens* y el *definiendum*. En la definición anterior, *bullying* es el *definiendum*, es decir, el término que designa la clase a ser definida; el resto de la proposición, donde se enuncian las condiciones para que algo sea aceptado como *bullying*, es el *definiens*. Para relacionar *definiens* y *definiendum* se suele utilizar el verbo *ser*. Algunos autores utilizan también el signo de igualdad (=); otros, el signo de igualdad seguido de la abreviación

“def.” (=def.), y otros, el signo lógico de la equivalencia material (\leftrightarrow). Así, la definición de *bullying* también podría expresarse de esta manera:

Bullying =def. Acoso físico o psicológico prolongado por parte de una persona o grupo de personas hacia otra.

No parece haber razones de peso para preferir una de estas fórmulas en lugar de las otras. En el presente trabajo utilizaremos casi siempre el verbo *ser* y, ocasionalmente, el símbolo =def.

Como ya se ha dicho, el *definiens* enuncia las propiedades o características que son condición para pertenecer a la clase designada por el *definiendum*. Es preciso ahora hacer algunas observaciones acerca del tipo de condición de la que estamos hablando. Debemos diferenciar propiedades que son condición para pertenecer a una clase (lo que los lógicos denominan “condiciones lógicas”), de propiedades que, aunque aparezcan regularmente asociadas a los miembros de una clase, no son una condición para pertenecer a ella.

La definición de *enfermedad*

Tomemos como ejemplo la clase a la que llamamos *enfermedades*. Enfermedad es el concepto central de la medicina. La razón de ser de la medicina es la existencia de enfermedades. Existen diversas concepciones sobre la naturaleza de las enfermedades, por ejemplo, que una enfermedad es cualquier estado físico indeseable para un individuo. En el siguiente pasaje de un texto sobre la definición de enfermedad, el filósofo de la medicina Christopher Boorse critica esta creencia:

Es claro que las enfermedades pueden ser deseables bajo ciertas circunstancias. Una persona que haya tenido viruela bovina puede salvarse de morir en una epidemia de viruela. La miopía puede ser ventajosa para evitar el enrolamiento. La esterilidad en un mundo sin contracepción podría ser una bendición para padres con familias numerosas (1977: 545).

Boorse argumenta aquí que el ser indeseable es algo común a muchas enfermedades, pero no es un rasgo definitorio. No se trata de una condición *necesaria* para pertenecer a la clase. Podríamos pensar ejemplos en los que todos los miembros de una clase comparten una cierta característica y, sin embargo, esta no es una condición necesaria para la pertenencia. Por ejemplo, todas las barras de oro existentes pesan menos de una tonelada. Pero tener determinado peso no es

una condición necesaria para que algo sea oro. Más adelante nos referiremos a un método (el método de los casos) que podemos utilizar para determinar si un rasgo presente universalmente en los miembros de una clase constituye además una condición lógica para pertenecer a esa clase.

Existen filósofos que creen que ser indeseable sí es una condición necesaria en la definición de enfermedad y, en relación con la crítica de Boorse, dirán tal vez que la noción de indeseabilidad debería ser entendida en un sentido diferente del que la entiende Boorse. No vamos a discutir ahora este punto. Pero aunque los adversarios de Boorse tuvieran razón, deberían responder a una segunda objeción, planteada por este filósofo:

Hay muchas clases de condiciones físicas indeseables, que restringen nuestro bienestar físico, que no aparecen como enfermedades en los textos médicos. Es indeseable estar por debajo de la media en relación con cualquier cualidad física significativa, por ejemplo, altura, fuerza, resistencia, coordinación, reflejos, velocidad, belleza, etcétera. Es indeseable tener debilidades humanas universales como la necesidad de sueño y el acceso regular a agua y comida. Esas condiciones no son enfermedades, y uno nunca podría diferenciarlas de las enfermedades únicamente en función de su disvalor. Como cualquier persona baja sabe, el ser bajo puede reducir la calidad de vida mucho más a largo plazo que una alergia menor o que una infección viral.

Lo que Boorse dice aquí es que, aunque fuera verdad que ser indeseable en algún sentido es una condición *necesaria* para que algo sea una enfermedad, no es, sin embargo, una condición *suficiente*. Hay estados físicos indeseables que no califican como enfermedades. Por lo tanto, debe haber alguna otra condición a ser cumplida. La crítica de Boorse puede resumirse así: ser indeseable o no es una condición necesaria para ser una enfermedad o, si es una condición necesaria, no es suficiente. Enunciar condiciones de pertenencia necesarias y colectivamente suficientes es un requisito que toda buena definición debe cumplir.

Describir y estipular

Al decir “x es cualquier cosa que cumpla con las condiciones A, B... N”, una persona puede estar haciendo dos cosas diferentes. Puede estar *dando reconocimiento* a una clase por primera vez (esto es lo que hizo Murphy al definir *pupkus* como “el residuo húmedo que deja la nariz de un animal al tocar una ventana”) o puede estar *describiendo* una clase ya reconocida (esto es lo que hace alguien

que define *círculo* como “curva cerrada cuyos puntos equidistan del centro”). En el primer caso, decimos que la persona está estipulando un significado para un concepto y que la definición es *estipulativa*. En el segundo, que la persona está describiendo un significado ya aceptado y que la definición es *descriptiva*. Se trata de una distinción pragmática, que depende del contexto. Cuando el matemático William Jones, en 1706, definió el número π como el número que se obtiene de dividir la longitud de una circunferencia por su diámetro, ofreció una definición estipulativa del término. Quien afirmara hoy que “el número π es el número que se obtiene al dividir la longitud de una circunferencia por su diámetro”, ofrecería una definición descriptiva. Encontramos un ejemplo similar en el siguiente fragmento de la *Crítica de la razón pura* de Kant:

En todos los juicios en que se piensa la relación de un sujeto con el predicado (refiriéndome solo a los afirmativos, pues la aplicación a los negativos es luego fácil), esa relación es posible de dos maneras. O bien el predicado B pertenece al sujeto A como algo contenido (ocultamente) en ese concepto A, o bien B está enteramente fuera del concepto A, si bien en enlace con este. Al primer caso lo llamo “el juicio analítico”.

Kant ofrece aquí una definición estipulativa del concepto-término *juicio analítico*. Sin embargo, quien dice actualmente que un juicio analítico es aquel donde el predicado del juicio expresa algo contenido implícitamente en el sujeto, no ofrece una definición estipulativa, sino descriptiva del uso filosófico aceptado de ese concepto.

Redefiniciones

La redefinición de un concepto es una forma de estipulación. La diferencia con otros tipos de estipulaciones reside en que, cuando redefinimos un término, estipulamos una definición para un término que se encuentra ya en uso. La definición de *razonamiento válido* como “razonamiento donde es imposible que las premisas sean verdaderas y la conclusión falsa” es una redefinición del concepto ordinario de validez. Cuando redefinimos un término podemos darle un significado completamente nuevo, pero puede ocurrir (es lo más común) que el significado nuevo tenga alguna relación con el significado ya aceptado. Es decir, al redefinir el concepto le damos reconocimiento a una clase que guarda alguna relación con la clase reconocida tradicionalmente a través de ese concepto. Lo más común es que nuestro objetivo sea dar visibilidad a una

subclase. Esto es lo que pasa en la definición técnica de validez. Los razonamientos válidos, en un sentido lógico, son una subclase de los razonamientos considerados válidos en un sentido preteórico. Pero al definir *válido* de manera tan restrictiva, muchos argumentos que serían considerados válidos, en un sentido coloquial, ahora quedarán excluidos de la clase así redefinida. También puede haber redefiniciones que hagan la clase originaria más amplia. Es raro que una redefinición realce una clase que no tenga ninguna conexión con la clase tradicionalmente reconocida.

Un ejemplo clásico es la definición de Hume del concepto de *causa*. ¿Cuál es el significado preteórico de *causa*? Para comenzar, la idea preteórica de causa contiene la idea de precedencia temporal: una condición necesaria para que un hecho pueda ser siquiera contemplado como posible causa de otro es que ocurra antes que él. Si el dolor de cabeza desapareció antes de haber tomado la aspirina, la aspirina no fue la causa de que el dolor se fuera. Debe haber además cierta contigüidad temporal. No diríamos que el impacto de una piedra causó la rotura de un vidrio si una cosa no hubiera ocurrido justo después de la otra. Si la rotura del vidrio se produjo mucho después del impacto, el responsable no fue el impacto (tal vez fue un rajadura, cuya causa fue el impacto). Debe haber también una contigüidad espacial: el impacto debe producirse en el lugar donde se produjo la rotura. Si la piedra impactó en una ventana y la rotura se produjo en otra, la causa de la rotura no fue el impacto de la piedra (a lo sumo, la causa pudo haber sido el cimbronazo producido por el impacto de la piedra en el otro vidrio).

Pero estas tres condiciones no son suficientes. Hay muchos pares de eventos A y B en los que uno es anterior al otro y donde existe contigüidad (temporal y espacial), y sin embargo no decimos que uno es causa del otro. Por ejemplo, alguien grita al verme lanzar una piedra. El grito es anterior y contiguo a la rotura del vidrio, pero esto no basta para decir que fue el grito la causa del estallido del vidrio. Es necesario que exista una conexión necesaria. Esto es: no es posible que la piedra golpee el vidrio sin romperlo. En cambio, sí es posible que, tras el grito, el vidrio no se rompa. Si alguien dijera que es perfectamente posible que el vidrio no se rompa al lanzar la piedra es porque, sin darse cuenta, está imaginando una ventana diferente a la que, de hecho, se rompió (por ejemplo, una ventana con un vidrio un poco más grueso) o una piedra lanzada con menos fuerza. Si tomamos la *misma* piedra y la lanzamos con la *misma* fuerza y de la *misma* manera sobre un vidrio del *mismo* espesor, etcétera, y el vidrio no se rompe, entonces la piedra no fue la causa de la rotura la primera vez. La idea de causa contiene en sí la idea de conexión necesaria. Podemos dar entonces la siguiente definición de causa:

A es causa de B si y solo si A es anterior a B, A es temporal y espacialmente contiguo a B, y la ocurrencia de A está conectada de manera necesaria con la ocurrencia de B.

La definición de causalidad que se acaba de dar pretende *describir* lo que en el lenguaje cotidiano entendemos como causa. Pero esa definición no implica que en el mundo físico haya verdaderamente relaciones a las que el concepto pueda ser aplicado legítimamente. Uno puede definir el concepto de unicornio como “caballo que posee un cuerno de añil en la frente”, pero esto no implica que exista en la realidad un animal al que se aplique concepto. ¿Existen en nuestro mundo relaciones causales tal como acaban de ser definidas? En principio, parecería que sí, pues todo el tiempo hacemos afirmaciones sobre relaciones causales entre fenómenos. Sin embargo, David Hume, el filósofo que inició la discusión moderna en torno al concepto de causalidad, dio una respuesta negativa a esta pregunta.

Nada de lo que observamos en el mundo físico y en nosotros mismos corresponde a esa noción de causalidad. Podremos observar la anterioridad y la contigüidad de ciertos fenómenos, pero ninguna conexión necesaria entre ellos. Observamos el descenso de la línea de mercurio en el termómetro y luego que el agua se congela, observamos que ambos sucesos son contiguos (espacial y temporalmente), pero nada en esas observaciones excluye la posibilidad de que, luego de la caída de la temperatura, el agua entre en ebullición, en lugar de congelarse. La conexión necesaria no es observable y, por lo tanto, de acuerdo con Hume, nunca estaríamos en condiciones de afirmar que un fenómeno es causa de otro. Si deseamos contar con un concepto “usable” de causalidad, esto es, que pueda ser empleado para hacer aseveraciones informativas acerca del mundo, el concepto de causalidad debe ser redefinido.

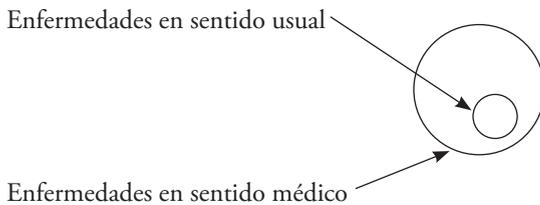
De acuerdo con esta experiencia, podemos, pues, definir una causa como *un objeto seguido de otro, cuando todos los objetos similares al primero son seguidos por objetos similares al segundo* (Hume, 1994: 101).

Hume redefine aquí el concepto de causa.¹ Su definición es más amplia que la preteórica. La clase de las causas, como la define Hume, incluye más cosas que la clase tal como es definida preteóricamente.

1 Hume ofrece una redefinición alternativa que, por falta de espacio, no podremos analizar en el presente contexto: “Podemos [...] dar otra definición de causa y llamarla *un objeto seguido por otro y cuya aparición siempre conduce al pensamiento a aquel otro*” (1994: 101).

Como ya se ha dicho, cuando se redefine un término es esperable que la nueva clase guarde alguna relación con la clase tradicionalmente reconocida con ese mismo término, pues si la clase a la que vamos a aplicar el término redefinido “no tiene nada que ver” con la antigua clase, la redefinición se verá expuesta inmediatamente a la siguiente objeción: ¿por qué estamos utilizando ese término para designar la nueva clase? Si lo que Hume denomina causa no tuviera ninguna relación con el significado tradicional del término, sería justo preguntar: ¿por qué darle a eso el nombre de causa en lugar de cualquier otro nombre?

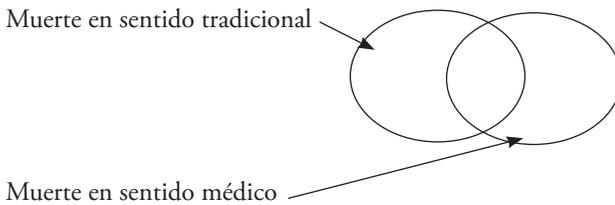
También en el campo de la medicina encontramos ejemplos de definiciones “amplias”. En el texto ya citado, Boorse señala que algunos manuales de medicina categorizan como enfermedades cosas como fracturas, heridas de armas de fuego, cuerpos extraños en el estómago, mordeduras de animales, electrocución, incineración, etcétera. Esos manuales incluyen en la clase de las enfermedades más cosas que las que incluye la clase de las enfermedades en sentido corriente:



Pero consideremos ahora el concepto *muerte*. De acuerdo con la definición tradicional de muerte, una persona está muerta si han cesado sus funciones cardiorrespiratorias. Una persona está muerta si no tiene pulso y su corazón no late. Dos hechos pusieron en tela de juicio esta definición. En primer lugar, la aparición de los respiradores artificiales a mediados del siglo pasado. Estos aparatos permiten mantener artificialmente en actividad las funciones cardíacas y respiratorias de una persona aun cuando sus funciones cerebrales hayan cesado de forma irreversible. Esas personas tienen pulso y sus corazones laten, pero nunca recuperarán la conciencia y dejarían de respirar y su corazón de latir si fueran desconectadas de las máquinas. ¿Tenemos derecho a decir que esas personas están vivas?

El otro hecho que conmocionó la definición tradicional de muerte fue la invención, en la misma época, de las técnicas de reanimación artificial (respiración boca a boca, compresiones cardíacas). Gracias a estas técnicas, es posible reanimar a una persona en los minutos siguientes de ocurrido un

paro cardiorrespiratorio. Ya no parecía que la mera ausencia de signos vitales fuera una razón suficiente para declarar muerta a una persona. Por eso, en la medicina actual el concepto de muerte ha sido redefinido: una persona es considerada muerta si y solo si sus funciones cerebrales se han detenido de forma completa e irreversible. Podemos describir la relación existente entre la antigua clase y la resultante de la redefinición como una relación de intersección de dos conjuntos. Muchos casos considerados como muerte, según la definición tradicional, son también casos de muerte según la definición médica. Pero hay casos de muerte según la definición tradicional que quedan excluidos de la definición médica (personas que podrían ser resucitadas mediante técnicas de reanimación), y hay casos que quedan excluidos de la definición tradicional que son considerados como casos de muerte en la definición médica (personas con muerte cerebral cuyas funciones cardiorrespiratorias son mantenidas activas mediante un respirador artificial).



Definiciones persuasivas

Hay un tipo de redefinición conocida como *definición persuasiva*. El siguiente ejemplo puede ser de ayuda. Todos sabemos qué significa ser una persona culta: alguien que ha leído mucho y que está familiarizado con las artes. No hay duda, por otro lado, de que ser culto es algo que está bien conceptualizado en nuestra sociedad. Si yo me refiero a alguien dando a entender que es una persona culta, esa persona suscitará la estima de mi interlocutor e indicará también que esa persona es objeto de mi estima. Esto mismo pasa con muchas clases, por ejemplo, con la de los sistemas políticos que llamamos *democracias* o con la de los sentimientos que denominamos *amor*. Esas clases son valoradas, estimadas, honradas, pero también hay clases que son objeto de los sentimientos opuestos (los ignorantes, las dictaduras, el odio) y otras que no generan ninguna actitud favorable ni desfavorable. Las actitudes en relación con una clase

pueden variar de individuo a individuo y pueden depender del grupo social al que pertenecen, de su ideología y de muchos otros factores. Los obreros son valorados por parte de los intelectuales de izquierda, de modo que si uno de ellos califica a alguien como obrero mostrará con ello respeto y consideración por esa persona. Un intelectual de derecha no expresará probablemente los mismos sentimientos al aplicar ese término a un individuo.¹¹

Supongamos ahora que uno quiere dar reconocimiento a una clase hasta ahora no reconocida y crear actitudes (favorables o desfavorables) en relación con ella. Si la clase en cuestión está emparentada con una clase ya existente que despierta el mismo tipo de sentimiento que uno quiere generar en relación con la clase nueva, podría ser conveniente designarla con el mismo término. Uno no debe hacer explícito que el término está siendo utilizado para designar un nuevo tipo de objeto, pues nuestra audiencia se podría preguntar si el nuevo tipo de objeto merece la misma valoración. El nuevo tipo de objeto debe ser introducido subrepticamente en la clase que ya goza de una determinada valoración. Es posible incluso que en la nueva clase ya no quede lugar para los antiguos miembros de la clase, y que los nuevos sean proclamados como los verdaderos miembros. Volviendo al ejemplo anterior, supongamos que una persona cree que lo que debe ser valorado no es tanto si una persona ha leído mucho o ha visitado muchos museos, sino que eso haya desarrollado en ella cierta sensibilidad imaginativa. Puede decidir entonces definir *culto* como “persona con sensibilidad imaginativa” (sea esto el resultado de lecturas, de visitas a museos o de cualquier otra cosa). ¿Está describiendo el significado aceptado del término o estipulando un nuevo significado? Claramente, lo segundo. Pero lo está haciendo de una manera velada y tendenciosa, pues está aplicando la palabra a algo que está relacionado con la clase a la que aplicamos normalmente el término (es uno de sus efectos), induciendo así a una equivocación. Con ello espera que la nueva clase reciba el beneplácito que normalmente reciben las personas cultas. Está formulando lo que Charles Stevenson denominó una “definición persuasiva” (1938). Si alguien descubriera la maniobra y objetara que no es eso lo que llamamos “una persona culta”, nuestro personaje podría responder que él está hablando de las “verdaderas” personas cultas. Es importante aclarar que, a diferencia de la palabra *culto*, muchas otras carecen de un significado lo suficientemente preciso como para determinar fehacientemente si una definición persuasiva ha tenido lugar. Si una clase no tiene límites precisos (por ejemplo democracia o libertad), habrá definiciones acerca de las cuales no podremos decir si son descriptivas o persuasivas.¹¹¹

Las definiciones persuasivas, por lo general, surgen del deseo de que un cierto individuo goce de la consideración de la que es objeto una cierta clase. Stevenson menciona como ejemplo la discusión acerca de si Alexander Pope es un poeta, siendo *poeta* una clase que goza de prestigio social. De acuerdo con Stevenson, la discusión acerca de si Pope era un poeta no era una mera discusión terminológica. Lo que estaba en juego era si la obra de Pope debía gozar o no del prestigio del que gozan las obras consideradas poéticas. Una discusión similar, más cercana a nosotros, es aquella acerca de si el psicoanálisis es o no una ciencia. No se trata de una mera cuestión clasificatoria. Sus defensores quieren que esta disciplina sea galardonada con el título de ciencia, y sus detractores desean que ese título le sea negado.^{IV} La pregunta sobre si el psicoanálisis es o no una ciencia conducirá inevitablemente a la cuestión de definir la naturaleza del método científico, y esa definición seguramente tendrá un carácter persuasivo.

Supongamos ahora que, en lugar de crear una nueva clase y proyectar sobre ella ciertos sentimientos, queremos conservar una clase ya existente, pero modificar los sentimientos y actitudes (positivos o negativos) que las personas tienen en relación con ella. La clase se encuentra ya socialmente reconocida y los hablantes expresan ciertas actitudes cuando hablan de ella o de los individuos que la componen. El primer paso es crear un nuevo término para designar la clase, un término que no tenga las resonancias negativas o positivas que la clase en cuestión normalmente tiene, un término que pueda ayudar a las personas a expresar el nuevo tipo de actitud buscada. Por ejemplo, si la clase de personas llamadas *indios* está valorada negativamente en una sociedad y deseamos luchar para modificar ese tipo de valoración, debemos encontrar un nuevo nombre que la designe, por ejemplo *pueblos originarios*. No tenemos aquí una redefinición, sino la búsqueda de un nuevo rótulo para una clase que ya ha sido reconocida.^V

Stevenson menciona algunos ejemplos interesantes del uso de definiciones persuasivas en filosofía. Uno de ellos involucra al filósofo Baruch Spinoza. En su *Ética demostrada según el orden geométrico*, Spinoza reconoce la existencia de una única sustancia que posee infinitos atributos, a la que llama “Dios”. Es interesante que Spinoza llame Dios a esta sustancia. “Su” Dios guarda poca o nula relación con lo que usualmente incluimos en esa categoría. Stevenson sostiene que, con la nueva definición de Dios, Spinoza pretendía inducir en sus lectores los mismos sentimientos de veneración que nos inspira la representación de Dios en la acepción corriente del término. Un caso igualmente ilustrativo es la definición de *justicia* que Platón ofrece en el libro IV de *La República*. Luego de describir las clases sociales de la república ideal (los gobernantes,

los guardianes, los productores) y las tareas que le correspondería a cada una, Sócrates define *justicia* como aquel estado en el “que cada una de las tres clases realiza a la perfección la tarea que le es propia” (401). Platón concordaría en que esa definición no describe el significado corriente del término, pero entonces, ¿por qué llamar justicia a ese estado? ¿Por qué no denominarlo con cualquier otro nombre? Lo que busca Sócrates es dignificar una cierta clase de conducta (hacer únicamente aquello que corresponde por la pertenencia de clase) dándole el nombre de algo que es muy apreciado socialmente.

La paradoja del análisis

Los lectores de este libro probablemente conozcan la célebre historia de Ulises y la tortuga, narrada por Zenón de Elea. Zenón imagina qué ocurriría si el héroe griego tuviera que correr una carrera contra una tortuga, a la que le diera unos metros de ventaja. La experiencia nos dice que Ulises alcanzaría rápidamente a la tortuga y ganaría la carrera sin ninguna dificultad. Pero Zenón, a partir de principios elementales de divisibilidad del espacio, concluye que sería imposible que Ulises alcanzara a la tortuga. El argumento propuesto por Zenón se conoce como la paradoja de Ulises y la tortuga. Una paradoja es un razonamiento aparentemente correcto que demuestra que algo, que la experiencia nos enseña que es falso, es verdadero. Una paradoja es, por así decir, una colisión entre lo que la razón nos dice y lo que nos dice la experiencia. Las definiciones (descriptivas) también son el origen de una paradoja muy conocida: la paradoja del análisis. La experiencia nos dice que, muchas veces, somos capaces de identificar objetos como pertenecientes a una cierta clase, pero que somos incapaces de definir esa clase correctamente. Que alguien esté en condiciones de identificar círculos, no quiere decir que esté en condiciones de definir *círculo*. Que alguien sea capaz de identificar y hablar de las causas de un fenómeno (“la lluvia causó un apagón generalizado” o “la causa del accidente no fue el mal tiempo, sino que el conductor pasó el semáforo en rojo”), no implica que podrá definir qué es una causa. Pero vayamos ahora al razonamiento que nos dice lo contrario. ¿Cómo es que alguien decide si un objeto pertenece o no a una cierta clase? Podemos suponer que chequea si el objeto en cuestión cumple con las condiciones que son necesarias para pertenecer a la clase, por ejemplo, identifica un fenómeno como la causa de otro porque el primero reúne las condiciones de lo que denominamos causa, lo cual implica que *sabe* cuáles son las condiciones necesarias

para que algo sea denominando causa. Pero, si es así, ¿cómo es posible que luego no sea capaz de decirnos cuáles son esas condiciones?

Hay una manera alternativa de plantear la paradoja del análisis. A pesar de que conocemos el significado del concepto definido, y a pesar de que la definición procura expresar ese significado ya conocido, el contenido de la definición puede tener para nosotros, hablantes nativos de la lengua, el carácter de un descubrimiento. Conocemos el significado de la palabra y, sin embargo, parecemos aprender algo nuevo cuando se nos enseña que la circunferencia es la figura cerrada donde todos los puntos equidistan de un mismo punto. Esto, que es válido para la definición de conceptos matemáticos, resulta aún más claro para los conceptos de otras áreas, en especial para los que interesan a la filosofía: causa, explicación, identidad, justicia, virtud y muchos otros. Si el razonamiento mencionado más arriba es correcto, esto no debería ser así. Si conocemos el significado de un concepto, su definición siempre debería resultarnos trivial.

Una tercera manera de expresar la paradoja consiste en decir que si nuestro primer argumento es correcto, entonces si dos personas conocen el significado de una palabra no debería haber ningún desacuerdo acerca de cuál es la definición correcta de esa palabra. Si dos personas utilizan *triángulo* con el mismo significado, deberían ponerse de acuerdo inmediatamente acerca de su definición. Si hay desacuerdo, entonces una de ellas no sabe realmente qué es un triángulo o el término tiene para cada una significados diferentes. Y lo que vale para el concepto *triángulo* deberá ser válido para cualquier otro concepto. Esta conclusión es contraria a lo que la experiencia nos enseña. A veces, la tarea de encontrar la definición correcta de un concepto puede resultar trabajosa. Podemos equivocarnos a la hora de proponer una definición o pueden surgir desacuerdos respecto de cuál es la definición correcta de un concepto. La filosofía es un claro ejemplo. Como veremos más adelante, los filósofos no han podido ponerse de acuerdo aún acerca de la naturaleza de la justicia, de las proposiciones necesarias, de la identidad personal, de las leyes científicas y de muchas otras cosas.

Cuando surge una paradoja se supone, en general, que quien se equivoca es la razón, esto es, se supone que quien formula la paradoja ha cometido algún error sutil en su razonamiento, lo que lo ha llevado inadvertidamente a una conclusión absurda. Varias explicaciones han sido propuestas para dar cuenta del error que habría cometido Zenón, y otras han sido ofrecidas también sobre el error que estaríamos cometiendo cuando caemos en la paradoja del análisis. Es así como las paradojas, en general, se resuelven: identificando el error de razonamiento que nos condujo a contradecir la experiencia. Voy a ofrecer a

continuación una posible solución a la paradoja del análisis. El error surge de no comprender que podemos tomar una decisión racional sin ser conscientes de cómo la hemos tomado. Un ejemplo sencillo puede aclarar esto. Los jugadores de fútbol suelen eludir a sus adversarios haciendo un *dribbling* (en nuestras pampas diríamos una gambeta), que consiste, básicamente, en engañar al adversario mediante un amague (dicho sea de paso, esto es una definición descriptiva del término *dribbling*). Existen muchos tipos de *dribblings* y cada buen delantero domina en general varios de ellos. Los buenos jugadores saben además cuál es el *dribbling* más adecuado para cada situación dada. Pero esto no significa que puedan describir verbalmente ese conocimiento (“en la situación de tipo x , el *dribbling* del tipo y es lo más aconsejable”) ni tampoco que apliquen ese conocimiento de manera consciente (“opté por el *dribbling* y porque la situación era de tipo x ”). Aceptando que esta descripción es correcta, podemos suponer que en el plano del uso de conceptos ocurre algo similar. Un hablante competente de castellano dispone de la teoría que guía el uso de conceptos como *causa* (“para que un fenómeno pueda ser considerado como la causa de otro debe reunir las condiciones x, y, z ”), pero casi nunca es capaz de poner en palabras esa teoría ni de explicar en una situación puntual por qué decidió calificar como causa un cierto fenómeno. Hacerlo requiere un trabajo de reflexión a partir de la práctica lingüística.

Evaluando definiciones

Tanto las estipulaciones como las descripciones pueden ser objeto de una evaluación. Pero tratándose de actos que persiguen objetivos diferentes, serán diferentes los criterios que utilicemos para evaluarlos. En cuanto a las estipulaciones, nos interesa saber para qué son útiles, para qué sirve darle visibilidad a una cierta clase. Por ejemplo, ¿para qué necesitamos el concepto *pupkus*? Nuestra capacidad de memorizar conceptos es limitada. No podemos disponer de un concepto para todas las clases. Solo aquellas que tengan una utilidad real deberán ser reconocidas. Las otras pueden pasar desapercibidas.

En cambio, si se trata de una descripción, nos interesará saber ante todo si esta es fiel, es decir, ajustada al significado aceptado de un concepto. ¿Cómo podemos determinar esto? En primer lugar, las definiciones descriptivas deben estar de acuerdo con nuestras intuiciones acerca de la aplicación del concepto definido a casos reales o hipotéticos. Recordemos la objeción de Boorse a la definición de enfermedad como estado indeseable. Él nos invita a imaginar a

alguien que, durante una guerra, tuviera miopía, lo cual le permite no ser enroldado. Nuestro conocimiento de la psicología humana nos permite comprender que, en esas circunstancias, la miopía podría ser algo deseable. ¿Diríamos por eso que la miopía ha dejado de ser una enfermedad? Creo que nadie diría algo así. Pero aquí surge una dificultad: ¿qué nos hace estar tan seguros de que la miopía seguiría siendo, aun en ese contexto, una enfermedad? Uno podría responder: el significado del término *enfermedad*. Pero ¿cómo puede ser el significado lo que nos guía en la respuesta si el significado del término es justamente algo que estamos tratando de develar? Recordemos lo que dijimos en relación con la paradoja del análisis: un buen jugador de fútbol conoce cada uno de los movimientos necesarios para hacer un *dribbling*, pero es incapaz de describirlos. Los conoce pero no en un plano consciente. Si quisiera describirlos debería, por ejemplo, filmarse haciendo un *dribbling* y, a partir de allí, reconstruir la técnica que utilizó para hacer la jugada. Con los conceptos sucede algo parecido. Somos capaces de “jugar” con los conceptos, pero no somos conscientes de cómo tomamos la decisión de aplicar un concepto a un determinado objeto y no a otro. Lo que debemos hacer, entonces, es observarnos “jugando” y, a partir de las decisiones que tomamos, reconstruir las reglas de aplicación subyacentes. Cuando le pedimos a un jugador que realice una jugada para que nosotros podamos entender cómo la hace, es importante que la realice sin pensar, que se deje llevar por su instinto de jugador. De la misma manera es importante que, al decidir si la miopía es o no una enfermedad en la situación imaginaria, nos dejemos llevar por nuestras intuiciones lingüísticas. Esa es la mejor guía para dar con el contenido del concepto.

El método para evaluar definiciones descriptivas que acabamos de comentar se conoce como *método de los casos*. Lo característico de este método consiste en imaginar ejemplos reales o hipotéticos acerca de los cuales debemos emitir un juicio. Este método tiene dos variantes. En la primera imaginamos una situación que cumple con las condiciones enunciadas en el *definiendum* y debemos juzgar si esa situación califica para que se le aplique el *definiens*. Una de las objeciones de Boorse a la definición de enfermedad como estado indeseable consiste precisamente en eso. Boorse toma como ejemplo la necesidad de los seres humanos de tener acceso regular a agua y comida. La necesidad de agua y comida nos limita espacial y temporalmente, pues donde vayamos tenemos que asegurarnos de que existan provisiones. Si no tuviéramos esas necesidades, nuestra existencia sería mucho más despreocupada. Por otro lado, la necesidad imperiosa de agua y comida nos obliga a tener que trabajar para conseguirlas. En conclusión, la necesidad constante de agua y comida es algo indeseable. Sin

embargo, ¿son enfermedades? Claramente no. La segunda variante del método de los casos consiste en imaginar primero una situación que califique para el *definiendum* y considerar luego si cumple con las condiciones enunciadas en el *definiens*. Esta variante también es ilustrada por algunas de las objeciones de Boorse. La miopía es una enfermedad. Pero ¿es necesariamente algo indeseable? Imaginemos a alguien a quien ser miope le permitiera salvarse de ser enrolado en el ejército. ¿No podría la miopía volverse algo deseable para esa persona? Eso sería posible. La primera variante muestra que las condiciones enunciadas en el *definiens* no son suficientes. La segunda muestra que las condiciones no son necesarias.

Algunas otras definiciones erróneas de enfermedad comentadas por Boorse podrán ayudarnos a aclarar este punto. Se suele suponer que una persona sana es aquella que no requiere de tratamiento médico, y que una persona enferma es la que necesita tratamiento. Esto podría llevarnos a definir la enfermedad como un “estado físico que un médico atiende”. Pero esta definición es incorrecta, pues ser tratable por un médico no es ni una condición necesaria (el mal de Chagas no tiene tratamiento médico y es una enfermedad) ni suficiente para que algo sea una enfermedad (los médicos estéticos tratan los signos del envejecimiento, que no son enfermedades). Otra definición común de enfermedad criticada por Boorse sostiene que esta es el estado en el que se encuentra una persona cuyos parámetros físicos están por fuera de los rangos considerados estadísticamente normales dentro de una sociedad. Por ejemplo, la hipertensión sería una enfermedad porque el nivel de presión arterial de las personas hipertensas está por encima del nivel de presión arterial estadísticamente normal. El enanismo sería una enfermedad porque la altura de las personas enanas está muy por debajo del rango de altura estadísticamente normal en los seres humanos. Esta definición sociológica tampoco puede ser correcta. Las personas pelirrojas o las que tienen sangre tipo cero negativo representan casos inusuales y, sin embargo, no son consideradas enfermedades. Esto muestra que la condición de ser anormal en términos estadísticos no es una condición suficiente para ser una enfermedad. Por otro lado, hay muchos estados que podrían volverse típicos (por ejemplo, la infertilidad en la película posapocalíptica *Children of men*), pero no por eso dejarían de ser enfermedades. Esto muestra que tampoco lo típico es una condición necesaria. Algunas personas consideran que una enfermedad es cualquier estado físico que produzca dolor. Pero, de nuevo, el dolor no es ni una condición suficiente (una rotura de ligamentos o un embarazo producen dolor y no son enfermedades) ni necesaria (hay enfermedades del corazón que son asintomáticas). Finalmente, se suele creer que una enfermedad es cualquier estado físico

que produzca algún grado de incapacidad. La miopía sería una enfermedad pues restringe la capacidad de ver, y el alzhéimer sería una enfermedad pues restringe capacidades cognitivas. Siendo la muerte la más extrema de las discapacidades, podríamos suponer que hay muchos estados que son considerados enfermedades no porque impliquen de manera inmediata una incapacidad, sino porque implican un riesgo (alto o bajo) de muerte, por ejemplo, el insomnio o el cáncer, si no son tratados. Pero esta definición tampoco es correcta. Hay enfermedades que no suponen ninguna discapacidad (por ejemplo, el pie de atleta) y hay estados que implican una discapacidad y no son enfermedades (la ausencia en los seres humanos de órganos visuales capaces de ver en la oscuridad). Como conclusión, podemos decir que características como tener un estado físico que los médicos atienden, ser atípico, ser doloroso o ser incapacitante no hacen parte de la naturaleza, de la esencia, de las enfermedades. Se trata, en cambio, de características concomitantes, que suelen acompañar a una enfermedad, pero no son *la* enfermedad.

Como ya se ha dicho, es importante que el juicio que hagamos acerca de la situación sea lo más espontáneo posible. Si nuestro interlocutor pensara demasiado acerca de si la situación califica o no para aplicarle el *definiendum*, su juicio podría quedar contaminado por nociones que no son las que queremos investigar. Por ejemplo, en nuestro caso queríamos investigar el significado preteórico de enfermedad. Para ello, es importante que responda sin pensar demasiado. Como ocurre en otros casos, las nociones socialmente compartidas acerca de un determinado asunto tienen más posibilidades de manifestarse cuando las personas no ejercen control sobre lo que deben decir.

En filosofía, los ejemplos más antiguos del uso de este método se encuentran en los diálogos platónicos. Sobre la búsqueda de definiciones en los diálogos de Platón hablaremos más adelante. Por el momento, mencionaremos apenas un ejemplo para clarificar lo dicho en el párrafo anterior. En *La República* de Platón, Sócrates discute con sus interlocutores sobre la naturaleza de la justicia. Uno de ellos, Céfalo, responde que es *justo* aquel que dice la verdad y devuelve siempre las cosas que ha recibido en préstamo. Pero Sócrates imagina un ejemplo para mostrarle a Céfalo que esa definición es incorrecta:

–De perlas –contesté yo–, es lo que dices, Céfalo; pero eso mismo de que hablamos, esto es, la justicia, ¿afirmaremos que es simplemente [...] el devolver a cada uno lo que de él se haya recibido, o estas mismas cosas se hacen unas veces con justicia y otras sin ella? Pongo por caso: si alguno recibe unas armas de un amigo estando este en su juicio, y ese amigo se las

pide después de vuelto loco, todo el mundo diría que no debe devolvérselas [...] a quien se halla en semejante estado.

–Bien dices –afirmó él.

–Por lo tanto, no se conforma la justicia [...] en devolver lo que se ha recibido (331, c-d).

Notemos que la objeción de Sócrates solo tiene sentido si presuponemos que Céfalo ha tratado de ofrecer una definición descriptiva de “ser justo”. Céfalo tiene que estar dispuesto a someter esa definición al examen de las intuiciones. Céfalo podría decir: “Yo sé que la intuición nos dice que devolver las armas a un amigo en esas circunstancias sería injusto. Pero, para mí, ser justo es decir la verdad y devolver las cosas que uno ha recibido en préstamo”. Estaría con ello redefiniendo el concepto de justicia. Por otro lado, el ejemplo imaginado por Sócrates ilustra la primera variante del método de los casos: devolver a cada uno lo que de él se ha tomado en préstamo, no es una condición suficiente para hablar de justicia.

Un ejemplo adicional de la utilización del método de los casos en Platón se encuentra en el *Teeteto*, un diálogo que tiene entre sus objetivos la definición de *conocimiento*. En un pasaje, el personaje Teeteto propone definir *conocimiento* como equivalente de *percepción*. Conocer algo es percibirlo. Sócrates no tiene dificultades en refutar esta definición mediante un caso hipotético:

Sócrates: Decimos que alguien que ve adquiere conocimiento de lo que ve, porque nos pusimos de acuerdo en que la vista o la percepción y el conocimiento son lo mismo.

Teeteto: Muy bien.

Sócrates: Pero supongamos que este hombre que ve y que adquiere conocimiento de lo que ve, pierde sus ojos; en tal caso, recuerda la cosa, pero no la ve, ¿no es así?

Teeteto: Así es.

Sócrates: Pero “no la ve” significa “no la conoce”, puesto que “ve” y “conoce” significan lo mismo.

Teeteto: Es verdad.

Sócrates: Entonces, concluimos que un hombre que llegó a conocer una cosa y aún la recuerda, no la conoce, ya que no la ve; y dijimos que esta conclusión es monstruosa.

Teeteto: Completamente de acuerdo.

Sócrates: En consecuencia, pues, si dices que percepción y conocimiento son lo mismo, nos vemos llevados, por esa afirmación, a sostener algo imposible.

Teeteto: Entonces tendremos que decir que percepción y conocimiento son diferentes (164, a-b).

Cuando alguien consigue montar de manera efectiva una escena imaginaria para refutar una definición decimos que ha ofrecido un *contraejemplo* para la definición que está siendo evaluada. Si encontramos un contraejemplo al aplicar el método de los casos, la definición obviamente queda refutada. Pero ¿qué decir si fallamos en encontrar un contraejemplo y todos los ejemplos que se nos ocurren son a favor de la definición? ¿Mostramos con ello que la definición es correcta? La imposibilidad de encontrar evidencia que desmienta una opinión no muestra que la opinión sea correcta. Creer lo contrario es cometer una falacia conocida como *ad ignoratium*. Pero cuando se han hecho intentos sistemáticos por desmentir un punto de vista y todos ellos han sido infructuosos, la ausencia de evidencia contradictoria podría ser tomada, si no como una razón concluyente a favor, al menos como un *indicio* de que el punto de vista podría ser correcto. Así, si hemos hecho reiterados intentos por probar la incorrección de una definición y todos ellos han sido infructuosos, podríamos tomar esto como indicio de que la definición propuesta es correcta.

Muchos de los contraejemplos que se utilizan en filosofía describen situaciones imaginarias, ficticias, contrafácticas. Sería absurdo que Céfalo rechazara el ejemplo de Sócrates por ser inventado. Un contraejemplo no pierde nada de su fuerza por el hecho de ser hipotético o ficticio. La única condición que se suele requerir es que las situaciones imaginadas sean coherentes, esto es, que no impliquen una contradicción. Por ejemplo, no sería aceptable postular escenarios en los que hubiera objetos imposibles, como círculos cuadrados u hombres que no fueran humanos. Más allá de esto, cualquier situación concebible es idónea para servir como contraejemplo.

Quiero que consideremos ahora el siguiente escenario. Un amigo me pregunta si sé el resultado del partido de ayer entre Cañuelas y El Porvenir. Le respondo que Cañuelas ganó 1 a 0. Mi amigo me pregunta entonces: “¿Sabés o creés que ganó 1 a 0?”. Mi amigo acaba de diferenciar dos clases de creencias: las meras creencias y aquellas que califican como saber. Me pregunta en cuál de esas dos categorías debe ubicar mi afirmación de que Cañuelas le ganó a El Porvenir (podríamos llamar a esto una pregunta de categorización). Pero ¿cuál es la diferencia entre creer algo y saberlo? Vamos a suponer que me equivoqué y que Cañuelas no le ganó a El Porvenir. Cuando mi amigo descubriera el error podría decir: “No lo sabías. En realidad, creías que Cañuelas había ganado”. Una diferencia crucial entre saber y creer es que, para “saber p”, mi creencia

en “p” debe ser verdadera. Pero supongamos que lo adiviné, tuve un palpito y ese palpito resultó correcto. Creo que nadie consideraría que, en ese caso, yo sabía de lo que estaba hablando. No basta que una creencia sea verdadera para calificarla como conocimiento, es necesario además que pueda justificarla a través de buenas razones. Tener un palpito no es una razón aceptable para afirmar que Cañuelas ganó. Que Cañuelas haya ganado los últimos partidos que disputó tampoco lo es. Que yo presencié el partido sí lo es. Yo podría responder entonces de manera categórica: “No lo creo. ¡Lo sé!”. Si analizáramos ejemplos rutinarios de *saber*, veremos que se trata siempre de creencias verdaderas y justificadas. Por ejemplo, sabemos que en caída libre los objetos se aceleran $9,8 \text{ m/s}^2$. Decimos que lo sabemos, en primer lugar, porque es algo en lo que creemos. En segundo lugar, porque esa creencia es verdadera. En tercer lugar, porque es una proposición que ha sido sometida a pruebas experimentales innumerables veces. Formulemos así la siguiente definición de conocimiento:

Conocimiento =def. Creencia verdadera justificada²

Pero consideremos ahora el siguiente escenario.³ Para saber la hora, miro el reloj que está colgado en la pared de la oficina. Es un reloj que siempre ha funcionado bien y al que le acabamos de cambiar las pilas. El reloj marca las 12. A partir de aquí formo la creencia de que son las 12 del mediodía. Esa creencia es verdadera porque, de hecho, son las 12 del mediodía. Pero nadie me ha dicho que el reloj no está funcionando y que, ¡vaya casualidad!, se descompuso justo a las 12 del día anterior y sus agujas quedaron marcando las 12 en punto. Y ahora, la pregunta: ¿diríamos en ese caso que *creo* o que *sé* que son las 12 en punto? La mayoría de las personas, cuando son confrontadas con este escenario, considera que yo no sé que son las 12 en punto. Es una creencia verdadera y está amparada en lo que parecen ser buenas razones (qué mejor razón para afirmar que es una cierta hora del día que lo marque un reloj que siempre funciona bien), pero sentimos que a esa creencia le falta algo para ser auténtico conocimiento. Si queremos excluir de la categoría de conocimiento la creencia recién mencionada, es necesario que agreguemos una condición extra a la definición que acabamos de formular.

Supongamos que encontramos la condición requerida. Esto no significa que no puedan aparecer nuevos contraejemplos. Y si esos contraejemplos aparecieran, nuevas condiciones deberían ser adicionadas para neutralizarlos. Lo

2 A idéntica conclusión llega Platón en el *Teeteto* (201 d).

3 El siguiente ejemplo es una adaptación de uno similar elaborado por Russell (1995).

que estaríamos buscando, entonces, sería una verdad necesaria, esto es, una proposición cuya refutación sea inconcebible. Tomemos por caso la proposición $2 + 2 = 4$. Es el paradigma de una proposición necesaria. Imaginemos que en una bolsa ponemos dos manzanas y luego dos más, pero cuando queremos sacarlas hay solo tres. Si eso ocurriera pensaríamos que alguien sacó una manzana de la bolsa sin que nos diéramos cuenta, o que la bolsa tiene un agujero o algo por el estilo. Pero si nos aseguráramos de que la bolsa no tiene ningún agujero y de que nadie la ha tocado, y si la misma situación se repitiera, pensaríamos que hemos contado mal o que no eran manzanas lo que pusimos en la bolsa (sino algún artilugio de magia), pero nunca concluiríamos que $2 + 2$ no siempre es 4. La cuestión de si las proposiciones necesarias resultan de la naturaleza de las cosas mismas, o si, por el contrario, expresan meramente una convención inquebrantable, es tema de debate en la filosofía y no necesita ser discutida aquí. Lo que está claro es que, si sometemos la definición de *conocimiento* a pruebas de la imaginación tan exigentes, es porque esperamos que el resultado de una definición sea una proposición necesaria. Hay un alto consenso entre los filósofos de que se trata de un objetivo razonable:

El análisis de un concepto no se aplica solo a cada cosa real que es un F. Se aplica a cada cosa que pudiera ser un F. El análisis del término *causa* no se aplica a todas las causas reales. El análisis se aplica a todas las cosas que pudieran ser una causa, a todas las causas posibles. Así, el análisis de un concepto debe ser entendido como afirmando que, para todo objeto posible, ese objeto es un F si y solo si es un G. Por eso, si el análisis es verdadero, será necesariamente verdadero (Daly, 2010: 45).

Un análisis correcto de *conocimiento* debe ser una verdad necesaria. Consecuentemente, los experimentos mentales hipotéticos deben ofrecer casos para poner a prueba apropiadamente diversas propuestas de análisis como las que veremos a continuación (Ichikawa y Steup, 2018).

Esta cuestión nos da pie para decir algo acerca de la forma de las definiciones. Desde un punto de vista formal, las definiciones son proposiciones universales. Por ejemplo, la definición de *conocimiento* nos dice que *todo* conocimiento es una creencia verdadera justificada. Pero nos dice también que toda creencia verdadera justificada es conocimiento. Esto marca una diferencia con otras proposiciones universales. Que todos los humanos son mamíferos no significa que todos los mamíferos son humanos. Si los enunciados universales son, desde el punto de vista de la lógica moderna, condicionales cuantificados uni-

versalmente, las definiciones pueden ser representadas como bicondicionales cuantificados universalmente:

$$(x) (Ax \longleftrightarrow Bx)$$

“A” representa aquí el *definiendum*, y “B”, el *definiens*. El símbolo del bicondicional indica que toda cosa que sea “A” será “B”, y que toda cosa que sea “B” será “A”. La definición estándar de *conocimiento* incluye en el *definiens* al menos tres propiedades: ser una creencia, ser verdadera y estar justificada. Su forma debería ser representada así:

$$(x) (Ax \longleftrightarrow Bx \ \& \ Cx \ \& \ Dx)$$

Dada una proposición universal del tipo “todos los A son B”, un caso particular que cumpla con las propiedades “A” y “B” es un ejemplo a favor de esa proposición. Un caso particular que sea “A” pero “no-B” es un *contraejemplo*. Una proposición universal y su contraejemplo son proposiciones inconsistentes, no pueden ser ambas verdaderas. Por lo tanto, si el contraejemplo es verdadero, la proposición universal es falsa. Utilizando las herramientas de la lógica moderna, podemos representar la forma de un contraejemplo de la siguiente manera:

$$(Ac \ . \ \text{no-}Bc)$$

La letra “c” simboliza un individuo. Esta forma de enunciado nos dice que ese individuo cumple con la propiedad “A” pero no con la propiedad “B”. Ahora bien, las definiciones son bicondicionales, y un bicondicional nos dice que “todo A es B” y que “todo B es A”. Por lo tanto, en el caso de las definiciones, será un contraejemplo también un caso de algo que sea “B” pero no “A”:

$$(\text{no-}Ac \ . \ Bc)$$

Pero acabamos de ver que las definiciones contienen un elemento adicional. Una definición no nos dice simplemente que “todos los A son B” y que “todos los B existentes son A”, sino que, *necesariamente*, “todos los A son B” y que “todos los B son A”. Lamentablemente, la lógica clásica es indiferente al valor modal de las proposiciones. De las proposiciones necesarias y de las inferencias que pueden involucrar este tipo de proposiciones se ocupa una rama de la lógica no clásica conocida como *lógica modal alética*. Esta puede ser entendida como una extensión de la lógica clásica y consiste en la adición de dos nuevos símbolos a los ya conocidos. Son los siguientes:

- $\Box p$ =def. p es necesario
- $\Diamond p$ =def. p es posible

Por ejemplo, la siguiente inferencia válida:

Dios existe necesariamente.
Luego, Dios existe,

se formaliza así:

$$\frac{\Box p}{p}$$

En realidad, los conceptos modales *necesario* y *posible* son interdefinibles:

$$\Diamond p = \neg \Box \neg p$$

$$\Box p = \neg \Diamond \neg p$$

A través de \Diamond podemos definir la noción de contingencia:

$$p \text{ es contingente} = \text{def. } (\Diamond p \cdot \Diamond \neg p)$$

Volvamos ahora al tema de la definición. Antes dijimos que una definición debía ser formalizada como un bicondicional. Acabamos de ver que esto es insuficiente. Una definición es un bicondicional que tiene carácter de verdad necesaria:

$$\Box (x) (Ax \longleftrightarrow Bx)$$

La postura según la cual el producto final de una definición es una proposición universal necesaria presenta varias dificultades. Un primer problema es cómo podemos estar seguros de que una vez que hemos excluido un contraejemplo mediante la adición de una nueva condición no aparecerá en el futuro algún nuevo contraejemplo que nos obligue a adicionar nuevas condiciones. El filósofo probablemente argumentará que se trata de un ideal que nunca podemos estar seguros de haber alcanzado, pero que las definiciones pueden ser aceptadas con carácter provisorio. Son aceptables mientras no aparezca un contraejemplo.

Consideremos la definición de *percepción*, uno de los conceptos centrales de la epistemología. ¿Qué significa que alguien percibe un objeto? Obviamente, la primera condición es que el sujeto tenga una experiencia visual, y la segunda es que exista en las inmediaciones del sujeto un objeto que corresponda a esa experiencia; de otro modo, no diríamos que está *viendo* el objeto, sino que lo

está recordando, imaginando o alucinando. Ver algo (y cualquier otro tipo de percepción) se compone indudablemente de esos dos tipos de elementos: uno subjetivo, experiencial, y uno objetivo, la presencia en el entorno del objeto correspondiente. ¿Son esas dos condiciones suficientes para afirmar que alguien ve un objeto? El filósofo Paul Grice mostró que no lo son (1985). Imaginemos que el objeto existe en el entorno del sujeto, pero que hay entre ambos una barrera opaca y que la experiencia del objeto es producida por un científico en el cerebro del sujeto a través de unos electrodos. No parece razonable decir que en ese caso el sujeto *ve* el objeto, aunque la situación satisfaga las dos condiciones recién mencionadas. Para excluir este caso de la clase de las percepciones, podríamos adicionar una condición en el *definiens*: la experiencia subjetiva debe ser causada por el objeto y por ninguna otra cosa. Pero supongamos que la imagen del objeto es captada por una cámara, y que esa imagen va al cerebro del sujeto mediante unos electrodos. Si bien el objeto es la causa de la experiencia, no es la causa inmediata. Podríamos agregar una nueva condición para excluir este caso: el objeto debe producir la experiencia de manera directa. Obtendríamos así la siguiente definición:

A percibe x =def. A tiene una experiencia cuyo contenido es el objeto x , y x existe en el entorno de A, y x es la causa directa de la experiencia de A.

¿Son estas condiciones suficientes? De acuerdo con la propuesta comentada en el párrafo anterior, podríamos considerarlas como suficientes en la medida en que no aparezca un contraejemplo que muestre lo contrario. Pero hay un problema adicional. Cuando buscamos la definición de un concepto, en general comenzamos por los casos más cotidianos y solemos arribar a una definición estándar que funciona bien para las situaciones ordinarias. Luego pasamos a casos que no se nos suelen presentar en contextos reales pero que son, con todo, concebibles (o, como dirían los filósofos, lógicamente posibles), y nos preguntamos si la definición estándar resiste las intuiciones acerca de esos casos extraordinarios. Si nuestras intuiciones contradicen la definición estándar, le agregamos una nueva condición a la definición para que se acomode a esas intuiciones. Preguntémonos: ¿tienen esas nuevas condiciones el carácter de una estipulación o de una descripción?; ¿somos nosotros, los filósofos, los que, partiendo del significado ordinario de *conocimiento* o *percepción*, hemos creado un nuevo concepto que es más preciso que el concepto estándar, o continuamos describiendo el significado corriente del concepto?; esas condiciones que han sido introducidas para responder a los casos extraordinarios, ¿son una creación o un descubrimiento?

Si uno opta por la opción descriptivista, debe estar dispuesto a ser un platonista (llamado también *realista*) respecto de las clases. Como señalamos en el inicio de este capítulo, el platonista es aquel que cree que ciertas clases no son convencionales, sino que existen objetivamente, sean reconocidas o no por los seres humanos. Si la clase *conocimiento*, por ejemplo, fuera una de esas clases, entonces estaría prefijado de antemano qué es conocimiento y qué no lo es. La tarea de adicionar condiciones sería descriptiva, si bien no de convenciones sociales, sino de cosas que existen por sí mismas. A los efectos de producir una posición coherente, el realista deberá argumentar que las intuiciones que tenemos respecto de los casos imaginarios surgen de una facultad racional especial que nos permite entrever qué condiciones hacen parte de esas clases subsistentes. Pero si uno no tiene inclinaciones platónicas y es más propenso a pensar que las clases son creaciones humanas, la posición descriptivista será poco atractiva. Pues si las clases son creaciones, no parece plausible suponer que hemos estipulado condiciones para dar cuenta de casos que nunca se nos han presentado en la experiencia y que tampoco creemos que se nos puedan presentar. Por ejemplo, la clase que denominamos *conocimiento* fue reconocida socialmente para decidir sobre casos ordinarios de creencias, y lo único que sabemos como hablantes de castellano es la manera en que esos casos deben ser categorizados. Sobre los casos extraordinarios no sabemos nada. No hay ninguna estipulación acerca de ellos porque la experiencia nunca nos dio ocasión de tener que decidir cómo categorizarlos. Por eso, quien tenga una visión convencionalista de las clases seguramente verá la tarea de adicionar las nuevas condiciones como un apartamiento de la tarea descriptiva. Considerará que se trata de una estipulación que tiene como punto de partida una definición descriptiva. De ser así, las intuiciones que tenemos acerca de los casos imaginarios no emanan del concepto ordinario. Son meras inclinaciones semánticas. Un desafío para esta posición es explicar la utilidad teórica o el interés de esas estipulaciones. ¿Para qué nos serviría, por ejemplo, disponer de una definición tan refinada de *conocimiento* si la definición estándar (más tosca) alcanza para discriminar los casos que se nos presentan cotidianamente? Por un lado, una opción para el convencionalista es aceptar que la búsqueda de definiciones que tengan el estatus de proposiciones necesarias es una tarea que carece de interés teórico, y, por otro lado, sostener que el objetivo de la filosofía debe ser elaborar definiciones que den cuenta de los casos que la experiencia empírica nos ofrece. Volveremos sobre este punto más adelante.

Reductio ad absurdum

En este apartado vamos a discutir sobre un método adicional para evaluar definiciones: *reductio ad absurdum*. Lo ilustraremos con un ejemplo tomado de la ética. Uno de los problemas centrales de la ética es determinar qué significa que una acción es moralmente buena, cuál es el común denominador de todas las acciones que llamamos buenas. Consideremos brevemente la respuesta que da el hedonismo para esta pregunta. El hedonismo es la doctrina según la cual lo único que tiene valor intrínseco es el placer. Todas las otras cosas son buenas en la medida en que sirven como medios para conseguirlo. Por el contrario, lo malo es lo que produce dolor, directamente o a través de sus consecuencias. El hedonismo no reivindica apenas el placer sensual (como el placer de comer o el placer sexual), sino, sobre todo, los placeres espirituales e intelectuales: el placer de ayudar a un amigo, de resolver un problema matemático, de leer un poema, de realizar una meditación budista. En general, los hedonistas consideran que el segundo tipo de placer es más importante que el primero. Pero ¿están los hedonistas ofreciendo una definición estipulativa o descriptiva de *bueno* al definir lo bueno como lo placentero o lo que conduce al placer? Cabrían las dos posibilidades. Algunos hedonistas podrían conceder que se trata de una estipulación. Otros podrían reivindicarla como una mera descripción de lo que el “hombre de la calle” denomina *bueno*. John Locke era este tipo de hedonista. Lo deja claro en el siguiente fragmento del *Ensayo del entendimiento humano*:

Eso que llamamos bueno es aquello que puede provocar o aumentar el placer, o bien disminuir el dolor en nosotros; o, también, lo que puede procurarnos o conservarnos la posesión de cualquier otro bien, o evitarnos un mal. Y, por el contrario, llamamos mal a lo que puede provocar o incrementar un dolor, o disminuir cualquier placer en nosotros; o bien a lo que nos produce cualquier mal o nos priva de un bien. Debe entenderse que por placer y dolor me refiero tanto al cuerpo como a la mente, según la distinción que comúnmente se establece, aunque en realidad no se trate sino de diferentes estados de la mente, provocados unas veces por desórdenes corporales, y otras, por pensamientos de la mente (2005, parte II, cap. xx, 2).

El nosotros de la frase inicial parece referirse a nosotros los hombres, y no a nosotros los hedonistas. La definición propuesta por Locke, que aquí entenderemos como una definición descriptiva, es la siguiente:

x es bueno =def. x nos produce placer, o es el medio para algo que nos produce placer, o es un medio para evitarnos un dolor.

Parece haber contraejemplos para esta definición. Por ejemplo, hay individuos que abandonan todas las comodidades del hogar, que abandonan a los amigos y a los familiares, y dedican su vida a ayudar a otras personas. Esos individuos parecen estar renunciando a la propia felicidad en aras de la felicidad de los otros, y no diríamos que lo que hacen es incorrecto. Si fuera verdad que obrar correctamente significa hacer aquello que nos proporciona placer, entonces lo que esos individuos hacen sería moralmente incorrecto. Como respuesta a esta objeción, el hedonista podría argumentar que, si juzgamos que el santo obra rectamente, es porque suponemos que es feliz al ayudar a los otros. Pero ¿es esto verdad? ¿Hacemos alguna suposición acerca de las vivencias de esos individuos cuando los tomamos como paradigma del obrar ético? De manera inversa, si fuera verdad que lo bueno equivale a lo placentero, diríamos que obra rectamente alguien que mata porque siente placer al hacerlo. Un contraejemplo más radical es ofrecido por Robert Nozick en un artículo titulado “La máquina de experiencias”:

Supongamos que hubiera una máquina de experiencias que fuera capaz de darnos cualquier experiencia que quisiéramos. Por medio de esta máquina, los neuropsicólogos serían capaces de estimular el cerebro para hacernos creer y sentir que estamos escribiendo una gran novela, o haciendo un amigo o leyendo un libro interesante. Todo el tiempo estaríamos flotando en un tanque con electrones sujetos a nuestro cerebro. ¿Sería correcto conectarnos de por vida a esta máquina, programando previamente todas nuestras experiencias? (1988: 44-45).

Si la respuesta que damos a esta pregunta es negativa, esto muestra que *bueno* no es sinónimo de *placentero*, y que la definición hedonista no describe correctamente el significado que ese concepto tiene en la moralidad ordinaria. Pero podríamos criticar también la definición hedonista tomando en cuenta algunas de sus consecuencias lógicas. Si fuera verdad que ser bueno equivale a ser placentero y que ser malo equivale a ser displacentero, uno nunca se equivocaría sobre lo que es bueno y lo que es malo, dado que es imposible equivocarse acerca de lo que es placentero y displacentero. Si ponemos este argumento en forma estándar, vemos que es un silogismo válido, conocido como *Cesare*:

- Todas las cosas buenas son placenteras.⁴
- Ninguna persona puede equivocarse acerca de si algo es placentero.
- Luego, ninguna persona puede equivocarse acerca de si algo es bueno.

Un argumento adicional en contra del hedonismo es que, si fuera verdad que todo lo doloroso es malo, nadie podría desear el mal, dado que nadie desea el dolor. Nuevamente, si colocamos este argumento en forma estándar, vemos que:

- Todo lo que es doloroso es malo.
- Ninguna persona desea lo que es doloroso.
- Luego, ninguna persona desea lo malo.

Vemos que es un argumento válido y, de hecho, también es un silogismo válido, conocido como *Celarent*. Lo que hemos hecho con ambos argumentos es reducir al absurdo la posición hedonista: a partir de la premisa hedonista inferimos una conclusión absurda. Dado que al hacerlo utilizamos una forma de razonamiento válida (en este caso, dos silogismos válidos), una de las premisas debe ser falsa (si un razonamiento es válido, es imposible que las premisas sean verdaderas y la conclusión falsa). Si la segunda premisa es indisputable, se sigue que la premisa falsa es la que enuncia la posición hedonista. Por lo tanto, el hedonismo queda refutado. Notemos la diferencia con el método de los casos. La reducción al absurdo opera a partir de un razonamiento válido. El método de los casos opera a partir del principio de no contradicción: una proposición universal es confrontada con una proposición particular (referida a un caso imaginario), que es contradictoria con ella. En el método de los casos, la verdad de la proposición particular es establecida mediante la intuición. En la reducción al absurdo se puede apelar a la intuición para determinar que la conclusión es absurda, pero, en algunos casos (como en los dos ejemplos anteriores), es a través de la experiencia empírica.

⁴ Antes dijimos que una definición es un bicondicional cuantificado $[(x) (Ax \leftrightarrow Bx)]$, lo cual equivale a decir que “todos los A son B y todos los B son A”. Para esta deducción basta tomar una de las partes del bicondicional.