

# Fallar es humano

## un análisis filosófico de la falibilidad

Axel Arturo Barceló Aspeitia

– borrador de abril 2017 / comentarios bienvenidos –

[abarcelo@filosoficas.unam.mx](mailto:abarcelo@filosoficas.unam.mx)

### I. Introducción

“No estoy convencido de que haya algo más que una etiqueta de moda [“normatividad”] que unifique de manera significativa esta gama de fenómenos que incluye la adhesión a los principios de la deducción, la toma de decisiones sobre cómo actuar, el adoptar creencias coherentes y útiles, el uso de palabras de acuerdo con sus significados, el pensamiento conceptual, el actuar de acuerdo a convenciones sociales, y el actuar como uno debe.”

Adam Morton (2006) Traducción mía.<sup>1</sup>

El objetivo de este libro es analizar la estructura de la falibilidad a través de diferentes campos normativos de interés filosófico y en especial, la epistemología, la ética y el lenguaje. Al terminar el libro, habremos construido lo que podríamos llamar una teoría general de la falibilidad dibujando las relaciones entre una serie de conceptos normativos fundamentales como los de suerte, virtud, responsabilidad, preparación, negligencia, juicio, y similares. Trabajaremos bajo la hipótesis de que esta estructura normativa es la misma en todos los campos normativos y lo ilustraremos con ejemplos provenientes de los tres campos ya mencionados y otros más. Si tenemos éxito, podremos dar cuenta de fenómenos tan aparentemente disímiles como como prepararse de manera responsable para un desastre natural hasta cómo comunicarse

---

<sup>1</sup>. “I am not convinced that there is anything more than a fashionable label in common [– “normativity” –] that makes a significant unity of [this] wide range [of phenomena] that includes adhering to principles of deduction, choosing actions, making beliefs coherent and serviceable, using words in accordance with their meanings, thinking conceptually, following social conventions, and acting as one ought.”

con lenguaje vago. Lo que no haremos será ofrecer una explicación de la normatividad en términos no normativos. Dudo que la normatividad no sea irreducible, pero esa no será una de las cuestiones que abordaremos. Otro monstruo que dejaremos dormir a nuestro lado será el del escepticismo. En otras palabras, no diremos nada sobre cómo “responder al reto escéptico”, lo que sea que eso signifique. Mas bien, trabajaremos bajo el supuesto de que el dislibro normativo tiene sentido, es decir, que podemos actuar de manera responsable o irresponsable, que el éxito y el fracaso son ambos posibles y que no es lo mismo hacer las cosas bien que mal. Muchos filósofos han presentado argumentos muy brillantes contra de estos supuestos, pero no los trataré de refutar aquí.

A fin de cuentas, defenderé un dualismo contextualista y materialista. Mi propuesta es un contextualismo en tanto sostiene que algunas de las nociones centrales de las que dependen nuestras evaluaciones normativas son sustancialmente sensibles al contexto. En particular, qué casos sean previsibles y cuales de ellos sean necesarios cubrir para actuar de manera responsable en una circunstancia dada dependerá de manera sustancial en qué relibros tengamos disponibles. Ppor ello, mi propuesta también es materialista (en el sentido Marxista presente en el trabajo de filosofas como Sally Haslanger, Iris Young o Ángeles Eraña, no en el sentido ontológico de rechazar la existencia de entidades no materiales).

Mi propuesta es un dualismo en tanto sostiene que tanto deontologistas como consecuencialistas tienen razón, pero cada uno de su dominio normativo propio e independiente. Del deontologismo recupero la tesis de que no podemos dar cuenta de todas nuestras prácticas normativas apelando sólo a nuestros fines, (y a nociones derivadas de ellas como las de confiabilidad, probabilidad, robustez, etc.) como sostienen los consecuencialistas. Sin embargo, mi dualismo recupera también las intuiciones consecuencialistas de que sería raro evaluar positivamente medios que no nos acercaran a nuestros fines (aunque también creo que aunque raros, estos casos sí existen, como sostienen los que llamaré deontologistas radicales). En otras palabras, creo que es razonable esperar de un agente virtuoso que actué bien y que sus actos no tengan consecuencias dañinas de la misma manera que es razonable esperar de quién no nos hace daño y actúa de manera impecable que sea una persona de carácter virtuoso, aun cuando el tener un carácter virtuoso no pueda reducirse a la tendencia confiable de actuar bien y no dañar a otros (y, de manera inversa, el actuar bien no pueda reducirse al actuar virtuoso). El que ambos aspectos normativos sean diferentes, aunque relacionados es lo que caracteriza mi postura dualista.

Sin embargo, también es uno de mis intereses centrales en este libro el presentar la estructura general de nuestras prácticas normativas y quiero dejar claro qué aspectos aceptan (o, aun mas, deberían aceptar) tanto deontologistas como consecuencialistas o relativistas, etc. Por ello, al mismo tiempo que desarrollo mi propia propuesta, señalaré los puntos y las direcciones en las cuales podrían desarrollarse propuestas alternativas. En otras palabras, trato de ser justo con las motivaciones y virtudes de otros tipos de dualismos y monismos, tanto deontologistas como consecuencialistas, tanto contextualistas como relativistas. Hago esto no sólo por razones de ética profesional – es mas honesto reconocer las fortalezas y razones de las propuestas alternativas que atacar muñecos de paja – sino también por razones propedeúicas. He podido llegar a mi propuesta principalmente porque las propuestas consecuencialistas y deontologistas actuales han alcanzado un grado de sofisticación superlativo. En otras palabras, aun cuando termino creyendo que tanto el deontologismo como el consecuencialismo están equivocados, creo que cada uno de ellos acierta a resaltar aspectos fundamentales de nuestras prácticas normativas que serían difíciles de apreciar bien fuera del tratamiento que le dan deontologistas y consecuencialistas.

El libro estará estructurado alrededor de seis conceptos centrales del fenómeno normativo:

## **I. Falibilidad**

En epistemología, se dice que el conocimiento es falible en tanto es posible que un agente conozca algo en base a una justificación que no garantiza la verdad de aquello que conoce. En base a esta caracterización, podemos definir un proceso como falible, en general, si su éxito no está garantizado, es decir, si efectuar el proceso de manera responsable e irreprochable no garantiza que se alcance el objetivo. El objetivo de este libro es explorar la estructura normativa de este tipo de procesos. Hoy en día, en filosofía, cuando se habla de falibilismo, se piensa automáticamente en la tesis epistemológica según la cual ninguna creencia humana está justificada de manera infalible. Sin embargo, me parece claro que la falibilidad epistémica es sólo un caso de un fenómeno mas general que puede darse en otros ámbitos. Que algo sea falible no significa mas que puede fallar y que algo sea infalible tampoco significa entonces más que

garantiza su objetivo de manera concluyente. Así pues cualquier actividad que tenga un fin u objetivo, puede ser falible – si es posible que no alcance su objetivo – o infalible – si lo garantiza.

### **A. Indeterminación, Riesgo y Normatividad**

Los procesos falibles generan un fenómeno muy interesante: dado que no garantizan su éxito, permiten que sea posible alcanzar el éxito, no en virtud de haber llevado a cabo el proceso de manera irreprochable, sino por suerte. En este segundo capítulo estudiaremos este fenómeno tal y como se manifiesta en diferentes espacios normativos, una vez más, poniendo especial atención al ámbito de lo epistemológico y de lo ético. En particular, identificaremos dos preguntas importantes que han dominado el debate alrededor de la fabilidad en estas dos disciplinas filosóficas: la pregunta descriptiva de ¿qué distingue un caso de éxito genuino de uno que se da solo por suerte? y la pregunta fundacional de ¿porqué preferimos los primeros sobre los otros. En este segundo capítulo veremos una serie de propuestas recientes según las cuales para que alcanzar un objetivo cuente como un verdadero logro es necesario que manifieste las virtudes relevantes del medio que se usó para alcanzar dichos logros.

### **B. Relativismo**

En esta cuarta sección abordamos la cuestión de ¿en qué sentido puede ser efectivamente irreprochable un proceso falible que no resulta en aquello que persigue? Esto significa, en el caso epistémico, preguntarse si realmente hay algo que podamos evaluar positivamente en un caso de creencia falsa justificada o, en ética, si podemos evaluar como moralmente correcto un acto que se realice de manera responsable pero tuvo consecuencias excesivamente dañinas para una víctima inocente. En particular, abordaremos la propuesta relativista según la cual nuestra evaluación de un proceso falible varía conforme vamos teniendo mayor o menor certeza de que alcanzara su éxito o no. En este sentido, lo correcto o incorrecto de un acto puede cambiar conforme va produciendo consecuencias positivas o negativas.

### **C. Dualismo**

En la quinta sección propondré otra respuesta a esta pregunta, según la cual es necesario distinguir dos diferentes tipos de normatividad – una dirigida hacia los fines y otra a los medios. Para ello, revisaremos evidencia empírica de que, efectivamente, contamos con dos sistemas cognitivos de

evaluación relativamente independientes los cuales, en casos normales, no generan resultados divergentes y que, por eso, los concebimos como operando sobre un solo dominio normativo.

#### **D. Presuposición y derrotabilidad**

En la última sección estudiaremos los presupuestos implícitos bajo los cuales esperamos que un proceso realizado de manera irreprochable deba resultar en éxito y cómo se relacionan con lo que en derecho (y en lógica y epistemología también) se han llamado *derrotadores*. Ilustraré la importancia de estos presupuestos mostrando como dan pie a algunas de nuestras paradojas más conocidas como la del prólogo o la de sorites.

## II. Falibilidad

La gente ... siempre se pregunta si sabías que iba a ser un fracaso. Dicen cosas como, "¿No lo sabían?" "¿Cómo podrían no haberlo sabido?" Mi experiencia es que no lo sabes. No lo sabes porque tienes mucho invertido y un interés personal [y] porque tienes esperanza.<sup>2</sup>

Nora Ephron (2010, 109-110)

### A. Infalibilismo y Falibilidad

La pregunta que guía esta reflexión es: ¿qué papel juega el mal, el error, la falla en un dislibro o práctica normativa? Aquí veo yo dos opciones: verla como un enemigo con el que hay que luchar a muerte, o tratar de incorporarla como algo de lo que nuestra teoría debe tomar en cuenta, con lo que nuestra práctica debe lidiar. La primera opción resulta en un infalibilismo en el que lo que se busca es desterrar y evitar en lo absoluto el mal, el error, la falla, etc. La segunda opción resulta en un dislibro y práctica normativa que le da a la falibilidad un lugar central. Como bien señala Dutant (2015), durante mucho tiempo, la filosofía occidental optó por la primera de estas alternativas, pero en años más reciente la falibilidad se ha instaurado como un aspecto ineludible de la normatividad. Lo que me interesa en este libro es estudiar precisamente las consecuencias de apelar a la falibilidad como manera de darle un lugar al mal, el error, la falla, el fracaso, la falsedad, el daño, etc. dentro de nuestro dislibro y práctica normativa. Para el infalibilista, lo que está mal, está mal y punto; no hay más. Para el no infalibilista, sin embargo, hay sentidos y circunstancias en las cuales parece que **está bien estar mal**, lo cual, por supuesto, suena paradójico; el objetivo de este libro es mostrar como no hay nada realmente paradójico aquí.

Para entender mejor el contraste entre un normatividad infalibilista y una que no lo es, pensemos en posibles ejemplos. El mas obvio es el epistemológico, donde el contraste se manifiesta entre una concepción de la epistemología en donde, para parafrasear el título completo del *Dislibro del Método* de Descartes (1637), conducir bien la razón no es mas que buscar la verdad. En otras palabras, el único objetivo es la verdad y lo que debemos de hacer es aquello que – de manera infalible – nos lleve a ello. En contraste, hoy en día, se

---

<sup>2</sup>. People... always wonder if you knew it was going to be a flop. They say things like, "Didn't they know?" "How could they not have known?" My experience is that you don't know. You don't know because you are invested [and] because you hope. Mi traducción.

piensa que esto, o bien es imposible, o es insuficiente para dar cuenta de lo que llamamos “conocimiento” y su relación con la verdad. Pero no quiero quedarme en el ejemplo epistemológico, por un lado porque ha sido ya muy elaborado, tanto en el texto de Durant como en otros lados, sino porque no quiero dar la impresión de que el fenómeno es eminentemente epistemológico – aunque diré más sobre esto más adelante.<sup>3</sup> Entonces pensemos en cualquier otro mal. Por ejemplo, durante una parte importante del siglo veinte se pensó que la lógica cumplir con su función normativa simplemente con señalar cuándo un argumento o inferencia era válido, es decir, cuando su conclusión se seguía lógicamente de sus premisas. Esto significaba que, frente a casos de argumentación o inferencia inválida, la lógica no tenía otra cosa que aportar más que decir eso, es decir, que son inválidos. En este sentido, su postura normativa era infalibilista. A finales del siglo pasado, varios filósofos y científicos cognitivos – provenientes predominantemente de la psicología cognitiva y de la inteligencia artificial, peor también de lo que ahora se llama la lógica informal –, propusieron adoptar una postura distinta que reconociera que no bastaba señalar que dichas inferencias o argumentos eran falaces, sino que era necesario ampliar el alcance del dislibro normativo sobre dichos argumentos para hacer distinciones más finas. En otras palabras, no aceptaban simplemente la aparente conclusión de que realizar dichas inferencias o apelar a dichos argumentos siempre era irracional, sino que a veces podría también ser racional – sin dejar de ser lógicamente inválidos. Nótese que lo que buscaban no era meramente explicar porqué a veces cometemos errores al razonar o argumentar, es decir, su proyecto no era meramente descriptivo, sino eminentemente normativo. Su objetivo fue ampliar el dislibro normativo sobre la argumentación y la inferencia para incorporar al error y distinguir entre errores racionales y otros genuinamente irracionales.

Para entender mejor el paso de una filosofía infalibilista a una falibilista, vale la pena revisar también el seminal trabajo de John L. Austin y, en especial, su conferencia “Un alegato en pro de las excusas” (1956, 1957). En esta seminal conferencia Austin señala que nuestras prácticas éticas no se reducen meramente a

---

<sup>3</sup>. Aunque tampoco quiero dar la impresión contraria, es decir, que el fenómeno no es epistemológico, o que es eminentemente ético o económico o evolutivo, etc. Por eso no me gusta hablar, como Bortolotti (2016) de “La Ética de la Creencia” como algo distinto de la epistemología. El que todos estos regímenes normativos tengan la misma estructura no hace que todos sean epistemología, o que todos sean ética, etc.

decir que ciertas cosas son buenas y otras malas, sino que son mucho más complejas: involucran justificaciones, excusas, perdón, etc. Y Austin nos invita a notar que lo único que se puede excusar es lo malo, es decir, los errores, las fallas. Las excusas tienen sentido precisamente porque sabemos que no basta señalar al mal y decir que es malo. Hay que hilar más fino y distinguir entre males justificados y no justificados, excusables e inexcusables, perdonables e imperdonables. Una vez más, la idea fundamental es que la normatividad no puede excluir al mal, a la falla, el error, sino que debe darle espacio (no como mero enemigo a evitar). En otras palabras, lo que señala Austin es que nuestras prácticas normativas morales no son infaliblistas, sino que reconocen nuestra falibilidad.

Desde una perspectiva infaliblista es difícil hacer sentido de nuestra práctica ordinaria de pedir disculpas u ofrecer excusas. Si pensamos que simplemente hay cosas buenas o malas, el que ofrece disculpas o bien hizo algo bueno o algo malo. Si hizo algo bueno, no necesita disculparse (a lo más necesitaría explicar porque lo que hizo es bueno); si hizo algo malo, tampoco tiene sentido que se disculpe, pues lo que hizo sigue siendo malo y nada que diga lo va a cambiar. De cualquier manera, no tendría sentido ofrecer disculpas. Sin embargo, aceptar que la práctica de pedir disculpas no tiene sentido es aceptar que una práctica normativa muy extendida no tiene sentido. Aceptar una conclusión tan fuerte debería requerir también evidencia muy fuerte. Es preferible buscar otra explicación que no nos presente a los humanos como sistemáticamente envueltos en una práctica sin sentido.

## B. Una Nueva Metodología para la Filosofía

Austin es conocido como uno de los principales proponentes de lo que hasta la fecha se conoce como la “Filosofía del Lenguaje Ordinario”. Una de las tesis principales de este movimiento filosófico era darle un lugar central en la reflexión filosófica a la práctica lingüística habitual de los hablantes comunes y en ese momento se presentaba como revolucionaria en frente a la tendencia, dentro de la filosofía analítica, de ‘reformular’ al lenguaje de manera artificial para que mejor reflejará las relaciones conceptuales de interés científico y filosófico. En otras palabras, frente a quienes pensaban que el lenguaje, aunque una buena herramienta para la comunicación ordinaria, era inapropiada para el análisis conceptual, Austin y sus colegas Ryle (1953) y Strawson (1959) proponían voltear a nuestro uso ordinario del lenguaje como guía, falible por supuesto, a esas relaciones conceptuales. Según Austin, ver porqué en nuestra práctica lingüística habitual, es

decir, en como hablamos y escribimos en nuestras lenguas naturales, hacemos ciertas distinciones – por ejemplo, entre “éxito” y “logro” – y relacionamos ciertas cosas – por ejemplo, todas las cosas para las que usamos el verbo “hacer” – nos permite ver distinciones y similitudes que de otra manera serían difíciles de identificar. Tal y como señala Austin:

“...nuestro surtido común de palabras encarna todas las distinciones que los hombres han considerado lo suficientemente importantes como para ser delineadas, y las conexiones más dignas de ser señaladas en la vida de muchas generaciones; éstas, seguramente, son más numerosas y sonoras desde que los hombres lograron imponerse contra la larga prueba de supervivencia del más apto; y más sutiles, por lo menos en los temas ordinarios y razonablemente prácticos, que los que tú o yo podemos pensar desde nuestro sillón en una tarde –el método alternativo más favorable.”

Este voltear a la manera como habitualmente usamos el lenguaje comparte con otro par de giros metodológicos en filosofía muy importantes del último siglo – el naturalismo y el giro pragmático – un interés por restituir el balance entre lo concreto y lo abstracto en filosofía. Llamamos “giro pragmático” a la tendencia, comúnmente asociada con el segundo Wittgenstein y el pragmatismo norteamericano, de estudiar los problemas filosóficos no ‘en abstracto’, sino en su uso práctico. Por ejemplo, en el caso de la normatividad, el consejo es dejar de preguntarnos qué es el bien, el mal, el deber, el valor, etc. en abstracto y empezar a pensar mas bien en cómo se manifiestan en lo que hacemos, es decir, poner las prácticas de evaluación al centro de la reflexión filosófica sobre la normatividad. Por el otro lado, el naturalismo puede caracterizarse como una reacción en contra de las excesivas idealizaciones que guiaron al análisis conceptual de principios del siglo XX y a favor de un renovado interés en voltear a ver las cosas en toda su complejidad. Para nuestros objetivos, adoptar una perspectiva naturalizada sobre las prácticas normativas significa reducir el número de idealizaciones que asumimos y voltear, de manera empírica, a nuestras prácticas normativas ordinarias, es decir, a cómo de hecho evaluamos. Al combinar el naturalismo con el giro pragmático, se abre una brecha metodológica por la cual podemos dejar a nuestras prácticas evaluativas ordinarias guiarnos hacia un mejor entendimiento de la normatividad.

Y aquí es muy importante señalar que cuando hablamos de nuestras prácticas evaluativas nos referimos a la manera en la que hacemos juicios evaluativos nosotros, aquí y ahora. Esto nos obliga también a voltear la cara hacia las disciplinas que mejor nos pueden servir para entender quienes somos estos nosotros

aquí y ahora que hacemos dichos juicios evaluativos. El seminal trabajo de John L. Austin aparece también aquí como un buen ejemplo. En él Austin ya apunta hacia dos fuentes importantes de información para pensar la normatividad: la psicología y las leyes. A lo largo de este texto yo también beberé de ambas fuentes. De la psicología será fundamental reconocer la importancia de los sesgos epistémicos sistemáticos de nuestras capacidades cognitivas y del derecho la importancia del concept de negligencia o descuido para asignar responsabilidad a agentes por las consecuencias negativas de sus actos. Pero también podríamos añadir a la lista otras disciplinas como la economía o la antropología.

Podríamos pensar a la psicología como la disciplina que nos ayudará a situar la evaluación dentro del marco general de la vida mental del sujeto y al derecho y la antropología como las que nos ayudarán a situarla en su marco social mas amplio. A fin de cuentas, la evaluación es algo que comúnmente hacemos los individuos, sobre individuos, pero también como colectividades y sobre colectividades; a veces en el foro interno de nuestras mentes y otras veces en el foro público. De ahí la relevancia de tomar en cuenta tanto lo que el derecho tiene que decir, como la psicología o la antropología.

De la psicología del razonamiento, es muy útil comprender el fenómeno de los sesgos epistémicos. Tal como lo define Claudia Lorena García (2004), “un procedimiento o mecanismo cognitivo de un determinado sujeto” muestra un sesgo epistémico cuando “hay un rango de condiciones o contextos en los cuales el procedimiento o mecanismo ... produciría sistemáticamente resultados cognitivos que serían epistémicamente inaceptables o injustificados, y otro rango de condiciones o contextos en donde el mecanismo en cuestión produciría en general resultados epistémicamente aceptables o justificados.” (García-Aguilar 2004, 82) Dada la sistematicidad de dichos errores, es difícil pensar que se deben a meros errores de ejecución (Eraña 2003), es decir, que nuestra mente tenga la capacidad de razonar siempre de manera lógica impecable y que, sin embargo, algunas veces, por distracción o descuido, nos equivocamos. Mas bien, lo que parece suceder es que nuestras capacidades cognitivas no están diseñadas para producir siempre dichos resultados lógicos. Es pues, necesario preguntarnos el porqué de dichos errores, es decir, si no es para producir siempre resultados lógicamente impecables, cuál es la función de nuestros mecanismos cognitivos. En otras palabras, es necesario, insisto, reintegrar el error lógico en nuestra teoría normativa del razonamiento y la cognición. No podemos simplemente decir que somos ilógicos. Tal y como lo explica Ángeles Eraña:

“A partir de la evidencia (empírica), los psicólogos cognoscitivos afirman que los seres humanos llevamos a cabo inferencia a través de una serie de atajos también llamados *reglas* o *procesos heurísticos* que se caracterizan por mostrar algunos sesgos en nuestro razonamiento [resultantes de ignorar aspectos relevantes del problema, o prestarle demasiada importancia a aspectos irrelevantes, que] deben atribuirse a la utilización de dichas reglas, esto es, a la utilización de sistemas de instrucciones que nos permiten solucionar un problema, no de manera exacta ni en todos los casos pero sí de manera que no [requiera] una computación excesiva desde el punto de vista de los recursos cognitivos que tenemos disponibles” (Eraña 2003)

Este mismo contraste aparecerá una y otra vez cuando consideremos ejemplos en otros campos normativos: el contraste entre un método (herramienta, procedimiento, mecanismo, acto, método, práctica, etc.) infalible pero muy caro (o hasta inaccesible) en términos de recursos (de tiempo, atención, esfuerzo, conocimiento, control, etc.) y otro más accesible (en el sentido de requerir de menos recursos para implementarse) pero falible. Si pensamos que la única manera de juzgar un método es preguntarnos si nos da los resultados correctos o no, entonces los métodos de primer tipo serían los únicos métodos “buenos”, mientras que una noción más amplia de la normatividad podría explicar porqué es preferible optar por métodos del segundo tipo, los cuales, aun cuando no siempre den los resultados correctos, siguen siendo “buenos” en algún sentido.

Tratar de domesticar este tipo de errores no significa que no se den de hecho errores de ejecución, es decir, que no cometamos errores por descuido sino que todos nuestros errores sean, en el fondo, racionales. Por supuesto que existe la negligencia, pero no todos los errores se deben a ella. Algunos errores se deben a la mala suerte, por ejemplo. Es por ello que en derecho se hacen tantas distinciones, entre crímenes y delitos, entre descuido y malicia, etc. No todos los errores son iguales y por eso no los tratamos igual. Algunos los dejamos ir, otros los señalamos; algunos se pueden perdonar tras una disculpa, mientras que otros son imperdonables; algunos involucran la suerte, otros la falta de preparación, etc. Una teoría adecuada de la normatividad, por lo tanto, debería de ser capaz de dar cuenta de esto y una teoría infabilista claramente no puede. En el resto del libro presentaremos una teoría general de la normatividad cosntruída ex-profeso para lidiar con este tipo de fenómenos, es decir, una teoría de la normatividad que le da un lugar central a la falibilidad.

De la economía tomamos la idea de que, una vez que nos proponemos algo, no hacemos todo lo que podemos hasta garantizar lograrlo. Si quiero un teléfono celular nuevo, no voy a pagar lo que sea por él, aun cuando cuente con el dinero. Si mi capital fuera infinito, sin lugar a dudas tal vez pagaría lo que fuera por él; pero no es así, mis relibros son limitados y como tales debo administrarlos. Esto significa determinar hasta cuanto sería responsable pagar por dicho teléfono, dadas mis otras necesidades, deseos, etc. En este libro no nos detendremos a estudiar todos los factores que afectan nuestra decisión de dedicarle ciertos relibros a ciertos fines, y no a otros. Lo único que es importante para nuestros objetivos es reconocer que nuestras decisiones y juicios dependen de nuestras capacidades y relibros limitados. Que el que no hagamos todo lo que está en nuestras manos para conseguir un fin no significa que no lo valoramos o que, en realidad, no es uno de nuestros fines. Para ilustrar la importancia filosófica de esta idea volvamos al ejemplo epistemológico y pensemos en una de las críticas medulares que, desde el pragmatismo (Peirce 1868), se ha hecho al infalibilismo de Descartes (1641): si tuviéramos relibros y capacidades ilimitados, tal vez un fundacionismo como el de Descartes tendría sentido, es decir, tendría sentido ir construyendo nuestro conocimiento a partir de cero basándonos sólo en certezas inamovibles que nos garanticen por completo la verdad de lo que creemos. Sin embargo, no somos agentes epistémico perfectos con relibros ilimitados, sino seres finitos que, además de buscar el conocimiento, tenemos otros fines en la vida; y a veces, para alcanzar estos otros fines – es decir, para vivir en este mundo – necesitamos tomar decisiones restringidas en el tiempo y con información escasa. Es decir, necesitamos decidir qué creer en estados de incertidumbre. Para dar cuenta de estas decisiones, para dar cuenta de estas creencias, una epistemología infalibilista no está bien preparada.

De manera similar, una concepción infalibilista de la racionalidad humana tampoco puede dar cuenta de las reglas heurísticas que de hecho subyacen a nuestra competencia cognitiva. Las adoptamos porque nos permiten ahorrar relibros, pero ello no significa que no valoremos la corrección lógica o epistémica, sólo significa que no estamos dispuestos a perseguirla a cualquier costo. Dado que tenemos otras necesidades y que vivimos con ciertas limitaciones, es útil contar con capacidades cognitivas de este tipo.

## C. Metas y Medios

Hace unos meses, tuve la oportunidad de escuchar a Miriam Schoenfield decir algo muy obvio, pero no por ello poco profundo, sobre la epistemología. Decía que, si uno valora la verdad, para la pregunta ¿qué debemos creer? había una respuesta simple y obvia: debemos creer lo que es verdad. Sin embargo, si la pregunta de ¿qué debemos creer? se respondiera tan fácilmente, los libros de epistemología serían mucho, mucho mas cortos. Para entender porqué, permítaseme poner otro ejemplo completamente análogo pero mas obvio y hasta gracioso. Supongamos que alguien presumiera saber la manera de hacerse rico en la lotería, es decir, qué debemos hacer para hacernos ricos jugando la lotería y le preguntaremos cual es esa manera. Supongamos entonces que su respuesta fuera “apostar siempre al número ganador”. En sentido estricto, lo que nos dice es cierto. Apostar siempre al número ganador es una manera infalible de hacerse rico jugando a la lotería. Efectivamente, si apostamos siempre al número ganador, haremos muchos dinero en la lotería. Sin embargo, es claro que eso no es lo que pedimos cuando pedimos por la manera de hacernos ricos en la lotería. Lo que pedimos es un *cómo*, es decir, algo como un método que nos guíe en qué hacer para ganar la lotería. Apostar al billete ganador es, mas bien, una definición de *lo que es* ganar la lotería. En otras palabras, lo que pedimos fue un medio, y lo que se nos dió fue (una redescipción de) la meta.

Lo mismo sucede con la pregunta ¿qué debemos creer? Responder que debemos creer la verdad, aun cuando es cierto, no nos da lo que pedimos; y esto es así porque no estamos meramente buscando una definición de creencia *buena*, una propiedad que distinga aquello que debemos creer de lo que no. Al igual que en el caso de la lo loteria, lo pedimos un *qué* sino un *cómo*: un método, un mecanismo, una manera de alcanzar nuestra meta. No pedimos una re-descripción de nuestras metas, lo que pedimos son los medios para alcanzarlas. Pedimos algo que nos permita guiar nuestro comportamiento y deliberación. Necesitamos una caracterización de qué hacer, qué creer, qué decidir, etc. que nos permita, a partir de nuestras capacidades y relibros reales, tomar deciciones, adquirir creencias, etc. apropiadas.

Comparemos dos posibles reglas epistémicas: “debemos creer cosas verdaderas”, “debemos creer sólo aquello de lo cual no podemos dudar”. Aun en una teoría epistémica en la que ambas fueran verdaderas, estos dos tipos de reglas deben jugar papeles distintos. Como señalé en el párrafo anterior, el primer enunciado puede servir para definir una de nuestras metas epistémicas, pero no sirve para guiar nuestra deliberación epistémica; esto se debe a que para aplicarla, deberíamos saber algo que no sabemos a la hora

de la deliberación. Al igual que la regla “para ganar, hay que apostar el número ganador”, es inservible como estrategia o método. Cuando lo que pedimos es una regla que nos pueda guiar en la acción, deliberación, etc. necesitamos reglas como la segunda, “debemos creer sólo aquello de lo cual no podemos dudar”, pues para aplicarla sólo necesitamos apelar a cosas que sí están a nuestra disposición, como nuestros propios estados mentales conscientes, que son accesibles a nuestra reflexión. En otras palabras, al momento de la deliberación, necesitamos reglas que apelen sólo a relibros y capacidades que tenemos de hecho disponibles, y no a cosas que trascienden nuestras capacidades y relibros.

Con esta idea en mente, permítaseme introducir una distinción entre dos maneras de hablar de lo que debemos de hacer (o, en general, de usar lenguaje normativo). Llamaré **trascendentes** a enunciados normativos como el ya mencionado “debemos creer cosas verdaderas”, para contrastarlos con enunciados normativos **instrumentales** como, por ejemplo, “debemos creer sólo aquello de lo cual no podemos dudar”. Aun cuando ambos tipos de enunciados nos dicen qué debemos de hacer, las diferencias entre ellos son sustanciales. Si bien los primeros nos sirven para establecer las **metas** de nuestro actuar, no sirven de mucho a la hora de determinar los **medios** para alcanzar dichos fines. Esto se debe a que el vocabulario que se usa para describir dichas metas refiere a



aspectos de la realidad que trascienden nuestras capacidades y/o están fuera de nuestro control – de ahí que haya decidido llamarlas “trascendentes”. La noción de ‘verdad’ nos trasciende de una manera que la de ‘duda’ no lo hace. Que una creencia que tengamos sea verdadera o no depende de manera sustancial en cosas que no están bajo nuestro control o requiere para su reconocimiento relibros o capacidades de las que no disponemos, mientras que la duda, en tanto carácter subjetivo de nuestro estado mental, es de fácil acceso directo, es decir, nos podemos dar fácilmente dar cuenta de ella.

Nótese que la distinción se aplica no sólo en el campo de la epistemología. Cuando juzgamos moralmente un acto por sus consecuencias, por ejemplo, usamos un criterio trascendente, pues las consecuencias de nuestros actos no están por completo bajo nuestro control ni tenemos los medios disponibles para preverlas por completo. Sin embargo, cuando las evaluamos por sus intenciones, o evaluamos el carácter de alguien por sus motivos, usamos criterios instrumentales, pues asumimos que

nuestras intenciones y motivos sí dependen de nosotros de una manera en que las consecuencias de nuestros actos no (Williams 1981).

Con esta distinción en mente podemos dar una caracterización más precisa de lo que queremos decir por “infalibilismo”. El infalibilismo es la tesis de que un enunciado normativo instrumental es verdadero sólo si garantiza el cumplimiento de una regla trascendente. En otras palabras, lo instrumental debe garantizar lo trascendente. Esta caracterización hace más precisa la idea de que, para el infalibilista, solamente son válidos aquellos medios que nos llevan de manera infalible al logro de nuestros objetivos. En contraste, el no infalibilista acepta que puede haber enunciados normativos instrumentales verdaderos pese a que no garantizan el cumplimiento de normas trascendentes.

Nótese, además, que esta noción de ‘disponible’ depende de factores muy contextuales. Como veremos más a detalle en la sección que hablaremos sobre *relativismo*, los relibros disponibles para una persona en un momento podrían ser muy distintos de los disponibles para otra o para la misma pero en otro momento. Conforme voy conociendo más sobre una situación más puedo prever cómo se desarrollará y actuar en consecuencia. Igualmente, los relibros disponibles para decidir cómo actuar comúnmente son muy distintos de los relibros disponibles para evaluar como se ha actuado, lo que hace que una caracterización que sea trascendente desde el punto de vista del agente pueda ser instrumental desde el punto de vista de quién evalúa la acción del agente. Lo mismo sucede, por ejemplo, en lógica, donde la cuestión de cómo caractericemos un buen argumento dependerá de si queremos una caracterización que nos permita construir nuestros propios argumentos o si lo que queremos es saber cómo evaluar los argumentos de otros (Morado 2000).

Quiero dejar intencionalmente impreciso qué entenderemos por ‘disponible’, ‘accesible’ y bajo nuestro ‘control’ en este primer acercamiento a la distinción por dos razones. En primer lugar, porque me parece que la distinción es lo suficientemente clara por el momento y creo que a lo largo del libro se irán aclarando estas nociones en su uso; en segundo lugar, porque, pese a que en los ejemplos paradigmáticos es suficientemente claro cuando una caracterización es instrumental y cuando es trascendente, es un debate profundo dónde dibujaremos la línea entre los relibros accesibles y los inaccesibles, entre aquello que cae bajo nuestro control y qué no. En meta-ética, por ejemplo, Fernando Rudy (manuscrito) ha recientemente propuesto una teoría de las capacidades que se opone a la más tradicional propuesta volicionista y el punto

central del debate es precisamente qué aspectos se encuentran bajo el control de agente en el sentido relevante para hacer evaluaciones morales: aquello que hace o puede dejar de hacer de manera voluntaria (como sostienen los volicionistas) o aquello para lo que tiene alguna capacidad (como sostiene Rudy). Quisiera permanecer neutro respecto a este tipo de debates y abierto a la posibilidad de que disponibilidad, control y acceso signifiquen diferentes cosas para casos distintos, de tal manera que para ciertos debates puede ser necesario entenderlos en un sentido muy restringido y en otros en uno más amplio. En particular, quiero trabajar con un sentido de ‘disponibilidad’, ‘control’ y ‘acceso’ que nos permita aplicarlos en casos en los que sería poco natural hablar de un **alguien** que tiene el control o acceso a dichos relibros, por ejemplo, cuando hablamos de procesos colectivos o naturales. Para ello, por supuesto, sería necesario extender también las nociones de ‘meta’ y ‘medio’.

Tradicionalmente (por ejemplo, en Ove, Edvarsson y Cantwell 2016) se presentan a las *metas* como teniendo como función típica la de

“...regular (dirigir y motivar) la acción de manera que promueva el logro de dicha meta. Bajo circunstancias normales, tener un objetivo significa que uno hará algo para lograrlo. Un agente que ha adoptado una meta realizará típicamente acciones que crea podrían acercarla a su meta y evitar las acciones que crea tendrán el efecto contrario (véase Nozick 1993 sobre principios).” (Ove, Edvarsson y Cantwell 2016, 492. Traducción mía)

En otras palabras, explicamos porqué la gente hace lo que hace asumiendo que muchas de nuestras acciones y decisiones tienen sentido y dirección, es decir, que apuntan hacia algo como meta u objetivo. Y esta presunción es lo que da sentido a nuestras prácticas normativas, es decir, precisamente porque pensamos que muchas veces hacemos lo que hacemos para lograr algo que tiene sentido evaluar dichas acciones respecto a su relación con la meta hacia la que están dirigidas.

Sin embargo, en este libro quiero extender la noción de meta de este, su sentido original, a uno más amplio que incluya, no sólo acciones sino también otro tipo de estados, mecanismos, herramientas, convenciones, instituciones, etc. no personales. ¿Qué quiero decir con esto? Que, en tanto hay ámbitos normativos en los cuales lo que estamos evaluando no son acciones intencionales o cosas similares, sino ámbitos en los cuales no tiene sentido siquiera hablar de un *alguien* que esté motivado y busque alcanzar dicha meta, y que, sin embargo, muestran fenómenos estructuralmente similares a aquellos en los cuales sí está

involucrada una persona, vale la pena eliminar esta condición de la concepción tradicional de “meta” y trabajar con una noción mas amplia. Por supuesto, no quiero hacer revisionismo teórico o señalar que los filósofos anteriores estaban equivocados al sostener que las metas tienen una función motivacional o que están ligadas a los deseos, intenciones y otros estados mentales de una persona. Claro que no, estos elementos personales juegan un papel importante para explicar ciertos aspectos de la acción humana; pero para los objetivos de este libro, a saber, determinar la estructura de nuestras prácticas evaluativas.

Tampoco quiero meterme en cuestiones del estatus metafísico de las metas. Es un problema filosófico importante el determinar si ciertas normas o valores son constitutivos de ciertas acciones, prácticas, etc. o si para que algo sea una meta es necesario que se actúe con la intención de alcanzarla. Pongamos un ejemplo sencillo. Supongamos que alguien decide entrar a jugar golf con el objetivo de perder el juego e intencionalmente lanza la pelota lejos del hoyo en turno. ¿Cuál es la meta relevante para evaluar este acto? Al no meter la pelota al hoyo, ¿ha alcanzado el jugador su meta o no? Uno podría pensar que, en tanto el jugador no tenía la intención de ganar el juego, la meta de su acción era **no** meter la pelota. Sin embargo, también podría argumentarse que, aunque la intención era no meter la pelota, no es inadecuado decir que al no meter la pelota se ha fallado en el objetivo del *juego*. A fin de cuentas, el objetivo del juego de golf es autónomo y relativamente independiente de las intenciones individuales de quienes lo juegan.<sup>4</sup> En otras palabras, para que podamos decir que lo que hace la persona es jugar golf, tenemos que aceptar como apropiado evaluar su actuar conforme a las reglas de dicho juego, incluyendo las que definen su meta. La meta de meter la pelota en el hoyo es constitutiva del jugar golf, de tal manera que quien juega golf acepta de manera tácita esta meta, aunque sus intenciones vayan en dirección contraria.

La idea de que ciertas normas y metas son constitutivas de ciertas prácticas y que, por lo tanto, son apropiadas para evaluar las acciones que se realizan dentro de dichas prácticas, independientemente de las intenciones de quienes las realizan, es una idea de profundas raíces Kantianas que ha sido revivida en años recientes (por ejemplo, por Sally Haslanger). En filosofía del lenguaje, por ejemplo, autores como Alston (2000), Brandom (1994), Douven (2006), Searle (1969, 2001), Sellars (1963), y Williamson (2000) han

---

<sup>4</sup>. Digo ‘relativamente’ independiente, porque en el fondo las reglas que de hecho tiene el juego también están fundadas, en última instancia, en las intenciones de quienes lo juegan en el sentido de que si los jugadores quisieran jugar el juego de otra manera, podrían convenir en otras, nuevas reglas.

sostenido que “la aserción es análoga a un juego, en el sentido de que implica intrínsecamente ciertas reglas” (Rescorla 2009, p. 99. Mi traducción)<sup>5</sup> y, en particular, que tiene entre sus reglas constitutivas el tener como meta el conocimiento (o la creencia, o la verdad, etc.). Esto significa que si afirmamos algo que no sabemos (o que no creemos, o que no es verdadero, etc.) habremos hecho algo mal, independientemente de nuestras intenciones. Alguien quien quiera decir una mentira habrá hecho algo malo si nos miente, aun si su intención era mentirnos. En meta-ética, algunos filósofos contemporáneos han defendido la posición meta-ética Kantiana de que por lo menos algunas de nuestras reglas morales genuinas son constitutivas de la acción humana libre (Korsgaard 2000), mientras que por otro lado, filósofos como Velleman (2000) y Steglich-Petersen (2006) han defendido que es constitutivo de la creencia el considerar su contenido como verdadero sólo si es verdadero.<sup>6</sup>

En años recientes, esta idea de que hay valores constitutivos asociados a ciertos tipos de actos, independientemente de las intenciones de quienes los realizan, ha adquirido también relevancia en ciertos debates en ética aplicada. Rae Langton y Caroline West (1999), por ejemplo, han argumentado que al consumir pornografía, independientemente de cuales sean nuestras intenciones, participamos en una práctica que tiene como meta subordinar a la mujer; mientras que Robin Zheng y Nils-Hennes Stear (manuscrito) han extendido el argumento de Langton para defender que el pintarse la cara de negro, en el contexto actual en los Estados Unidos, es participar en una práctica de opresión hacia la comunidad afro-americana, aun si nuestras intenciones son inocentes y no tenemos ninguna animadversión hacia dicha comunidad. Si bien estas discusiones tienen clara relevancia filosófica, son irrelevantes para mis propósitos, pues mi teoría se aplica tanto en los casos en los que las normas son constitutivas como en las que no, e incluso a casos en los cuales la evaluación es completamente externa a aquello que se evalúa. por ejemplo, no tenemos que dirimir

---

<sup>5</sup>. “On this view, assertion is analogous to a game, in that it intrinsically involves certain rules. Someone who exhibits no sensitivity to the rules of a game does not grasp what it is to play the game, and someone who exhibits no sensitivity to assertion’s constitutive norms does not grasp what it is to assert a proposition.”

<sup>6</sup>. Nótese que aquí no estoy haciendo una diferencia entre lo que Steglich-Petersen (2006) llama los sentidos “teleológico” y “normativo” de las normas constitutivas relevantes. Aunque la distinción es importante, no es relevante para los objetivos de este trabajo.

el debate de si el valor adecuado para evaluar un platillo de cocina es su sabor o su valor nutricional para reconocer que existen prácticas normativas asociadas a ambos valores y que, en cada una de ellas, tiene sentido distinguir entre metas y medios, enunciados normativos trascendentes e instrumentales, relibros disponibles para la persecución de dicho valor, etc.

## D. Sesgos Cognitivos

Permítaseme entonces explicar cómo funciona este sentido extendido de ‘meta’ (y las nociones relacionadas de ‘medio’, ‘disponibilidad’, ‘acceso’, ‘control’, etc.) desarrollando un ejemplo que ya había mencionado anteriormente: los sesgos epistémicos:

Cuando recién ingresé al Instituto de Investigaciones Filosóficas, yo venía de hacer lógica, es decir, me consideraba un lógico y en mi posgrado me había dedicado principalmente a la lógica (y la filosofía de las matemáticas). En ese entonces mi concepción de la racionalidad era que ésta consistía en razonar de manera lógica, es decir, según las reglas que la lógica buscaba describir y que, en gran parte, había descrito ya con éxito. Por ejemplo, esto significa que si uno quiere determinar si una proposición de la forma *Para-toda-x-(P(x)-implica-Q(x))* es verdadera o falsa, es necesario y suficiente eliminar la posibilidad de un contraejemplo *a* que sea *P* pero no *Q*. Por ejemplo, si queremos saber si en un bar se respeta la regla de que sólo los mayores de edad beben alcohol (que es una regla de la forma *Para-toda-x-(P(x)-implica-Q(x))* donde *P* es la propiedad de beber alcohol y *Q* la propiedad de ser mayor de edad), sabemos fácilmente que hay que checar a los menores de edad para ver si están bebiendo alcohol o a los que beben alcohol para ver si son mayores de edad o no. A decir verdad, la evidencia empírica muestra que casi cualquier persona puede efectuar esta simple inferencia que manifiesta su capacidad racional. Desafortunadamente para esta concepción de la racionalidad, también es cierto que si le pedimos a la gente que realice otras inferencias *con exactamente la misma forma*, no nos dan la respuesta correcta. Por ejemplo, si ponemos frente a la gente cuatro tarjetas, cuyas caras visibles de las cartas muestran, las primeras dos los números 3 y 8, y las otras dos las letras R y A y les decimos que cada una de ellas tiene un número de un lado y una letra del otro. Al preguntarles qué cartas deberían dar vuelta para comprobar si es cierto que si una carta muestra un número par por un lado, entonces la cara opuesta muestra una vocal, la gente comúnmente responde que se deberían voltear la carta con el número 8 y la carta con la letra E. Sin embargo, un poco de atención nos permite dar cuenta de que la

respuesta correcta (si entendemos al condicional como expresando una implicación material) es que se deberían voltear la carta con el número  $\mathcal{8}$  y la carta con la letra  $R$ .

Lo sorprendente de este ejemplo es que este problema tiene exactamente la misma forma lógica que el anterior y, sin embargo, mientras que el primero lo resuelve la gente de manera correcta sin muchos problemas, este segundo casi nadie lo resuelve de manera correcta aun después de mucha reflexión. Además, otro punto clave a notar era que este error no era una excepción rara que podríamos explicar apelando a un descuido o falta de pericia lógica, pues el experimento da prácticamente los mismos resultados cuando se aplica a estudiantes de lógica que cuando se aplica a otras personas. Pero lo más importante, es que era un *error* que se reproducía de manera sistemática, lo que mostraba que era natural y, posiblemente constitutivo de nuestra forma de razonar. Pero, ¿qué debemos concluir de esto? ¿Qué la gente es irracional porque no puede ni siquiera hacer una operación tan simple y básica como detectar las condiciones de verdad de un condicional? ¿O que más bien la racionalidad no tiene que ver con reglas lógicas sino con otras cosas que más bien sólo la psicología nos puede revelar?

Obviamente, ambas respuestas son inaceptables. Como buen estudiante de Frege, sabía muy bien que la validez lógica era independiente de cómo de hecho razonamos y que ningún experimento psicológico podría convertir una inferencia no válida en una válida, por más que mostrara que todo mundo infiere así. Pero por otro lado, sabía también que una teoría de la racionalidad que tuviera como consecuencia que casi todos somos irracionales tampoco podía ser una teoría correcta. Era necesario, pues entender mejor la relación correcta entre la validez lógica y la racionalidad humana.

Es aquí que la distinción entre reglas trascendentes e instrumentales nos sirve para entender la relación entre validez lógica y racionalidad. En particular, podemos concebir a las reglas lógicas formales como reglas trascendentes de la racionalidad y las que de hecho guían nuestro razonamiento como instrumentales. No es necesario entrar en detalles, pero ahora entiendo que lo que sucede en estos casos es que este tipo de “errores” lo que muestran es que nuestras capacidades cognitivas no están diseñadas para siempre producir validez formal, pues hacer esto requiere ejecutar operaciones muy complejas (por ejemplo, detectar la forma lógica de la información que se nos presenta) que, además, nos tomaría demasiado tiempo. En la mayor parte de los casos, y en los más relevantes para nuestra supervivencia, no tenemos el tiempo para realizar dichas operaciones y debemos distribuir nuestros recursos cognitivos (memoria, atención, etc.) entre

muchas tareas cognitivas, de tal manera que comúnmente no podemos dedicar demasiados recursos a una sola inferencia. Por eso, nuestros mecanismos están diseñados para procesar información de manera eficiente, es decir, rápida y usando recursos cognitivos mínimos a costo de sacrificar cierta efectividad. En otras palabras, aun cuando en principio podríamos actuar y tomar decisiones siempre de manera lógicamente correcta, los recursos necesarios para hacerlo no nos están disponibles y por eso nos sirve de mucho contar con mecanismos cognitivos que nos permiten realizar inferencias que, no siempre son lógicamente válidas, pero son lo suficientemente certeras en suficientes casos importantes, a partir de recursos que sí tenemos accesibles. A esto me refiero cuando digo que las reglas de la lógica formal son trascendentes, mientras que nuestros mecanismos cognitivos siguen más bien reglas heurísticas que son instrumentales, pues están diseñadas para optimizar nuestros recursos cognitivos reales. Una vez más, en definir que entendemos por ‘disponible’ es importante tomar en cuenta el aspecto económico que mencionamos en la sección anterior. Cuando nuestros recursos son limitados (y nuestros recursos cognitivos son, de hecho, limitados) hay ventajas importantes (en este caso, evolutivas) en contar con procesos y mecanismos menos demandantes. Esto significa que, aun cuando en un sentido absoluto tenemos disponibles los recursos necesarios para realizar, por ejemplo, inferencias lógicamente impecables (y por eso es que podemos hacer, por ejemplo, pruebas lógicas absolutamente válidas), no siempre podemos darnos ese lujo y en este sentido dichos recursos no deben contarse como realmente disponibles en el sentido relevante.

De esta manera podemos dar cuenta de lo que sucede en el caos cognitivo de una manera que evite ambos cuernos del dilema. Por un lado, reconocemos la validez lógica como valor con fuerza normativa a la hora de evaluar nuestras inferencias, pero la restringimos a casos en los que hay pocas restricciones de acceso y control, es decir, cuando está justificado dedicarle los muchos recursos de tiempo, control y atención que se necesitan para ejecutarla. Por el otro, evitamos también la conclusión apresurada de que somos constitutivamente irracionales al desarrollar una concepción de la racionalidad acotada, donde por lo menos algunos de nuestros errores lógicos no se deben a fallas de ejecución o al descuido, sino que son plenamente racionales.

En este caso se ve claro cómo nos sirve extender el lenguaje de ‘disponibilidad’, ‘acceso’, ‘metas’ y ‘medios’ a casos donde literalmente no hay una persona de la que decimos adopta un medio o un fin, de la que decimos tiene disponibles ciertos recursos limitados que debe administrar. No quiero comprometerme a

que esta manera de usar estos términos sea metafórica o no, siempre y cuando se reconozcan las fuertes semejanzas entre el caso personal y el caso no-personal, de tal manera que las conclusiones que alcance de ahora en adelante puedan aplicarse tanto a unos casos como a los otros.

## E. Falibilidad mas allá de la epistemología

Hoy en día, en filosofía, cuando se habla de falibilismo, se piensa automáticamente en la tesis epistemológica según la cual, “ninguna creencia (teoría, concepción, tesis, etc.) puede nunca tener un apoyo o justificación racional concluyente [es decir,] nunca podremos excluir por completo el poder dudar de su verdad.” (Hetherington 2017. Mi traducción) Esta tesis a veces suele formularse apelando a la noción misma de falibilidad así: “Toda creencia está justificada de manera falible” es decir, si una creencia está justificada, esta justificación es falible, es decir, puede fallar. Así formulado el falibilismo, queda claro que la falibilidad no es una propiedad de la creencia en sí misma, sino de su justificación o apoyo racional. Lo contrario de lo falible, entonces, es lo infalible. Una creencia estaría justificada de manera infalible si su justificación por sí misma garantizara por completo su verdad. Durante mucho tiempo se creía que, por lo menos para cierto tipo de creencias podría lograrse una justificación de este tipo. Para Descartes, por ejemplo, nuestras ideas claras y distintas estaban justificadas de manera infalible, es decir, su propia claridad y distinción garantizaban por completo su verdad.

Espero sea claro, a estas alturas del libro, que la falibilidad de la justificación de nuestras creencias es sólo un caso de un fenómeno mas general que puede darse en otros ámbitos además del epistemológico. Que algo sea falible no significa mas que puede fallar y que algo sea infalible tampoco significa entonces más que garantiza su objetivo de manera concluyente. Así pues cualquier actividad que tenga un fin u objetivo, puede ser falible – si es posible que no alcance su objetivo – o infalible – si lo garantiza.

Pongamos un ejemplo de interés filosófico: la posección de conceptos. Podemos entender el debate sobre si es necesario para poseer un concepto dado el ser capaz de entenderlo por completo o no como un debate sobre la falibilidad de la posección de conceptos. En otras palabras, si la posección de conceptos fuera infalible, quien posee un concepto deber ser siempre capaz de evaluar de manera acertada, dada suficiente información sobre un objeto cualquiera, si cae o no dentro de su extensión, y ser capaz de efectuar de manera competente las inferencias necesarias que definen las relaciones lógicas del concepto – por ejemplo,

inferir de que algo es un buey a que no es una vaca hembra o de que algo es un lobato a que es un lobo joven. En contraste, si la posesión de un concepto fuera falible, entonces sería posible poseer un concepto y fallar al aplicarlo, ya sea en inferencia o clasificación. (Earl 2017)

Sin embargo, alguien podría pensar que no es sorprendente que tenga sentido hablar de falibilidad aquí, pues el fenómeno de la posesión de conceptos sigue siendo un fenómeno epistemológico (Earl 2017). Sin embargo, creo que debe ser importante dejar claro que no siempre que hablamos de falibilidad lo hacemos en un plano epistemológico. Para ello, sería útil tratar de encontrar un ejemplo no epistemológico. Para lograr esto, es más fácil empezar fuera de la filosofía. Consideremos el siguiente caso de la vida real:

Hace unos días, un camión de gran peso embistió un mercado navideño en Berlín, asesinando a 12 personas e hiriendo a casi 50. Meses antes, otro camión de gran peso fue usado en la ciudad de Niza para un ataque similar que resultó en 86 personas fallecidas. En condición de anonimidad, unos días después del ataque en Berlín, un alto funcionario de fuerzas antiterroristas declaró a la *Agence France-Presse* que no existían medidas infalibles contra este tipo de ataques, aunque era posible tomar precauciones como prohibir la circulación de automóviles y colocar bloques de hormigón alrededor de las calles donde se esperan se junten multitudes. En otras palabras, aunque las medidas podrían evitar ciertos ataques, no pueden garantizar que este tipo de ataques dejen de suceder. "Desgraciadamente, habrá que acostumbrarse a vivir bajo amenaza" declaró a la *AFP* Ludovic Guérineau, ex miembro de los servicios secretos franceses y ahora director de operaciones de la firma de seguridad privada Anticip.

En este caso, podemos hablar de medidas falibles o infalibles, no porque haya una dimensión epistemológica, sino porque hay un objetivo y este puede estar garantizado por la medida o no. Por supuesto, alguien podría re-formular el ejemplo de manera epistemológica, por ejemplo, diciendo que lo que hace a las medidas sugeridas a la *AFP* falibles es que, dado que *no podemos saber* cuando o donde habrá un intento de ataque con camiones, no podemos evitar por completo, o que el tomar dichas medidas *no nos permite saber con certeza* que no habrá ataques con camión; sin embargo, espero sea claro que el lenguaje epistémico en esta formulación del fenómeno es completamente artificiosa y no añade nada a decir simplemente que las medidas serían adecuadas aun si no eliminaran por completo la amenaza de un ataque de ese tipo. Lo que hace a una medida, técnica, proceso, etc. falible o no no es el de si, al aplicarlo, sabremos o no que logrará su objetivo, sino que si dicho objetivo está o no garantizado. Si tengo razón, entonces, todo medio puede ser

falible o infalible dependiendo de si garantiza o no su fin. Así es como voy a entender la falibilidad de aquí en adelante. [Agradezco a Ricardo Mena el presionarme en la importancia de este punto.]

En mi trabajo de (Barceló 2015) presenté otro ejemplo de relación filosófica importante que es útil pensemos en términos de falibilidad en vez de infalibilidad: la relación entre una verdad y su base metafísica, es decir, aquello que la hace verdadera. Mi ejemplo ahí era la verdad de que mis sobrinos son traviesos, lo cual es verdadero porque mis sobrinos, Aimee y Balam, de hecho son traviesos. Sin embargo, la relación es falible – o derrotable que es otra manera de decir lo mismo, como veremos mas adelante – porque el que mis sobrinos Aimee y Balam sean traviesos no garantiza por completo que mis sobrinos sean traviesos, ya que sigue existiendo la posibilidad de que tuviera otro sobrino que no fuera travieso. En ese caso, seguiría siendo el caso que mis sobrinos Aimee y Balam sean traviesos y, sin embargo, sería falso que mis sobrinos son traviesos.

Podría parecer muy extraño el que use el ejemplo de la relación entre hechos y verdades aquí dado que he dicho desde el principio que lo que busco en este libro es presentar una teoría sobre la estructura de nuestras prácticas normativas y no hay nada normativo sobre esta relación. Lo acepto. Sin embargo, como quedará cada vez mas claro, especialmente en la última sección de este libro, mi meta es aun más ambiciosa y abarca cualquier tipo de relación que se da entre lo determinado y lo indeterminado y que, por lo tanto, es falible. Esto significa que se aplica también a relaciones como la de causalidad, inferencia, consecuencia, fundamentalidad metafísica, referencia, etc. Los casos normativos son paradigmáticos, pero no son los únicos.

Si bien la infalibilidad es una virtud extrema de las herramientas, reglas, convenciones, métodos, procesos, mecanismos, esquemas de comportamiento, medios, etc.<sup>7</sup> esto no significa que las herramientas

---

<sup>7</sup>. En general, para ser mas conciso a lo largo del libro hablo simplemente de “medios” para referirme a cualquier miembro de este heterogéneo grupo. En este sentido, un medio es cualquier cosa cuyo comportamiento tratamos de explicar teleológicamente, es decir, apelando a algún fin que de sentido a su comportamiento.

(métodos, medios, etc.) falibles sean inútiles.<sup>8</sup> Es posible que una herramienta sea falible y aun así sea útil y su existencia valga la pena. Aun cuando las medidas sugeridas a la AFP arriba no garanticen detener los ataques a civiles por medio de camiones, esto no significa que no deban de adoptarse. Pensar de otra manera es caer en lo que se llama la falacia de confundir lo bueno con lo perfecto. Una medida puede ser buena, aunque no sea perfecta, es decir, infalible. Pero uno cae en la misma falacia si piensa que todo método falible vale la pena. Cuando es racional o no usar un método falible es una cuestión fundamental para entender la falibilidad y de ello hablaremos en el siguiente capítulo. En él veremos cómo tomarse en serio el fenómeno de la suerte requiere una transformación sustancial de nuestra concepción del error, en general, y del error lingüístico en particular.

---

<sup>8</sup>. Cuando nos preguntamos por la utilidad de una herramienta, es importante distinguir entre tres preguntas relacionadas pero sustancialmente distintas:

1. ¿Por qué es útil que **exista** la herramienta – es decir, no sólo por qué fue inventada, sino por qué ha seguido siendo tan exitosa como para haber sobrevivido?
2. ¿Por qué es útil **poseer** una herramienta?
3. ¿Por qué es útil **usar** una herramienta en una circunstancia dada?

Continuando con el ejemplo del martillo, no es lo mismo preguntarse por qué existen los martillos, que preguntarse por qué tenemos uno en nuestra casa o por qué lo usamos para clavar algo en una ocasión dada.

### III. Riesgo y Suerte

A mediados del siglo pasado, el re-descubrimiento de dos fenómenos importantes en el campo de la epistemología y la ética obligaron a los filósofos interesados en estos campos normativos a tomarse en serio el tema de la suerte. El primero de ellos es lo que ahora conocemos como los “casos Gettier” en honor al filósofo que los presentó, de manera muy elegante y sugerente, en un breve artículo de título “¿Es el conocimiento creencia verdadera justificada?” (Gettier 1963), el segundo es el fenómeno de la suerte moral, introducido a la meta-ética por Thomas Nagel (1979) y Bernard Williams (1981). Aún cuando, en realidad, ambos fenómenos no son sino casos particulares de un fenómeno más general, nos acercaremos al fenómeno primeramente a través del argumento de Gettier en la siguiente sección.

#### A. ¿Cuál es la moral del argumento de Gettier?

Como habíamos señalado en sesiones anteriores, una de las funciones centrales de los dislibros normativos es guiarnos en nuestra toma de decisiones y deliberación, tanto privada como pública. Nos involucramos en prácticas normativas precisamente en aquellos entornos en que hay algo en juego, aquellos en los cuales no solamente hay algo que ganar, sino también algo que perder. Esto hace que los dislibros normativos cobren un papel importante en circunstancias en las que hay un riesgo, y lo que esperamos de un medio adecuado es una **reducción sustancial de dicho riesgo**. El actuar responsable es el hacer lo suficiente, dadas nuestras capacidades y relibros, para reducir riesgos para nosotros mismos y para aquellos que puedan ser afectados por nuestros actos. En otras palabras, nos preguntamos qué hacer precisamente en aquellas circunstancias en las que **no da lo mismo** qué hagamos, sino que cómo actuemos puede tener consecuencias de gravedad tal que justifiquen le dediquemos cuidado a nuestras decisiones. Existe la normatividad epistémica, por ejemplo, porque no nos es indiferente qué pensemos, es decir, qué creamos, qué inferamos, qué dudemos, etc. Pensar de manera incoherente, no aprovechar nuestra información para tener una imagen adecuada de nuestra realidad, etc. son cosas que pueden tener consecuencias negativas, no sólo para nuestra supervivencia como individuos, sino también como especies, por ejemplo. Es por eso que nos preguntamos qué debemos de hacer, como debemos de actuar, para ser epistémicamente responsables, con nosotros mismos y con otros. Eso es lo

que da sentido a la epistemología. De la misma manera, el lenguaje también tiene una dimensión normativa porque es importante seguir ciertas reglas para poder aprovechar la existencia del lenguaje para alcanzar nuestros fines colectivos e individuales. Si no importara en absoluto como usáramos las palabras, no habría normatividad en el lenguaje. Pero claro que importa. Si no pudiéramos esperar que los otros usasen las palabras de acuerdo a su significado, no podríamos usarlas para comunicarnos, por ejemplo.

La normatividad tiene sentido, por lo tanto, en los espacios en los que no podemos dejar las cosas meramente a la suerte, sino que buscamos algún tipo de certeza (en el caso del infalibilismo), o por lo menos confianza (en el caso no infalibilista), entendiendo ambos “certeza” y “confianza” en un sentido no-epistemológico. En este sentido, un método, medio, herramienta, etc. es adecuado para perseguir ciertos fines, si, a partir de recursos que están de hecho disponibles (es decir, que no son muy costosos, de los que tenemos acceso y están bajo nuestro control) reduce el riesgo de falla.

Ahora bien, una vez que reconocemos que los medios pueden ser falibles, es decir, que un medio puede ser adecuado aun cuando no garantice al cien por ciento su fin, uno podría pensar que tenemos cuatro opciones:

1. Usamos un medio adecuado y obtenemos nuestro fin.
2. No usamos un medio adecuado y no obtenemos nuestro fin.
3. Usamos un medio adecuado y, sin embargo, no obtenemos nuestro fin.
4. No usamos un medio adecuado y, aun así, obtenemos nuestro fin.

Para el infabilista, la opción (3) no es genuina. Un medio es adecuado sólo si nos lleva a obtener nuestro fin. Pero el infabilista y el fabilista comúnmente no dicen nada sobre (4). En esta sección del libro veremos que esta división de cuatro tipos de situaciones no es suficientemente fina para dar cuenta de nuestras verdaderas prácticas normativas. Veremos que no todos los casos de tipo (1), por ejemplo, son normativamente equivalentes, por ejemplo, veremos que no siempre que se da (1) podemos hablar de un logro. Esta es la conclusión del famoso argumento de Gettier.

El argumento de Gettier se basa en dos premisas sencillas: Primero, que la justificación es falible, es decir, que es posible que una persona esté justificada en creer una proposición que es de hecho falsa (con los

ajustes adecuados para las proposiciones necesarias);<sup>9</sup> en segundo lugar, que la deducción (competente) extiende la justificación, es decir, para cualquier proposición  $P$ , si  $S$  está justificado en creer  $P$ , y  $P$  implica lógicamente  $Q$ , y  $S$  hábilmente deduce  $Q$  de  $P$  y acepta  $Q$  como resultado de esta deducción, entonces el que  $S$  crea que  $Q$  está justificado. Presumiblemente, ambas premisas se aplican tanto a la justificación *a priori* como a la *a posteriori*.

Para aquellos que puedan no estar familiarizados con el argumento de Gettier, permítanme ofrecer una reconstrucción muy general y abstracta (o, al menos, una reconstrucción de su caso disyuntivo II, el caso conjuntivo I es completamente análogo y dual):

- 1 Hipótesis a reducir: El conocimiento es creencia verdadera justificada
- 2 Premisa 1: La justificación es falible. Es posible que una persona esté justificada en creer una proposición que es de hecho falsa. Sea  $S$  tal persona y  $p$  la proposición falsa que  $S$  cree justificadamente.
- 3 Sea  $q$  otra proposición que  $S$  no sabe, aunque sea verdadera.
- 4  $p$  implica lógicamente  $(p \vee q)$ .
- 5 Supongamos que  $S$  ve la implicación de  $p$  a  $(p \vee q)$ , y acepta  $(p \vee q)$  en base a  $(p)$ .
- 6 Premisa 2: La justificación se transmite a través de, al menos deducciones lógicas básicas como la que va de  $p$  a  $(p \vee q)$ .
- 7 De 1, 5 y 6,  $S$  está justificado a creer  $(p \vee q)$ .
- 8  $q$  implica lógicamente  $s (p \vee q)$ .
- 9 De 3 y 8,  $(p \vee q)$  es verdadero.
- 10 De 7 y 9,  $S$  está justificado a creer la proposición verdadera  $(p \vee q)$ .
- 11 Sin embargo,  $(p \vee q)$  es verdadero porque  $q$  es verdadero (de 3, 8 y 9), y  $S$  no sabe que  $q$  es verdadero (de 3).
- 12 De 11,  $S$  no sabe que  $(p \vee q)$ .
- 13 Por lo tanto,  $S$  está justificado en creer una proposición verdadera –  $(p \vee q)$  – que no conoce, lo que contradice nuestra hipótesis inicial (1).

---

<sup>9</sup>. En el sentido relevante de "justificado" en el cual el que un agente  $S$  esté justificado en creer que  $P$  es una condición necesaria de que  $S$  sepa que  $P$

El argumento es perfectamente sólido, por lo que la única manera de resistirse a su conclusión sería atacar una de sus premisas. La premisa 2 parece razonable sin duda, pero ¿qué razones debemos tener para aceptar la primera premisa, es decir, la falibilidad de la justificación? Gettier no lo dice explícitamente, pero claramente parece una presuposición de la definición misma de conocimiento que se discute. Para que el conocimiento sea una verdadera creencia justificada, cada elemento de la definición debe hacer algún trabajo. Si la justificación no fuera falible, implicaría la verdad, y ésta sería descartable de la definición. Además, puesto que la justificación presupone la creencia, justificarse sería una condición necesaria y suficiente para el conocimiento. En otras palabras, la justificación no sería más que una palabra rara para hablar del conocimiento y como tal no podría definirla. En consecuencia, la justificación no puede implicar verdad, es decir, debe ser falible.

Algunos filósofos han considerado que la moraleja de los argumentos de Gettier es que hay un ingrediente o condición extra que falta ser añadido a la justificación y la verdad para obtener conocimiento. Sin embargo, uno debe tener cuidado de no agregar ninguna condición adicional que también se transfiera a través de deducciones lógicas inmediatas y sea falible, pues el argumento de Gettier también se aplicaría a tal condición adicional. Supongamos que alguien sostuviera la hipótesis de que el conocimiento no es una creencia verdadera y justificada, sino una creencia verdadera justificada más alguna otra condición. Llamemos a esta condición adicional C. Ahora, dado que C es falible, es posible que alguien S tenga una creencia falsa p que satisfaga C. Toma de nuevo otra proposición verdadera q que S no sepa. Si S no está justificado ni en creer p ni en creer q, toma entonces una tercera proposición r que S crea de manera justificada, pero no sepa. Una vez más, si S hace la inferencia de p a  $(pvq)$  o  $((pvq) vr)$ , tendría una creencia verdadera justificada que también satisface la condición C de la que, sin embargo, no tiene conocimiento. Por lo tanto, ninguna tercera condición de este tipo puede ayudar a construir el puente entre la creencia verdadera justificada y el conocimiento. Esto significa que si hubiera una nueva condición que añadir esta tendría que ser una que no se transmitiera por deducciones básicas competentes – lo cual sería muy raro – o que fuera infalible. Sin embargo, ya hemos visto que sería difícil encontrar una condición instrumental – es decir, una que hiciera más que reformular la noción de conocimiento, sino que nos ayudara a guiar nuestro comportamiento – que fuera infalible. Nos queda la tesis negativa de que la creencia verdadera y justificada no es conocimiento, puede ser una condición necesaria, pero no suficiente. Pero la probabilidad de encontrar

una condición adicional para completar el análisis es muy baja, por lo que el conocimiento no puede ser analizado en términos de justificación.

Nótese que aunque el argumento está tradicionalmente formulado en términos de conocimiento proposicional, es decir, conocimiento **que**, es también válido cuando se aplica al conocimiento práctico, es decir, al conocimiento **cómo**, pues éste también es falible. En este caso, la conclusión relevante sería que que si algo es el tipo de cosa que alguien podría saber cómo hacer, entonces es posible hacerlo, aun para quienes saben como hacerlo, no como manifestación de que saben hacerlo, sino por mera suerte.

Premisas:

1. Saber cómo es falible, es decir, un agente puede saber cómo hacer X y, sin embargo, fallar al intentar hacer X.
2. Si hacer X es una forma de hacer Y, entonces saber cómo hacer X implica saber cómo hacer Y.

Argumento:

3. Sea *a* un agente que sepa hacer X and sin embargo falle al intentar hacerlo.
4. Sea Y algo que *a* hace, pero sin saber cómo hacerlo.
5. Hacer X y hacer Y son maneras de hacer (X o Y).
6. *a* sabe como hacer (X o Y) (por 2, 3 y 5), hace (X o Y) (por 4 y 5) y, sin embargo, su hacer (X o Y) no manifestó su capacidad de hacer (X o Y).

Por lo tanto, *a* hace (X or Y) por pura suerte.

La idea fundamental detrás del argumento es que si hay (por lo menos) dos maneras distintas de hacer algo, basta que uno sepa hacerla de *alguna* manera para que sepa hacerla; esto significa que es posible que uno, que sabe hacerla de una manera, intente hacerla de dicha manera y al fallar, por suerte, terminé haciéndola de la otra. De esta manera, habrá hecho algo que sabe hacer, aunque por suerte. Esto demuestra que algunas cosas que se puede saber cómo hacer, también se pueden hacer (incluso por los que saben cómo hacerlo) por pura suerte.<sup>10</sup>

Aun más, creo que el argumento de Gettier es lo suficientemente amplio como para aplicarse no sólo al conocimiento sino a nociones normativas similares, es decir, aquellas en las que existe una meta (en este

---

<sup>10</sup>. Agradezco a Miguel Ángel Sebastián el haberme corregido sobre la formulación correcta de la conclusión de este argumento.

caso, la verdad) y un medio para alcanzar dicha meta que puede ser evaluada positivamente pese a no ser infalible. En otras palabras, el argumento de Gettier es pertinente siempre que hay riesgos involucrados, como lo desarrollaré en la siguiente sub-sección.

## B. Riesgo y Normatividad

"El futuro tiene muchos nombres.  
Para los débiles es lo inalcanzable.  
Para los temerosos, lo desconocido.  
Para los valientes es la oportunidad".

Victor Hugo

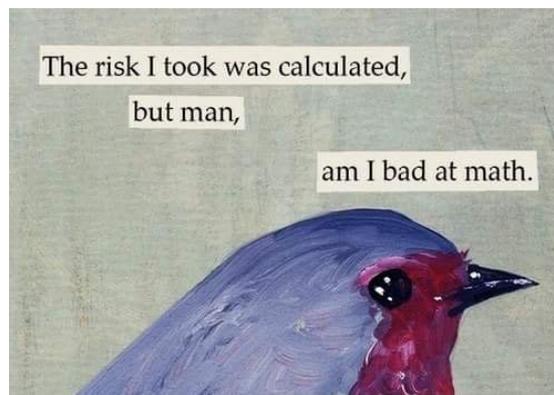
Cada vez que se toma un riesgo, existen por lo menos dos maneras de evaluar el acto: uno, juzgar si estuvo justificado asumir el riesgo (dado el estado de conocimiento del sujeto, la información disponible, las probabilidades, las ganancias posibles de los posibles resultados, etc.) y dos, simplemente juzgar si el resultado fue el deseado. Arriesgar todo nuestro patrimonio en una carrera de caballos de cuyos participantes no sabemos nada es insensato, aun cuando ganemos la puesta y la riqueza que obtengamos de la misma nos traiga enormes beneficios. Inversamente, hay un sentido en el que, aún cuando estamos justificados en tomar un riesgo, *post facto* sentimos que estuvo mal tomarlo si el resultado que obtuvimos no fue el deseado.<sup>11</sup> Distinguir entre estos dos tipos de juicios es uno de los puntos centrales de Bernard Williams en su seminal artículo "Suerte Moral" (1981). Cada vez que tomamos el volante de un auto, estamos corriendo el riesgo de atropellar a alguien. Si somos cuidadosos al manejar, reducimos este riesgo, pero no siempre lo eliminamos. Es decir, es posible manejar con cuidado y aun así atropellar a alguien. Cuando esto sucede, tenemos razón en pensar que hicimos algo horrible, aun cuando no podamos ser acusados de negligencia al no encontrarse ningún descuido o irresponsabilidad de nuestra parte. La primera intuición (la de que hicimos algo horrible, a saber, atropellar a alguien) es capturada por el primer criterio normativo, mientras que el segundo criterio busca capturar la intuición detrás del segundo juicio (de que somos inocentes, pues actuamos de manera

---

<sup>11</sup>. Desde el punto de vista psicológico, la evidencia parece indicar que ambos tipos de riesgos son evaluados por mecanismos cognitivos distintos, por lo que presumiblemente tienen historias evolutivas distintas (Cushman 2015).

cuidadosa y responsable). Cuando un acto es in-impugnable según el primer criterio, pero no según el segundo, tenemos un desafortunado caso de *mala suerte*.

Como ya he sugerido, esta distinción es muy importante para la discusión ética contemporánea, pues se encuentra al centro de la discusión alrededor de la suerte moral (Nagel 1979, Williams 1981). Sin embargo, también es importante para la epistemología y la semántica filosófica. Para entender su importancia en epistemología, basta reconocer que toda creencia involucra un riesgo. Si hemos aprendido algo de veinticinco siglos de escepticismo filosófico en occidente es que no podemos garantizar por completo la verdad de ninguna o prácticamente ninguna de nuestras creencias. En otras palabras, aunque sea de manera implícita, cada creencia nuestra involucra el riesgo de que lo que creamos no sea verdadero. Si nuestra creencia es justificada, hemos satisfecho los requisitos que impone el primer criterio de normatividad (hemos tratado de reducir el riesgo de falsedad, cumplido nuestras obligaciones epistémicas, etc.). Si nuestra creencia es verdadera, hemos satisfecho el segundo. La diferencia entre verdad y justificación, por lo tanto, no es sino la versión epistemológica de la misma distinción general entre ambos tipos de juicios.



Poner la distinción en términos epistémicos nos sirve además para recordarnos que existe un tercer criterio normativo involucrado en nuestra evaluación de actos arriesgados. Si bien es cierto que hay algo malo en tomar riesgos injustificados, o que algo malo sucede si no obtenemos el resultado deseado al correr un riesgo, como Gettier (1963) ya bien señaló, muchas veces no es suficiente estar justificado en asumir un riesgo *y además* obtener el resultado deseado, sino que también es deseable que la satisfacción de ambos criterios esté adecuadamente conectada. En otras palabras, para satisfacer este tercer criterio normativo, no es suficiente que uno haya tenido el cuidado suficiente para reducir los riesgos a un grado razonable, o que se haya obtenido el resultado que se esperaba, sino que es necesario también que dicho resultado se haya obtenido precisamente porque se tomaron dichos cuidados. El contraste aquí es con actos arriesgados

(posiblemente cuidadosos) en los cuales el resultado esperado se obtiene por *buena suerte*.

La distinción también es importante en semántica, pues todo lo dicho hasta ahora sobre la creencia se aplica *mutatis mutandi* también a la afirmación. Afirmar siempre involucra asumir un riesgo: el riesgo de que lo que digamos resulte ser falso. Esto es lo que significa decir que la verdad es la norma de la aserción. Sin embargo, en tanto que hay un riesgo involucrado, las tres fuentes de normatividad anteriormente mencionadas vuelven a aparecer: una afirmación puede estar justificada, ser verdadera, o ser verdadera gracias a que está justificada. Idealmente, una afirmación debe satisfacer los tres criterios; pero si no es así, uno debe ser cuidadoso distinguir las diferentes razones por las cuales la afirmación no es óptima. (En 2012, y en sesiones posteriores de este mismo libro, argumento que ignorar esta distinción es un tipo de error detrás del relativismo de John MacFarlane (2014), por lo menos, respecto a modalidades epistémicas). Es por eso que retractarse de una afirmación no es lo mismo que afirmar que lo dicho era falso. Uno puede también retractarse porque lo dicho no estaba justificado, por ejemplo. No sé si existen casos en los que uno deba retractarse de una afirmación gettierizada, es decir, verdadera, justificada, pero donde la verdad no esté adecuadamente vinculada a la justificación, pero supongo que sí.

Como mencioné al final de la sección pasada, la distinción se aplica también fuera de la epistemología en cualquier situación en la que hay un riesgo – calculable o no. Permítanme ilustrar esto con un ejemplo. Supongamos que efectivamente en Berlín se decide colocar bloques de hormigón alrededor de las calles donde se esperan se junten multitudes para impedir otro ataque con camiones. Como bien señalaba el oficial anónimo de las fuerzas anti-terroristas, esto es una medida razonable aun cuando no nos pueda garantizar al cien por ciento que no sucederá otro ataque de este tipo. Después de todo, sigue siendo posible que los terroristas decidan atacar una zona sin bloques de hormigón. Si esto sucede, los bloques de hormigón no servirán de nada. En este sentido, aun cuando se coloquen los bloques de hormigón, no se ha eliminado el riesgo de ataque en camión. Aun así, esto no significa que es seguro que habrá un ataque y que este será exitoso en la zona donde no hay bloques de hormigón. Si había planes para atacar una zona donde se pusieron bloques y estos sirvieron para disuadir a los atacantes de intentarlo, podremos decir que fue un logro de las fuerzas anti-terroristas el evitar dicho posible ataque. Pero si el ataque sucede en otra area, su éxito o fracaso ya no tendrá nada que ver con los bloques sino con factores externos. Quedará a la suerte. Puede ser que los atacantes sean tan incompetentes que el ataque falle. En ese caso podríamos decir que las fuerzas

anti-terroristas tuvieron buena suerte. Pero si mas bien lo logran, no podremos acusar a dichas fuerzas de negligentes, sino que lo más adecuado sería decir que se tuvo mala suerte.

## C. ¿Qué es lograr algo?

Los temas de la suerte y los accidentes han alcanzado una nueva importancia en años recientes debido al proyecto de ciertos filósofos de definir la agenda epistemológica en estos términos. La idea básica es sencilla: No se puede SABER por accidente; lo que sí se puede es tener una creencia que sea verdadera por accidente. Por ejemplo, si un estudiante copia mal en un examen de matemáticas, pero por suerte le atina el resultado correcto, podemos decir que tiene una creencia matemática verdadera que no es conocimiento.

Aunque los diccionarios los presenten como sinónimos, intuitivamente para lograr algo, no basta conseguirlo, obtenerlo o alcanzarlo. En otras palabras, una vez que nos proponemos algo como objetivo o emprendemos una tarea, no basta tener éxito y conseguir eso que nos proponemos para decir que lo logramos. Es necesario también que el logro no haya sido accidental o por suerte, sino el resultado directo de ejercer alguna capacidad o hacer un esfuerzo dirigido a alcanzar dicho fin. Por ejemplo, si alguien se propone comprar un coche nuevo, se pone a trabajar ahorrar, etc. hasta que después de muchos años de trabajo se gana el coche en una lotería, no diríamos que por fin *logro* obtener el coche que quería. Podríamos decir que por fin *obtuvo* lo que quería, pero no diríamos que hubo ningún logro de su parte. Tuvo suerte, buena suerte.

Hasta el momento, esta concepción de lo que es un logro no es controversial. La disputa filosófica empieza a partir de aquí. En general, dos preguntas permanecen abiertas: la cuestión *descriptiva* de explicar en qué consiste exactamente el que lograr un objetivo se deba al ejercicio de alguna capacidad o sea el resultado directo de hacer un esfuerzo para alcanzarlo, es decir, explicar exactamente qué tipo de relación debe haber entre el esfuerzo o ejercicio de una capacidad y la obtención del objetivo para decir que alcanzar dicho objetivo fue un logro y no un afortunado accidente, y la cuestión *fundacional* de porqué valoramos más un logro (debido al esfuerzo propio o al ejercicio de alguna capacidad) que el obtener lo que deseamos de manera accidental o por buena suerte, mientras que castigamos igual los casos en los que el mal es resultado de mala suerte que cuando no lo es (Martin & Cushman 2015). Tratar de responder esta segunda pregunta fundacional ha sido uno de los objetivos centrales del trabajo de filósofos como John Grecco, Peter

Graham, Ernest Sosa, etc. La cuestión es especialmente complicada porque al tratar de definir un logro como, por ejemplo, aquel que se debe al ejercicio de una capacidad virtuosa, el problema pasa a determinar (de una manera no circular) qué es una capacidad virtuosa y cuando se ejerce. Una vieja propuesta de solución ha sido tratar de definir suerte y/o accidente de manera independiente y luego definir el logro como la realización no accidental de una tarea. El esfuerzo más conocido en esta dirección tal vez sea el de definir conocimiento en términos de seguridad (Sosa 2007). Bajo esta perspectiva general, un resultado es accidental si no es seguro que hubiera pasado en circunstancias similares (o mas bien, similares en los aspectos relevantes). Un resultado es seguro, en este sentido, si depende condicionalmente del esfuerzo que se hizo para obtenerse, es decir, si en los mundos posibles adecuados (adecuados para el ejercicio de la aptitud relevante) en que se hace el esfuerzo, se da el resultado. Por ejemplo, un experto sexador de pollos no identifica por accidente el sexo de un pollo, ya que sería capaz de identificar el sexo del mismo pollo (u otro suficientemente parecido) en casi cualquier circunstancia similar. En otras palabras, esperamos (de manera derrotable) que sea raro que quién esté bien preparado fallé, que si tenemos suerte, ésta sea endeble, en vez de robusta, es decir, que pequeños cambios cambien el resultado positivo o negativo.

Otras propuesta más reciente de resolver el problema descriptivo ha sido tratar de definir las habilidades como un tipo de disposición, y luego definir el logro en términos del ejercicio de una habilidad. En esta propuesta, una persona tiene la habilidad de realizar una tarea si tiene la disposición a realizar exitosamente dicha tarea en las circunstancias relevantes con un grado sustancial de confianza (Greco 2010). Bajo esta concepción de lo que es una habilidad, una persona puede no lograr realizar una tarea por una de tres razones: porque no tiene la habilidad de realizar la tarea, porque las condiciones necesarias para el ejercicio de la habilidad no se dan o porque -- aunque el sujeto tenga la habilidad y las circunstancias sean las adecuadas -- dicha habilidad sigue siendo falible. Finalmente, también se ha tratado definir la posesión de una capacidad (o por lo menos, de nuestras capacidades básicas) en términos evolutivos: tenemos la habilidad de lograr ciertas tareas porque uno o varios de nuestros órganos (u otros rasgos fenotípicos) fueron seleccionados con dicha función. En este sentido, por ejemplo, tenemos la habilidad de recordar hechos pasados (en las circunstancias relevantes, y de manera falible) porque parte de nuestro aparato cognitivo fue seleccionado para realizar esta función. Sin embargo, aun cuando este tipo de estrategias puedan servir para

explicar algunas de nuestras capacidades – las mas “naturales” por decirlo de alguna manera –, dudo que sea viable encontrar una explicación adaptativa para todas ellas.

Por mi parte, yo prefiero modelar la noción de justificación en términos de “estar preparado” en vez de “tener la capacidad” o “ejercer una habilidad”, porque esta noción claramente revela aspectos importantes de la justificación, algunos de los cuales filósofos de la virtud como Sosa y Grecco habían ya señalado. En particular, porque revela que, detrás de la noción absoluta de justificación – según la cual, o bien se está justificado/se está preparado/se tiene la capacidad o no se tiene – hay otra noción relativa – según la cual, uno está preparado para una situación o no, relacionadas de tal manera que uno está preparado (en el sentido absoluto) cuando se está preparado (en el sentido relativo) para no sólo las situaciones que más se esperan, sino también para enfrentar por lo menos algunas eventualidades **previsibles**. ¿Cuáles y cuantas? Las suficientes, y aquí, por supuesto, la noción de suficiente convierte a la noción de justificación en un aspecto gradual: uno puede estar más o menos justificado si actuó preparado para más o menos situaciones que se pudieran presentar.

Ha habido intentos de modelar la suerte de manera modal o probabilista, pero han fallado, pues aunque son nociones relacionadas, la suerte es determinante sólo en casos imprevisibles y lo previsible en este sentido no es lo mismo que lo posible o lo probable. En otras palabras, lo más posible o probable no es siempre también lo mas previsible, ni vice versa. La que hace a una situación posible imprevisible no es su ser meramente poco probable, sino su ser racionalmente (y, por lo tanto, también responsablemente) insignificante, y ésta es una noción irreduciblemente normativa, no sólo descriptiva.

Desafortunadamente, la palabra “previsible”, en español, es ambigua.<sup>12</sup> La noción relevante para entender lo que yo llamo situaciones **previsibles** *no* es la noción epistémica de ‘prever’, sino la noción mas fuerte de ‘**prevenir**’. Prefiero el termino “previsible”, a “prevenible” porque este último se usa de manera común en el sentido de “evitable”, pero lo que se busca evitar no es la situación previsible misma, sino las consecuencias negativas que podrían resultar de no estar preparado para ella. En otras palabras, so “previsible” para referirme, no a lo que se puede prever, sino lo que se debe prevenir. Prevenir no es sólo

---

<sup>12</sup>. Además, etimológicamente refiere al futuro, aunque nuestras metas pueden estar en el pasado o en el presente también. Agradezco a Catherine Elgin el haberme llamado la atención en este respecto.

darse cuenta de que algo es posible – que es como comúnmente entendemos la noción de ‘prever’, sino responder de manera responsable ante este conocimiento, es decir, **prepararse** para que se dicha posibilidad. Así pues, una situación es mas previsible que otra, en el sentido relevante para entender la normatividad, no si es mas subjetivamente probable o si se da en un mundo posible mas cercano en algún sentido puramente descriptivo, sino si sería mas irresponsable no estar prevenido. Para ser mas precisos, una situación *A* es mas previsible que otra situación *B*, para un fin *F*, si una persona que persigue *F* y no previene la posibilidad de *A* (es decir, no hace lo suficiente para estar adecuadamente preparado para *A*) es más irresponsable que una persona que persigue *F* y no previene la posibilidad de *B*. Esto se deba a que la noción misma de “suerte” no es meramente descriptiva, sino también normativa. Tener suerte es haber obtenido algo que no se ha ganado. Esperamos que sea raro que quién esté bien preparado falle, es decir, esperamos que la suerte no sea la regla, sino la excepción. Pero lo que hace a los casos de suerte excepcionales no es su ser poco probables. En sentido estricto, lo que distingue a quién está preparado (en el sentido que aquí me interesa) de quien no lo está no es la pura **cantidad** de situaciones para las que está preparado, sino el **tipo** de situaciones para las que está preparado. Prepararse para un posible huracán cuando se vive en zona de huracanes es racional y responsable en un sentido que no lo es prepararse para una invasión extraterrestre. Por eso, a la persona que se preparo para una invasión extraterrestre que no sucedió, pero no para un huracán previsible no le decimos que tuvo mala suerte, sino que fue negligente; mientras que a la persona que se preparo para un huracán previsible que nunca llegó, pero no para una invasión extraterrestre imprevisible no le juzgamos de negligente, sino a lo más, de tener mala suerte.

Desde esta perspectiva, decimos que estamos preparados (en el sentido relativo) para una situación si podemos garantizar que, si se da dicha situación, obtendremos el resultado (si así lo deseamos), pero esto no significa que si se da una situación para la que no estamos preparados, podemos garantizar que **no** se va a obtener el resultado, sino que **no** podemos garantizar que **sí** se va a dar, es decir, se puede obtener o no, pero ya no depende de nosotros, sino de la suerte. En otras palabras, una vez que se da una situación para la que no estamos preparados, está fuera de nuestro control qué vaya a suceder – el mundo es comúnmente indiferente a nuestros deseos: se puede dar el resultado que deseamos o puede que no, dependiendo de nuestra suerte, buena o mala.

Comúnmente invocamos el estar preparados para algo (y nociones similares) cuando queremos explicar cosas, especialmente cuando queremos explicar el éxito o fracaso de una empresa. Decimos cosas como "La charla salió muy bien porque la preparé con cuidado", "El Canelo perdió la pelea porque aún no estaba listo para pelear con un campeón", etc. Esa es la noción que quiero recuperar aquí y, en este respecto, es importante notar que la relación entre el estar preparado para algo y el éxito o fracaso que ello explica no es una relación meramente causal. Efectivamente hay algún tipo de relación causal entre el estado resultante de mi preparación para X y X. Por ejemplo, para prepararme para mi próximo viaje a Ciudad Obregón, tendría que asegurarme de que mis maletas estén hechas con las cosas que necesitaré, que he hecho los arreglos necesarios para llegar a tiempo, comprando un boleto de avión, estando en el aeropuerto a tiempo, etc. Si efectivamente logro llegar a tiempo a mis compromisos en Ciudad Obregón, debe haber una relación causal entre algunos de los estados resultantes de mi preparación como el que tenga mi billete de avión a la hora de presentarme al aeropuerto y el estado de yo estar en Ciudad Obregón algunas horas después, pero aquí parece que lo causalmente eficaz no es mi preparación misma, sino algunos de los hechos físicos subyacentes. Además, probablemente habrá muchos hechos involucrados en mi preparación para X (en el sentido absoluto) que podrían ser ineficientes para causar X, pero esenciales para estar listo para X (en el sentido absoluto). Por ejemplo, como veremos mas adelante, podría sobrevivir a un huracán porque me preparé de manera adecuada sin que todo lo que hice (y necesitaba hacer) para estar listo para el huracán sea causalmente responsable de mi supervivencia real.

Por otro lado, la relación entre el estar preparado (en el sentido absoluto, pero no en el sentido relativo) y el éxito resultante tampoco puede ser una relación de fundamentación metafísica, si uno piensa – como es costumbre, pero creo yo que de manera equivocada (Barceló 2015) – que esta relación implica necesidad metafísica ya que, como he insistido a lo largo del libro, el vínculo entre la preparación (absoluta) y el éxito resultante es falible. Además, no parece que la preparación sea (ni siquiera parcialmente) constitutiva del éxito. Los fundamentos metafísicos de que mi equipo haya perdido su juego son todos los hechos que sucedieron durante el juego (en particular, que el balón cruzó nuestra línea de gol más veces de las que cruzó la de nuestro oponente), no lo que sucedió antes del partido. Por lo tanto, a menos que adoptemos una concepción falible de la fundamentación metafísica, como la que defiendo en (Barceló 2015) – parece haber un tipo distinto de vínculo que no es ni causal, ni metafísico (y presumiblemente tampoco

conceptual ni lógico) entre el estar preparado (en el sentido absoluto) para algo y que hayamos logrado ese algo.

Dado que en la mayoría de los casos cuando uno está preparado (en el sentido absoluto) para algo, no se está completamente preparado (en el sentido relativo) para cualquier situación eventual que se pueda dar, es compatible con el estar preparado que no se logre el resultado esperado. Estos son los casos de mala suerte. Cuando creo, afirmo o, en general realizo una acción dirigida a un objetivo o que puede ser evaluada respecto a su resultado, puedo actuar preparado no preparado, pero dicha preparación rara vez va a garantizar de manera total el éxito de la empresa. De tal manera que, aunque esté preparado, mi empresa puede terminar en fracaso.

Por otro lado, y por supuesto, así como no es suficiente estar preparado para obtener el resultado deseado, tampoco es necesario. Si uno tiene buena suerte, puede obtener éxito sin haberse preparado, como el proverbial burro que tocó la flauta. Pero aún más, y esta es la moraleja principal del famoso argumento de Gettier, es posible estar preparado (en el sentido absoluto), que se de una situación para la cual uno no está preparado (en el sentido relativo) y sin embargo, se dé el resultado deseado. Es decir, así como se puede tener mala suerte, se puede tener buena suerte, aun cuando se esté preparado. Como había mencionado antes, ambos casos de suerte, buena y mala, tienen en común que resulten de que se dé una situación para la que uno no está preparado.

Pongamos un ejemplo para ilustrar esto: Supongamos que vivo en una zona de tormentas severas, donde es un riesgo presente la posibilidad de inundaciones. Si soy una persona responsable, contactaré a las autoridades adecuadas para enterarme si mi vivienda se encuentra en una zona de derrumbes o inundación repentina, además de informarme sobre los planes y señales de emergencia, incluyendo información sobre rutas de escape y localización de albergues. También realizaré regularmente simulacros de evacuación y tendré siempre a la mano provisiones de emergencia como agua, alimentos no-perecederos, equipo de primeros auxilios, etc. Cada uno de las cosas que haré para prepararme me ayudará para lidiar con algunas de las posibilidades que se pueden dar en caso de inundación. Por si se producen fisuras en las tuberías de gas, por ejemplo, habré cerrado las llaves maestras de abastecimiento de gas tan pronto como sepa del riesgo inminente de inundación, por ejemplo. Es posible que la tormenta no produzca fisuras en las tuberías de gas, pero es responsable estar preparado para dicha posibilidad y es por ello que cierro la llave maestra.

Entre mas me preparo para una posible inundación, mayores serán mis posibilidades de sobrevivir con el menor daño a mi propiedad y a la gente bajo mi responsabilidad. Y así como yo me preparo para la posible inundación, mis vecinos también se prepararán si son responsables. Es más, es posible que mi vecino se prepare **más** que yo. Si yo guardo comida para mantener a mi familia por cuatro o cinco días, es posible que él guarde comida para mantener a la suya durante dos semanas; si yo reviso la caducidad de mi extintor de incendios cada seis meses, él puede hacerlo cada tres, etc. Aun así, eso no hace que él sí esté preparado y yo no. Hay un umbral de ‘preparacionez’ tal que si uno lo pasa, no importa cuanto más se prepare, es correcto decir que uno ya está preparado.

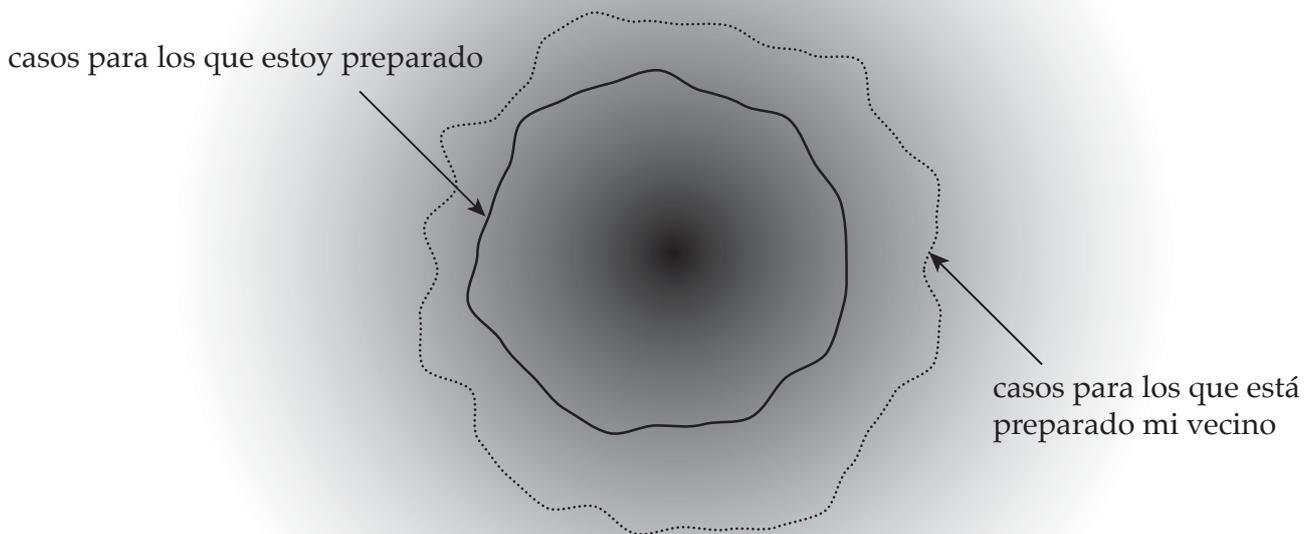
Podemos usar el siguiente diagrama para explicar mejor esta idea.<sup>13</sup> Imagínense este plano como el espacio de posibilidades de tal manera que cada punto de él representa una situación futura relevante ordenada de tal manera que entre más cercana al centro, más previsible es. Uno podría garantizar por completo el éxito sólo si estuviera preparado para cualquier situación futura posible. Sin embargo, el esfuerzo necesario para ello sería excesivo, así que es racional estar preparado sólo para las situaciones futuras más previsibles, es decir, para los puntos más cercanos al centro. Por supuesto, dónde dibujar la línea entre las posibilidades para las que vale la pena prepararse y las que no es una cuestión imprecisa. Aun así, queda claro que no es necesario prepararse para las posibilidades más remotas para estar adecuadamente preparado. Hay un umbral vago, pero existe, entre las posibilidades que hay que tomar en cuenta y aquellas que no.

Supongamos entonces que yo he cruzado ese umbral, y aunque no estoy tan preparado como mi vecino, estoy adecuadamente preparado para una posible inundación. Esto significa que aunque mi riesgo es mayor que el de mi vecino, ambos hemos sido responsables en tanto ambos riesgos son razonables. En el diagrama, la línea oscura representa las situaciones futuras posibles para las que estoy preparado y la línea punteada las situaciones para las que mi vecino está preparado. Aunque la región cubierta por la de mi vecino es mas grande que la mía, ambas cubren las posibilidades más previsibles. Pero esto significa también que habrá situaciones en las cuales, pese a estar ambos preparados en el sentido absoluto, yo no esté preparado y mi vecino sí. Esto puede significar que él y su familia sobrevivan el desastre sin mayor dificultad y mi familia

---

<sup>13</sup>. Tomo el diagrama de Miguel Ángel Fernández, aunque creo que no lo ha usado en su obra escrita, sino sólo en sus presentaciones orales.

sufra o ni siquiera sobreviva. El desastre puede ser tal que, por ejemplo, estemos aislados no por tres o cuatro días, sino por dos semanas o más. En dicho caso, mis provisiones no serán suficientes, aunque las de mi vecino sí. Si eso sucede, uno estaría tentado a pensar que en realidad **no estaba** preparado, pero lo mejor sería decir que tuve muy mala suerte.



Cambiamos un poco el ejemplo e imaginemos que la inundación es de tal magnitud que estamos aislados por más tiempo del que estábamos preparados, pero mi vecino decide compartir sus provisiones conmigo y mi familia y así nos ayuda a sobrevivir hasta que somos rescatados. En ese caso debemos decir que estábamos preparados para la inundación y que la sobrevivimos, pero es claro que no podemos decir que la sobrevivimos **porque** estábamos preparados. En otras palabras, nos encontramos en un caso de Gettier. Estábamos preparados para las situaciones más normales que esperar, pero no para caso extraordinarios como el que de hecho se dio, y es por eso que si sobrevivimos o no lo hacemos, no será porque estábamos preparados o porque no lo estábamos, sino por suerte: buena suerte si sobrevivimos, mala suerte si no lo hacemos.

Traslademos lo que acabamos de decir sobre inundaciones al caso epistémico. Creer, había dicho ya, involucra siempre un riesgo: el riesgo de que lo que creamos sea falso, y lo epistémicamente responsable es hacer algo para evitar que esto suceda. Si me baso en evidencias y buenas razones, podré sentirme satisfecho en haber sido responsable con mis creencias aunque no pueda garantizar al cien por ciento que ellas sean verdaderas (no puedo, por ejemplo, asegurarme que no soy un cerebro en una cubeta o que no estoy en alguna otra situación escéptica). Al creer, me comprometo a estar preparado para que el mundo sea como creo que es. Si el mundo es como es racional esperar que sea, entonces mis creencias serán verdaderas,<sup>14</sup> pero no podemos garantizarlo.

Tener justificación para una creencia particular, digamos la creencia de que  $p$  es estar preparado para que el mundo sea de ciertas maneras (las maneras en las que sería racional esperar que el mundo sea y por lo menos algunas maneras excepcionales), maneras en las que  $p$  es verdadero. Pero no es necesario estar preparado para cualquier manera que sea el mundo.<sup>15</sup> Ni el que tiene una creencia justificada ni el que la tiene sin justificación alguna están preparados para cualquier manera que sea el mundo, de tal manera que si el mundo es de alguna de estas maneras para las que no estamos preparados, no podremos garantizar que nuestra creencia será verdadera. Es posible que sí lo sea, pero también es posible que no. De cualquier manera, no dependerá de nosotros, sino de la suerte. Si tenemos buena suerte, nuestra creencia será verdadera; si tenemos mala suerte, será falsa. La justificación no nos vacuna contra la suerte; a lo más, sólo disminuye la probabilidad de que dependamos de ella. Esto es lo que significa decir que la justificación reduce el riesgo de fallo. Tanto la buena como la mala suerte tienen en común que resulten de que se dé una situación para la que uno no está preparado.

Lo interesante es que esto se puede extender no sólo a otros casos donde se persigue un objetivo, sino a cualquier estado humano o acción que pueda evaluarse positiva o negativamente. Si creemos, por ejemplo,

---

<sup>14</sup>. Esta relación entre la situación para la que estoy preparado y la verdad de lo que digo no es una inferencia (es decir, no infiero que el mundo es como creo que es de que el mundo es como es racional esperar que sea, sino que simplemente creo que el mundo es como creo que es y presupongo que el mundo es como es racional esperar que sea, pero esta presuposición no es un estado o acto mental separado) aunque bien puede haber una implicación lógica o metafísica aquí.

<sup>15</sup>. Es mas, ni siquiera es necesario estar preparado para todas las maneras en las que el mundo podría ser y donde  $p$  es verdadero.

que es bueno estar delgado, por ejemplo, podemos distinguir entre casos en los que se es afortunado ser delgado de aquellos en los que es un logro; si creemos que la sopa que estamos comiendo es deliciosa, podemos distinguir entre casos en los que por suerte nos tocó una sopa deliciosa de casos en los que es el producto de la capacidad del cocinero de preparar una sopa deliciosa. Igualmente, podemos distinguir casos en los que “hasta el mejor cazador se le escapa la liebre” de casos en los que la liebre se escapa precisamente porque somos malos cazadores. Y para cualquiera de estos valores podemos generar fácilmente casos de Gettier. Podemos imaginar casos en los que el buen cazador caza la liebre, pero no porque es buen cazador sino porque tuvo suerte. O casos en los que hicimos todo lo que sería razonable esperar de una persona responsable para ser delgados y aun así nuestra delgadez sea más el resultado de suerte que de nuestro esfuerzo por mantenernos delgados. Todos estos son casos Gettier.

CREENCIA	VERDADERA	JUSTIFICADA	CONOCIMIIENTO
ACCIÓN DIRIGIDA HACIA UN OBJETIVO	ALCANZAR EL OBJETIVO	Haber hecho lo suficiente para lograr el objetivo	LOGRO
EMPRESA	ÉXITO	PREPARACIÓN	LOGRO
ACCIÓN MORALMENTE EVALUABLE	BUENO EN SUS CONSECUENCIAS	RESPONSABILIDAD (NO NEGLIGENCIA)	
Aserción	verdad	justificación	conocimiento

Aunque todos estos son espacios en los cuales podemos aplicar el esquema que aquí he dibujado, a los filósofos le interesan sólo ciertos tipos de acciones y estados y sólo ciertos tipos de normas y valores. Por ejemplo, a la epistemología suele interesarle la creencia y tomar a la verdad como criterio fundamental para evaluarla. A la filosofía del lenguaje, en contraste, suelen interesarle los actos de aserción, aunque también suelen tomar a la verdad como criterio fundamental para evaluar dichos actos. Cuando afirmo que el mundo es de cierta manera, me comprometo a que, a menos que suceda algo poco previsible, entonces será como afirmo que es. En ambos ámbitos este modelo se aplica de manera directa. Sin embargo, también podemos aplicarla a otros campos del lenguaje de interés filosófico. Por ejemplo, como ya mencionaba en la sección anterior, también podemos entender el uso que se da de los predicados o conceptos bajo este esquema.

Idealmente, poseer un concepto debería de implicar ser capaz de determinar para cualquier objeto si cae bajo su extensión o no. Sin embargo, es racional usar un concepto aun cuando no seamos capaces de determinar su extensión para cualquier objeto posible, siempre y cuando estemos preparados para hacerlo en los casos más previsibles. Es por ello que nuestros conceptos son vagos. Así pues, podemos concebir la posesión de un concepto como el estar preparado para usarlo y a la vaguedad como una manifestación previsible de la falibilidad de nuestras capacidades. En el proximo capítulo, desarrollaré algunas de estas aplicaciones esperando que mi propuesta adquiriera mayor definición a través de este tipo de ejemplos.

## IV. Aplicaciones y Ejemplos

### A. Casos Gettier en Matemáticas

Una vez que hemos extraído las moralejas de la falibilidad de nuestros medios y del argumento de Gettier, la pregunta relevante para la filosofía de la matemática es si es posible explotar estos nuevos espacios teóricos para renovar su agenda epistemológica. ¿Podemos entender mejor el conocimiento matemático si empezamos por plantearnos que el quehacer matemático no ha dado los resultados que ha dado por accidente no suerte? Si es así ¿cuales son las habilidades o capacidades relevantes? ¿Probar teoremas? ¿Sistematizar intuiciones? Dadas las obvias limitaciones de espacio y tiempo, en este texto me enfocaré sólo a la primera de estas putativas habilidades: el probar teoremas. Además, el abordar la cuestión de si el probar teoremas es un método infalible de producción de conocimiento es muy relevante para la cuestión de si efectivamente todos nuestros mecanismos y métodos de producción de conocimiento son falibles, como sostiene el falibilista, o existe por lo menos alguno que sea infalible. A fin de cuentas, casi siempre que se aborda este problema suele salir a colación el caso de las pruebas matemáticas formales como un posible contraejemplo al falibilismo.

Para comenzar, vale la pena empezar preguntándose si la noción de “prueba” relevante para la epistemología de las matemáticas es una noción trascendente o instrumental, es decir, si la regla de que debemos creer sólo en proposiciones matemáticas probadas es una regla trascendente que sólo puede servirnos para definir una de nuestras metas epistémicas, o si es una regla instrumental que puede servirnos para guiar nuestro actuar como matemáticos. La cuestión es fundamental y no es sorprendente que haya sido una de las preocupaciones centrales de filósofos y matemáticos durante el período conocido como “crisis de los fundamentos” a finales del siglo XIX y principios del siglo XX. Esta preocupación es central al proyecto epistemológico de, por ejemplo, David Hilbert. Su programa tenía por objetivo central asegurarse de que la noción de “prueba” operativa en matemáticas fuera, al mismo tiempo, infalible y caracterizable en términos no-trascendentes, es decir, que el método de pruebas requiriera para su implementación sólo de capacidades cognitivas básicas que estuvieran a disposición de cualquier matemático. Por eso Hilbert restringió su noción de ‘prueba’ sólo a aquellas que pertenecieran a sistemas formales cuyas reglas de deducción y axiomas

podrían definirse de manera recursiva con relibros finitos (a decir verdad, la noción misma de recursividad tenía como objetivo capturar de manera rigurosa nuestra noción de calculable por cualquier persona). Desafortunadamente, gracias al trabajo de Gödel y otros, hoy sabemos que es imposible producir una noción de prueba que satisfaga los tres requisitos que buscaba Hilbert, es decir, que fuera infalible, instrumental y nos permitiera conocer todas las proposiciones matemáticas verdaderas. Esto significa que la noción de prueba operativa en matemáticas es una noción trascendente o falible, es decir, que cuando en matemáticas se prueba algo, de hecho, no se trabaja con una noción de prueba formal en el sentido de Hilbert.

Hay, sin embargo, otras razones para pensar que la noción formal de ‘prueba’ es trascendente. Por principio de cuentas, hablamos de “probar un teorema”, no cuando se hace el esfuerzo, sino cuando se alcanza obtener la prueba. Si el resultado no es una prueba, no decimos que hemos probado un teorema, sino sólo que lo hemos intentado. En otras palabras, para que podamos llamar a una acción “probar un teorema” es necesario que su resultado sea una prueba. Cabe por lo tanto preguntar si es suficiente que el resultado de una actividad sea una prueba para hablar de probar o es necesario algo más, algo así como el empleo de una habilidad. En otras palabras ¿es posible probar algo por accidente (en el sentido relevante a este ensayo)? Supongamos que un pésimo estudiante, que no ha estudiado en clase ni ha desarrollado la habilidad de probar teoremas en, digamos, geometría diferencial, desesperado decide improvisar en su examen una secuencia de fórmulas al azar las cuales, afortunadamente para él, terminan formando una prueba perfecta del teorema de Stokes. ¿Diríamos que el estudiante sabe que la integral de una función  $f$  en el intervalo  $[a, b]$  puede ser calculada por medio de una antiderivada  $F$  de  $f$ ? Seguramente no. Parece que algo le falta. Lo mismo diríamos de alguien quién, sin el menor conocimiento de geometría diferencial, se propusiera aprenderse de memoria la secuencia de fórmulas que forman una prueba de dicho teorema. Podemos decir que este último personaje es capaz de producir a voluntad pruebas del teorema de Stokes. Sin embargo, me parece que en este caso tampoco podríamos decir que sabe que dicho teorema es verdadero.

¿Qué consecuencias epistemológicas podríamos sacar de estos ejemplos? Por un lado, parece ser que para probar un teorema no es suficiente producir una prueba. ¿Significa esto que el probar es un *logro* en el sentido arriba caracterizado? La respuesta no es muy obvia. En el primer ejemplo, podríamos decir que precisamente la razón por la cual lo que hizo el estudiante no es probar el teorema de Stokes es por que su

prueba no fue el resultado de ninguna habilidad. No tenemos la menor seguridad de que, en otras circunstancias similares, el estudiante pueda volver a ser capaz de producir dicha prueba. Sin embargo, en el segundo caso sí tenemos a alguien ejerciendo una habilidad para producir la prueba. Nuestro segundo personaje sí puede volver a probar el teorema a voluntad (en las circunstancias adecuadas, obviamente). El resultado no es accidental, y por lo tanto, debería poder ser considerado un logro, no una mera obtención de un resultado positivo.

Tal parece que enfrentamos un horrible dilema: o bien tenemos que rechazar la hipótesis de que para probar un teorema basta lograr probar un teorema de manera no accidental, o rechazamos el presupuesto de que lo que es capaz de producir el segundo de nuestros personajes es un prueba propiamente dicha. Aunque sea una teoría del error, este segundo cuerno del dilema no es tan ad-hoc como podría parecer. Después de todo, no es raro escuchar a filósofos decir que las secuencias de fórmulas o en general el tipo de cosas que se pueden escribir en exámenes y aprender de memoria sin haber entendido **no** son pruebas propiamente dichas. Para fundamentar esta concepción de la prueba matemática, revisemos un poco más cierta concepción tradicional de la prueba como justificación epistémica. Tradicionalmente, la justificación que otorgan las pruebas a nuestras creencias en la verdad de los teoremas probados se considera una justificación indirecta o inferencial. Comúnmente se piensa que el problema epistemológico fundamental para las matemáticas es el de la justificación de nuestras creencias en los axiomas de nuestras teorías, ya que se piensa que nuestras creencias en el resto de los teoremas están justificadas por las pruebas. Bajo esta concepción, las pruebas son un tipo de inferencia, es decir, procesos psicológicos. Por eso, no podemos decir que el estudiante de nuestro ejemplo ni quién aprendió de memoria las fórmulas efectivamente probaron un teorema, ya que la inferencia no sucedió en la mente de ninguno de ellos.

Ahora bien, si las pruebas son procesos inferenciales, entonces regresamos a nuestra pregunta anterior ¿es suficiente haber obtenido el resultado buscado (la inferencia válida) para decir que se probó el teorema? Es decir ¿es posible probar un teorema por accidente? ¿Es posible encontrar escenarios parecidos a los anteriores, pero no sobre la prueba escrita, sino sobre la prueba psicológica? Supongamos ahora un segundo estudiante, uno que sí sabe un poco de geometría diferencial, pero que como ser falible que es, aún se le van algunas inferencias, es más, a veces se le van algunas inferencias lógicas básicas. Por ejemplo, a veces

se confunde acerca de cómo funciona la implicación material y a veces cree que es válido inferir la negación del consecuente de la negación del antecedente. Además, también llega a confundir cuál es el antecedente y cuál el consecuente. Supongamos ahora que a la hora que se propone probar el teorema de Stokes realiza todos los pasos a la perfección, excepto en un paso en el cual tiene ya probados un condicional y la negación de su consecuente. Por error, identifica mal el antecedente y el consecuente del condicional, por lo que cree haber probado el condicional y la negación de su antecedente. En ese momento también se confunde sobre cómo aplicar el *Modus Tollens* y cree que éste valida inferir la negación del consecuente a partir de la negación del antecedente. El resultado, por supuesto, es una inferencia válida, un *Modus Ponens*. Si no vuelve a cometer un error, el estudiante habrá realizado puras inferencias válidas y así inferido válidamente el teorema de las premisas relevantes. ¿Diríamos por lo tanto que el estudiante probó el teorema? Admito que mis intuiciones a este respecto no son tan firmes y estables como las de los ejemplos anteriores, pero el epistemólogo tradicional tendrá que aceptar que sí, que el estudiante ha probado el teorema. Según la epistemología tradicional, no importa si la prueba es resultado del azar o del esfuerzo, efectuarla es suficiente para tener conocimiento. En consecuencia, el estudiante de nuestro ejemplo efectuó una inferencia válida, y aunque creía que su inferencia era válida, no lo **sabía**. Carece de un conocimiento de segundo orden sobre la fuente de su conocimiento. Cree correctamente que su inferencia fue válida, pero no lo sabe. No sabe que ha probado lo que cree correctamente haber probado.

Hay por lo menos otro problema sustancial para quienes quieran defender este tipo de epistemología. Por un lado, tenemos el problema de que si una prueba es una inferencia (o una serie de inferencias), hay pruebas tan largas que la inferencia sería humanamente imposible, y parece tener poco sentido hablar de una inferencia que no se puede realizar y que, sin embargo, tenga algún valor epistémico.

Una manera de tratar de rescatar la intuición de que el estudiante de nuestro ejemplo no sabe el teorema que ha probado, aunque tenga una prueba correcta de él, y que pruebas como la del mapa de cuatro colores tienen valor epistémico, es pedir de un sujeto, para saber la verdad de un teorema matemático, no sólo que posea o haya realizado una prueba, sino que también **sepa** que dicha prueba sea correcta. Sin embargo, esto debilitaría la motivación central detrás de esta opción, la cual era establecer una analogía fuerte entre la epistemología de las matemáticas y el resto de la epistemología. Por el contrario, lo que se establecería sería una separación infundada entre el papel epistémico de la prueba matemática y la inferencia

en el resto del conocimiento, donde es ampliamente aceptado que -- excepto en casos extraordinarios -- la inferencia transmite conocimiento.

## B. Casos Gettier Morales

Ya he argumentado que, dadas las fuertes similitudes entre los ámbitos de lo moral y lo epistémico (Barceló 2012), debemos esperar encontrar casos de Gettier en el caso de juicios morales, es decir, casos en los que se actúe con malas intenciones, y en general, el agente haga todo lo que esté en su disposición para producir un mal,<sup>16</sup> el mal que se buscaba se de hecho como efecto de los esfuerzos y actos del agente y, sin embargo, haya algún sentido o perspectiva desde la cual el acto no es moralmente reprobable. Esta tercera perspectiva, he argumentado, es la perspectiva del vínculo adecuado entre las intenciones (y los preparativos, esfuerzos, etc. del agente) y los resultados de la acción. Así como un caso Gettier en epistemología es uno donde uno tiene una creencia verdadera justificada que no es conocimiento porque la justificación y la verdad no están adecuadamente vinculadas, un caso Gettier moral sería uno donde se tienen malas intenciones, un mal resultado, y sin embargo, no estaríamos dispuestos a calificar la acción de mala en un tercer sentido (aún no explorado adecuadamente en la ética) porque el resultado no está adecuadamente ligado a la intención. En otras palabras, si mi hipótesis central de este libro es correcta, es decir, si todos los ámbitos normativos tienen la misma estructura, es de esperar encontrar casos de Gettier en todos los ámbitos normativos, incluido el moral.

Recientemente, Carolina Sartorio (2015) ha publicado un interesante artículo sobre lo que ella llama una "nueva forma" de suerte moral. Su caso proviene de la literatura sobre el papel de la causalidad en el derecho, y ha llegado a ser conocido como el acertijo del "viajero en el desierto". La versión de Christopher Taylor y Daniel Dennett es la siguiente:

“Todos en el puesto de avanzada de la Legión Extranjera Francesa odian a Fred y lo quieren muerto. La noche anterior a que Fred realice una caminata por el desierto, Tom envenena el agua de su cantimplora. Luego, Dick, sin saber de la intervención de Tom,

---

<sup>16</sup>. Por economía, de ahora en adelante en esta sección, cuando hable de intenciones me referiré no sólo a las intenciones, sino también a los esfuerzos, habilidades, capacidades, etc. del agente relevantes para determinar que el agente hizo todo lo que estaba de su parte para lograr sus nefastos fines.

vierte el agua (envenenada) y la reemplaza con arena. Por último, Harry llega y perfora la cantimplora, para que el "agua" se escape lentamente. Más tarde, Fred se despierta y emprende su viaje, provisto de su cantimplora. Demasiado tarde se da cuenta de que su cantimplora está casi vacía, pero además, lo que queda es arena, no agua, ni siquiera agua envenenada. Fred muere de sed.” (Christopher Taylor and Daniel Dennett 2000. Mi traducción)<sup>17</sup>

Este se trata de un caso de Gettier moral. Lo que hizo Tom es malo desde la perspectiva de sus intenciones y medios (el análogo moral de la justificación epistémica), es decir, Tom ha hecho lo suficiente de lo que le correspondía como para estar justificado en esperar que Fred muriera de sed como resultado de sus acciones. Así como la justificación es falible, esta manera de actuar mal no garantiza plenamente que resulte el mal esperado. Además, Fred murió de sed, lo que significa que se alcanzó el objetivo de Tom al envenenar el agua de Fred. Sin embargo, no diríamos que Tom logró matar a Fred. La muerte de Fred no fue un logro de Tom (y mutatis mutandi para Harry o Dick) porque no hubo un vínculo apropiado entre las acciones de Tom y el resultado deseado, es decir, porque Fred no murió de sed debido a que Tom envenenó a su agua, sino por suerte.

Carolina Sartorio tiene razón al señalar que este es un caso de suerte moral en el sentido de Nagel, pero substancialmente diferente de los que suelen presentarse en la literatura (donde una acción es mal intencionada pero tiene buenas consecuencias o viceversa), porque aquí la acción es mal intencionada, produce un mal, pero no es mala en el sentido de que el mal que se produce no es un logro basado en las malas intenciones del agente. Por lo tanto, podemos mantener la hipótesis central de este libro de que todo campo normativo involucra (por lo menos) estas tres perspectivas de evaluación distintas que corresponden, en el caso epistémico, a la justificación, la verdad y el conocimiento. En el caso moral, si tengo razón en mi

---

<sup>17</sup>. Everybody in the French Foreign Legion outpost hates Fred and wants him dead. During the night before Fred's trek across the desert, Tom poisons the water in his canteen. Then, Dick, not knowing of Tom's intervention, pours out the (poisoned) water and replaces it with sand. Finally, Harry comes along and pokes holes in the canteen, so that the "water" will slowly run out. Later, Fred awakens and sets out on his trek, provisioned with his canteen. Too late he finds his canteen is nearly empty, but besides, what remains is sand, not water, not even poisoned water. Fred dies of thirst.

interpretación del acertijo del viajero por el desierto, la acción de Tom es justificadamente malvada (actuó mal), tiene un mal resultado pero no hay un vínculo adecuado entre la justificación maligna y el mal resultante, por lo que no estamos hablando de un acto malo pleno, pues a fin de cuentas Tom no logró matar a Fred.

## G. Hugo Padilla sobre Gettier

Hugo Padilla (1971) ha sugerido que parte de la razón por la cual durante tanto tiempo se nos escaparon casos como los de Gettier es lingüística. En particular se debe a que, en muchos contextos normales, cuando alguien afirma una frase de la forma “ $P$  y  $Q$ ” (para dos enunciados  $P$  y  $Q$ ), los escuchas entienden que la persona ha tratado de comunicar algo más que la mera conjunción de  $P$  y  $Q$ , ya sea que  $Q$  sucedió después de  $P$  o que fue un efecto o consecuencia de  $P$ . Es por ello, piensa Padilla, que muchos filósofos les cuesta trabajo separar las hipótesis de que el conocimiento es creencia verdadera y justificada de la hipótesis de que el conocimiento es creencia que es verdadera *porque* está justificada, y así darse cuenta de que los ejemplos con los que Gettier ilustra su argumento son contraejemplos de la primera hipótesis pero no de la segunda.

El punto, sin embargo, es más general; nos muestra la importancia de ser muy cuidadosos al presentar casos tipo de Gettier, tanto dentro como fuera del campo epistemológico. Consideremos un ejemplo como el de los posibles ataques en Berlín. Si presentamos los hechos del caso así:

1. Hubo un intento de ataque con camiones.
2. Estábamos preparados para un ataque con camiones.
3. El intento de ataque falló.

Sería decir algo verdadero, pero engañoso pues sugiere fuertemente que el fallo se debió a los esfuerzos de las fuerzas anti-terroristas de Berlín, en vez de la incompetencia de los atacantes, lo cual es falso. Pero esto no se debe a ninguna cuestión profunda sobre la naturaleza de la normatividad o en particular sobre la falibilidad de las medidas adoptadas en Berlín, sino más bien a imprecisiones en nuestro lenguaje. Siguiendo la sugerencia de Padilla, si mi propuesta es correcta, aquí hay una especie de falacia de equivocación, la cual se hace aún más manifiesta cuando se hace explícito donde sería el ataque, por ejemplo, así:

1'. Hubo un intento de ataque con camiones (en la zona despoblada X de Berlín).

2'. Estábamos preparados para un ataque con camiones (en la zona poblada Y distinta de X).

3'. El intento de ataque falló.

Así presentado el ejemplo, desaparece la sugerencia de que el intento falló gracias a que estábamos preparados y nos hace preguntarnos ¿porqué falló el intento de ataque en X? En otras palabras, mientras que la descripción en (1), (2) y (3) parece completa, la de (1'), (2') y (3') es claramente incompleta.

Lo mismo se aplica a los otros ejemplos que hemos ya visto:

1. Hubo una inundación.

2. Estábamos preparados para una inundación.

3. Sobrevivimos.

Presentar así los hechos es decir cosas verdaderas pero engañosas pues se sugiere que la razón por la cual sobrevivimos es porque estábamos preparados. Sin embargo, una descripción mas adecuada – no porque la primera no fuera verdadera, sino porque escondía elementos relevantes para darse cuenta de lo que realmente sucedió sería:

1'. Hubo una inundación (de diez días).

2'. Estábamos preparados para una inundación (de cuatro o cinco días).

3'. Sobrevivimos.

Presentar así los hechos nos hace preguntarnos inmediatamente ¿cómo sobrevivimos si no estábamos preparados para una inundación de diez días? En otras palabras, presentar así las cosas muestra que lo que falta es precisamente una conexión entre lo que de hecho sucedió y *para qué* estábamos preparados y que esta conexión se puede perder cuando la posibilidad que se actualiza en los enunciados de tipo (1) – y, por lo tanto, en la que se funden los hechos descritos en los enunciados de tipo (3) – no pertenecen a las posibilidades de las que dependen los hechos descritos en los enunciados de tipo (2).

Si mi análisis es correcto, lo mismo debe suceder en los ejemplos que usa Gettier. Tomemos el primer contraejemplo. En él, Smith y Jones han solicitado el mismo empleo. El gerente de la empresa le ha

asegurado a Smith que Jones se quedará con el empleo. Luego, Smith sabe que el número de monedas que Jones tiene en el bolsillo son diez, porque las ha contado (de manera competente). Finalmente, sin que Smith mismo lo sepa, él también tiene diez monedas en el bolsillo y será quien se quedará con el empleo. Una vez mas, presentar este caso de la siguiente manera es tramposo:

1. El hombre que se quedará con el empleo tiene diez monedas en su bolsillo.
2. Smith cree justificadamente que el hombre que se quedará con el empleo tiene diez monedas en su bolsillo.
3. La creencia de Smith de que el hombre que se quedará con el empleo tiene diez monedas en su bolsillo es verdadera.

Es engañoso, no porque sea falso – tanto (1) como (2) y (3) son verdaderas en la situación descrita por Gettier – sino porque sugiere fuertemente que el que Smith haya logrado tener una creencia verdadera es consecuencia de su justificación (o, si se quiere, de aquello en lo que descansa su justificación). Una vez mas, sostiene Padilla, tenemos una equivocación que se manifiesta con claridad una vez que describimos el caso con mayor detalle, así:

1. El hombre que se quedará con el empleo (Smith) tiene diez monedas en su bolsillo.
2. Smith cree justificadamente que el hombre que se quedará con el empleo (Jones) tiene diez monedas en su bolsillo.
3. La creencia de Smith de que el hombre que se quedará con el empleo tiene diez monedas en su bolsillo es verdadera.

Esto deja claro que la situación relevante en (1) – el que Smith se quedará con el empleo y tiene diez monedas en la bolsa – no pertenece al conjunto de situaciones previstas por Smith en su justificación (2) – el contar las monedas en la bolsa de Jones, por ejemplo, no le da ninguna información sobre las monedas en su propia bolsa (la justificación de Smith es insensible a los hechos que hacen a su creencia verdadera). Esto fortalece la tesis que he venido defendiendo de que todos estos casos son estructuralmente idénticos.

## D. El Lenguaje como Tecnología

Antes de terminar de estudiar la falibilidad, quisiera presentar un ejemplo de importancia filosófica en el que, como espero mostrar a lo largo del libro, es útil hablar de falibilidad y no suele hacerse: las palabras. No seré el primero en sostener que el lenguaje es una tecnología y que las palabras son sus herramientas, pero me parece que las consecuencias filosóficas de esta afirmación aún no han sido exploradas del todo. Para los objetivos de este libro, es útil pensar así a las palabras para poder entender los límites de la semántica tradicional y cómo adoptar un anti-infalibilismo nos permitiría superarlos.



Existe una manera tradicional, y de sentido común, de pensar el lenguaje, entre otras cosas, como un medio de representación, es decir, como una manera de representar cómo es el mundo. Los predicados, por ejemplo, nos sirven para describir maneras en que las cosas pueden ser o no ser. Hay cosas que son verdes y otras que no lo son y la función del predicado “verde” es captar esta diferencia. El significado de “verde”, pues, es la propiedad que tienen las cosas verdes y que las distingue de las que no lo son. En consecuencia, aprender el significado de “verde”, es adquirir la capacidad de usarla para marcar dicha diferencia.

En oposición a esta manera de concebir el lenguaje, filósofos del lenguaje ordinario como Wittgenstein o Strawson nos invitan a pensar al lenguaje como una tecnología y a las palabras como sus herramientas. Como tales, son instrumentos que usamos para diferentes fines, entre ellos, el de

representar cómo es el mundo. Pero rechazan que esto se logre meramente a través de una correspondencia directa entre palabra y realidad de tal manera que, por ejemplo, a cada predicado le corresponda una propiedad y a cada nombre un objeto. En su lugar, proponen pensar al lenguaje en su uso. Esto significa, entre otras cosas, que cuando decimos que usamos al lenguaje para comunicar y representar al mundo lo que queremos decir es que comunicar y representar al mundo son cosas que los humanos hacemos con la ayuda del lenguaje, es decir, que las palabras son las herramientas que usamos para realizar esta tarea y como tales hay que pensarlas. En este sentido, en vez de decir que los nombres refieren a objetos, lo más correcto sería decir que somos nosotros los que nos referimos a objetos usando sus nombres. De la misma manera, el significado de un predicado como “verde”, por ejemplo, no es la presunta propiedad que tienen las cosas verdes y que las distingue de las que no lo son, sino el uso que le damos a dicha palabra para distinguir ciertas cosas de otras. Como Charles Travis (1985) de manera célebre ya señaló, probablemente no haya ninguna propiedad que tengan en común una fruta cuando es verde, un semáforo cuando está en verde ni el planeta tierra que haga adecuado que a todas ellas les apliquemos el adjetivo “verde”.

Esto no significa, por supuesto, que las palabras no tengan significado y que podamos usarlas de cualquier manera. Si alguien me pregunta cómo me pareció la última película que vi, no puedo usar el adjetivo “verde” para comunicar que me pareció muy bien lograda y fiel al libro en el que está basada. Para explicar porque podemos usar las palabras para lo que de hecho las usamos es esencial apelar a sus significados convencionales –el giro pragmático en filosofía del lenguaje no propone lanzar todos los diccionarios a la hoguera – pero este significado convencional no puede por sí mismo fijar de manera completa, para cada posible circunstancia de uso, a qué objetos nos podemos referir y a cuales no. A esto es a lo que me referiré cuando diga que las palabras como herramientas no son instrumentos infalibles de referencia.

Ahora bien, para entender cabalmente qué significa sostener que las palabras son herramientas, uno debe tener claro qué son y cómo funcionan las herramientas. Por ejemplo, si decimos que conocer el significado de una palabra es saber usarla como herramienta, esto significa que para poder desarrollar una teoría del significado de las palabras necesitamos saber qué implica saber usar una herramienta en general. Tomemos un ejemplo sencillo de herramienta: un martillo.

Supongamos que alguien quiere enseñarle a otra persona a usar un martillo. Probablemente lo que haga sea demostrarle como se agarra, como se toma el clavo, y como se levanta para tomar fuerza y ejercerla sobre la cabeza del clavo para que este se inserte en la superficie sobre la que queremos clavarlo, etc. Sin embargo, imaginemos que el aprendiz le pregunte a su maestro: “Muy bien, veo lo que hiciste ahí, pero no me has dicho exactamente cuanta fuerza debo aplicar ni cuanto exactamente debo levantarlo, o siquiera me has dicho de qué tamaño deben ser los clavos.” Es posible que uno trate de resolver la preocupación del aprendiz explicando como es necesaria la suficiente fuerza para vencer la resistencia de la superficie a clavar o como diferentes tamaños y tipos de claves se usan para diferentes materiales y con diferentes funciones. Pero, ¿qué sucedería si el aprendiz insistiera en que fuéramos mas precisos? Uno podría decir, bueno, si insertas el clavo de manera perpendicular al grano de la madera, debes meterlo por lo menos 12 veces el valor de su diámetro, Por ejemplo, si usas un clavo de 4mm de diámetro, debes introducirlo en la madera por lo menos 4.8 centímetros. Si no penetra lo suficiente, la pieza se puede zafar, mientras que si lo introduces demasiado, la madera se puede partir.

Hasta aquí todo suena muy sensato. Sin embargo, es posible imaginar un extraño aprendiz que quisiera insistir en preguntar exactamente cuanto debe entrar del clavo en la madera. “Me hablas de ‘demasiado’ y ‘lo suficiente’, pero ¿exactamente cuanto es eso? – puede insistir nuestro extraño aprendiz – ¿debo introducirlo exactamente 12 veces su diámetro?” En este caso es claro que uno nunca va a poder dar una respuesta absolutamente exacta a su aprendiz y va a llegar un momento en que deba reconocer que estas reglas son aproximadas y que o bien no hay tal cosa como la distancia exacta que el clavo deba introducirse a la manera o que si la hay uno no la conoce. Sería aun mas extraño, pero concebible que al escuchar esto el aprendiz concluya entonces que, en realidad, el maestro no sabe como usar el martillo, en tanto no sabe cuanto – es decir, exactamente cuanto – debe introducirse el clavo en la madera. Aun peor sería si el aprendiz concluyera que el problema es con el martillo (y el clavo) y sostuviera que es inservible en tanto no se conoce su uso y que es necesario desarrollar una nueva herramienta, una especie de martillo perfecto, para el cual sí haya parámetros exactos y conocidos de cuanta fuerza debe ejercerse, de qué tamaño debe ser el clavo, cuanto debe penetrar cada material, etc.

Pero por supuesto que no hay ningún problema con el martillo ni con nuestro conocimiento de su uso. Sabemos usar el martillo aun cuando este conocimiento no pueda expresarse sino con reglas aproximadas. Cuando aprendemos a usar el martillo, como casi cualquier herramienta, no lo hacemos a través de reglas exactas, sino a través de una combinación de demostración de uso, adquisición de reglas aproximadas, ensayo y error.

Ahora bien, para los filósofos del lenguaje ordinario, la comunicación lingüística es como la carpintería y las palabras son como los clavos y el martillo, nuestras herramientas. Para saber usarlas no necesitamos reglas precisas. Por ejemplo, no necesitamos saber exactamente a qué temperatura debe estar un objeto para saber si se le aplica la palabra “caliente” o no. No necesitamos saber exactamente qué tan lejos debe estar un objeto para poder decirle “ese” en vez de tener que decirle “este”, para decir que está “allá” en vez de “aquí”. Y si hubiera un objeto en esa area gris para la cual no tenemos claro si es apropiado decirle “aquí” o “allá”, basta con buscar otra manera de identificarlo, ya que para eso tenemos otras palabras y otros mecanismos. El que haya una zona gris entre el “aquí” y el “allá” no denota ningún problema con estos términos, ningún defecto que los haga inservibles, aunque sí denota cierta limitación a su uso. Denota que no es posible determinar, en cualquier situación, para cualquier objeto espacial, si es adecuado describir su ubicación usando “allá” o no (y lo mismo sucede con “aquí”), es decir, denota que como herramienta no es infalible sino que tiene limitaciones. Pero conocer este tipo de limitaciones es parte del saber usar dichos términos, aun cuando este conocimiento sea él mismo vago – es decir, aunque tampoco podamos decir exactamente a qué distancia de nosotros se encuentra esa línea gris. Sería igual de necio que el aprendiz de carpintero de nuestra historia anterior el filósofo que insistiera en reglas exactas para el uso de nuestras palabras. Probablemente es una cuestión de control la que lleva a algunos filósofos del lenguaje a tomar una actitud frente al lenguaje similar a la que toma nuestro aprendiz de carpintero frente al clavo y al martillo.

Pero lo importante para nuestros objetivos de pensar a las palabras como herramientas es que nos permite entender mejor su aspecto normativo. Cuando decimos que el significado de una palabra es su uso, obviamente no queremos decir que podemos usar cualquier palabra cómo queramos y que cómo sea que la usemos, la habremos usado bien, en tanto ése es su significado.

Como cualquier otra herramienta, las palabras pueden usarse bien o mal, y por lo general, no basta la intención de ningún usuario individual para hacerla hacer algo para o que no tiene la capacidad. Así como mi propia voluntad no es suficiente para convertir esta extensión eléctrica en un sacapuntas, así tampoco puedo decidir un día usar la palabra “máscara” para referirme al aroma de la tierra recién mojada tras un chubasco. En tanto el uso de las palabras es comunitario, como bien insistió tanto Wittgenstein en su obra tardía, se necesitan por lo menos dos para usar una palabra. Pero es sólo en esa situación que se pone a prueba su efectividad y eficiencia como herramienta comunicativa. En tanto herramientas de coordinación humana, las palabras tienen significados públicos y externos, es decir, al igual que cualquier otra herramienta, las palabras obtienen su significado en respuesta a presiones externas, materiales y sociales en un proceso complejo, dinámico y fragmentado. Y aunque los significados son externos en este sentido, también son internalizados por sus usuarios en tanto les sirven para alcanzar sus fines de comunicación y coordinación.

En el modelo tradicional, saber qué significa una palabra significa saber en qué situación es correcto usarla, de tal manera que si, por ejemplo, llamo “verde” a algo que no es verde, este error puede deberse a una de dos cosas: a que pienso equivocadamente (o quiero hacer creer a los demás que pienso) que es verde o a que me equivoco sobre lo que significa “verde”. En otras palabras, si yo sé que algo no es verde y aun así le llamo “verde” esto se debe a que no sé bien qué significa la palabra verde. En el modelo de palabras como herramientas, en contraste, saber qué significa una palabra significa saber en qué situación es correcto usarla, de tal manera que si uso la palabras “verde” por ejemplo, para tratar de describir un objeto, algo habra salido mal si no logro que mi escucha identifique la propiedad a la que me refiero; pero esto no significa necesariamente que el error haya sido resultado de mi falta de pericia en el uso del lenguaje, es decir, no necesariamente debe evidenciar que no conozco el significado de verde o que no soy capaz de detectar correctamente el color del objeto del que hablo. Tal vez puede ser que tuve mala suerte y cosas que razonables asumí sobre el contexto no eran tal y como las pensaba. En otras palabras, surge una tercera fuente de posible error que no proviene, ni del significado, ni del referente. Proviene de la suerte.

El significado de una palabra garantiza la referencia exitosa sólo en condiciones normales, es decir racionalmente previsible. El que estamos en una situación racionalmente previsible de

comunicación es lo que comúnmente llamamos, en filosofía del lenguaje y lingüística, una presuposición. En ausencia de la presuposición, la referencia escapa de nuestras manos y se convierte en un asunto de suerte. Esto significa que es posible que haya errores de comunicación aun en el intercambio comunicativo de dos personas bien informadas y hábiles en su uso del lenguaje.

# V. Deontologismo y Consecuencialismo

## A. El Valor de Lograr Algo

Habíamos señalado ya que reconocer la falibilidad de nuestros logros nos obliga a distinguir entre lograr algo y obtenerlo por mera suerte. Esto nos ha llevado a dos importantes preguntas filosóficamente sustanciales: la cuestión *fundacional* de porqué valoramos más un logro (debido al esfuerzo propio o al ejercicio de alguna capacidad) que el obtener lo que deseamos de manera accidental o por buena suerte, mientras que castigamos igual los casos en los que el mal es resultado de mala suerte que cuando no lo es (Martin & Cushman 2015), y la cuestión *descriptiva* de explicar en qué consiste exactamente el que lograr un objetivo se deba al ejercicio de alguna capacidad o sea el resultado directo de hacer un esfuerzo para alcanzarlo. El



capítulo pasado estuvo dedicado a la pregunta fundacional, mientras que éste lo dedicaremos a dar respuesta a la cuestión descriptiva, es decir, explicar exactamente qué tipo de relación debe haber entre el esfuerzo o ejercicio de una capacidad y la obtención del objetivo para decir que alcanzar dicho objetivo fue un logro y no un afortunado accidente. Nótese que para ello no basta que el resultado deseado sea causado

por la posesión de la capacidad o la realización del esfuerzo. Regresemos al ejemplo del que trabaja para comprarse un coche. Supongamos que para no gastarse el dinero que iba guardando para comprar su coche, nuestro personaje abre una cuenta bancaria de esas que automáticamente te registra a una lotería. Si se gana el coche en dicha lotería, habrá una relación causal entre su esfuerzo por ahorrar para comprarse el coche y la obtención de dicho coche. Sin embargo, es claro que su obtención del coche no se debió de manera directa a su esfuerzo o al ejercicio de alguna capacidad propia. En otras palabras, no fue un logro personal, sino producto de la suerte. Adoptemos, como es costumbre en la literatura, la convención de llamar “consecuencia” a la relación que hay entre un esfuerzo virtuoso por lograr algo y el logro que se obtiene de él (Berker 2013). El problema fundacional, entonces, es preguntarse porqué valoramos más lograr algo como

consecuencia de nuestro esfuerzo que por suerte, y el problema descriptivo es explicar precisamente en qué consiste esta relación de consecuencia.

Respecto a la pregunta fundacional, en un extremo tenemos lo que llamaré el “**consecuencialismo radical**” – que no es sino lo que Hilary Kornblith llama “consecuencialismo burdo” y Bortolotti (2016) llama “universalismo” – según el cual es un error (muy generalizado) el pensar que el hacer el esfuerzo para lograr algo o tener la capacidad de lograrlo tengan algún valor, cuando lo único que tiene realmente valor es aquello que se logra (Gardner 2007, Littlejohn *en prensa*). otras palabras, el esfuerzo que no termina en éxito no vale nada. Por lo tanto, para este tipo de consecuencialismo, no debemos valorar diferente un logro de un accidente afortunado.

Más generalizado es un tipo más moderado de consecuencialismo, según el cual sólo las metas tienen un valor fundamental, y los logros y medios confiables sólo tienen valor de manera derivada en tanto nos acercan a esas metas valiosas. Este tipo de consecuencialismo también sostiene que es un error muy generalizado el pensar que el hacer el esfuerzo para lograr algo o tener la capacidad de lograrlo tengan algún valor independiente del valor de aquello que logran; sin embargo, no piensan que no tenga ningún valor. Aunque si valoramos la posesión o el ejercicio de una capacidad no es sino porque valoramos obtener aquello que podemos alcanzar con el ejercicio de dicha capacidad. En otras palabras, y tal y como lo caracteriza Berker (2013):

Todas las teorías consecuencialistas [moderadas] sostienen que (i) hay ciertos estados de cosas que son inherentemente buenos, y otros estados de cosas que son inherentemente malos, y (ii) todas las demás nociones normativas están determinadas por lo bien que conducen hacia ellos de manera directa o indirecta, en virtud de estar adecuadamente conectados con alguna otra entidad que conduzca directamente hacia ellos. Si aplicamos esta idea general al ámbito ético y tomamos al bienestar individual como el valor meta (sin tomar en cuenta cómo se distribuye ese bienestar), lo que resulta es el utilitarismo. Si aplicamos esta idea general al reino epistémico y tomamos a la creencia verdadera como el valor meta, lo que resulta es el confiabilismo. Lo que importa epistémicamente, para los confiabilistas, son las consecuencias (en términos de alcanzar creencias verdaderas y evitar creencias falsas) de formar creencias de

una manera determinada. (Berker 2013. Traducción Mia)<sup>18</sup>

En el otro extremo se encuentran filósofos de inclinación Kantiana (pienso en Christine Korsgaard 2000, por ejemplo, pero también Bortolotti 2016, entre otros) para quienes la diferencia fundamental entre el logro y el accidente afortunado es que sólo el primero puede adjudicarse normativamente al sujeto. En otras palabras, para este tipo de *deontologistas* – por llamarlos de alguna manera –, hay una diferencia sustancial entre que algo sea malo/bueno y que alguien haya actuado mal/bien, es decir, que algo sea correcto/incorrecto. Si alcanzamos nuestras metas esto es algo bueno, pero no habremos actuado bien (es decir, de manera correcta) a menos que dicha meta haya sido lograda en base a nuestros esfuerzos y cuidados. Por eso no elogiamos a quienes ganan la lotería (aunque, como veremos a detalle mas adelante, sí asignamos responsabilidad e imputamos castigos a quienes desafortunadamente actúan o causan males sin que estos sean intencionales o el resultado de negligencia por su parte). Para entender mejor este tipo de deontologismo, vale la pena distinguir dos sentidos en los que evaluamos positivamente algo. Por ejemplo, cuando queremos ir de picnic, podemos decir que es bueno que el clima esté agradable, pero al hacerlo no estamos haciendo un juicio moral genuino, ya que no estamos evaluando ningún acto ni ningún agente. No hay nada eminentemente humano en la bondad del clima, aun cuando su bondad dependa de nuestras necesidades y preferencias humanas y que también sea muy humano nuestro aprecio. Sin embargo, cuando decimos que es bueno que alguien haya traído, repelente de mosquitos al picnic, lo que estamos evaluando sí es un acto genuino ya que estamos adjudicando la bondad de un estado conveniente – el que contemos con repelente de mosquitos en

---

<sup>18</sup>. “All consequentialist theories hold that (i) there are certain states of affairs that are inherently good, and other states of affairs that are inherently bad, and (ii) all other normative notions under the theory’s purview are determined by how well the objects of assessment conduce toward or promote those states of affairs (either by directly conducing toward them, or by indirectly conducing toward them, in virtue of being suitably connected to some other entity that directly conduces toward them). If we apply this general idea to the ethical realm and take the value to-be-promoted to be individual well-being (with no regard to how that well-being is distributed), what results is utilitarianism. If we apply this general idea to the epistemic realm and take the value to-be-promoted to be true belief and the avoidance of false belief, what results is reliabilism. What matters epistemically, for reliabilists, are the consequences (in terms of attaining true beliefs and avoiding false beliefs) of forming beliefs in a certain way.”

el picnic – a una persona – a quién lo trajo. Para que este trasladar la bondad de la consecuencia al agente sea correcto, sostienen los deontologistas, es necesario que la consecuencia sea un logro del agente, es decir, es necesario que sea resultado de sus esfuerzos, capacidades, etc. y no de la suerte. Al trasladar esta bondad al agente, se vuelven apropiadas nuevas respuestas que antes no estaban disponibles. Por ejemplo, podemos felicitar o reprender al agente por sus la bondad o maldad de sus actos (y las consecuencias de sus actos), y si nosotros somos el agente, podemos sentirnos orgullosos o avengorizados de ellos. Por eso no significa lo mismo decir que el clima es bueno para un picnic (es decir, es propicio para un picnic) a decir que el clima es el correcto para un picnic (es decir, es correcto irse de picnic en un día así); lo que estamos evaluando en cada caso es algo muy distinto, y sólo en el segundo caso lo que estamos evaluando es un acto.<sup>19</sup>

Finalmente, hay una posición aun mas **radical** en oposición al consecuencialista en la cual los valores con los cuales evaluamos a los medios son completamente independientes de los valores con los que evaluamos los fines, de tal manera que un mecanismo, regla o método puede ser valioso aun si **no** nos acerca a nuestros fines, es decir, si **no** reduce el riesgo de falla. Estoy pensando en filósofos como los ya mencionados Ángeles Eraña y Hilary Kornblith, pero también Catherine Elgin y Sally Haslanger para quienes nuestras obligaciones y nuestro sentido de ‘cuidado’ se han construido a partir de contingencias históricas y/o evolutivas que no son todas responsivas a nuestros fines. En este sentido, por lo tanto, el mero llamarlos medios es ya confundente, pues denota un lugar secundario y derivado con respecto a los fines. Así por ejemplo, dado que nuestras capacidades cognitivas son el resultado de un proceso evolutivo que ha sido el resultado de muchas contingencias históricas que se han dado a lo largo de nuestra historia biológica, es completamente posible que ellas funcionen a la perfección sin llevar de manera confiable a la verdad. Extendiendo esta idea de Eraña y Kornblith a otros ámbitos, podemos decir que los métodos que adoptamos

---

<sup>19</sup>. Dado que a mí me interesa extender este modelo normativo a casos donde lo que queremos evaluar no son actos o estados de agentes, es necesario relativizar este principio también a entidades que no sean agentes. Así pues, aun cuando hablamos de una herramienta, por ejemplo, podemos distinguir entre las bondades que podemos apreciar de las consecuencias positivas de usar dicha herramienta, de las que podemos ascribir a la herramienta misma por estar basadas en su funcionamiento correcto. Sobre las similitudes y relaciones entre la normatividad apropiada para evaluar a agentes humanos y la normatividad relevante para evaluar herramientas, véase Franssen (2006).

para alcanzar nuestros fines son muchas veces el resultado de contingencias históricas, psicológicas, etc., muchas de las cuales no responden a qué tanto dichos métodos nos acercan o no a nuestros fines. Por ejemplo, es una regla común entre los bibliotecarios de Estados Unidos poner libros de calidad en la sección de libros para niños para formar lectores mas exigentes, aunque no hay muchos datos empíricos rigurosos que sustenten dicha regla (y hay razones para pensar que no sirve de nada para mejorar la calidad de los libros que los niños que asisten a la biblioteca de hecho leen, Edmonds 1987); esto significa que un bibliotecario responsable en Estados Unidos pondrá libros de calidad en los libreros de la sección para niños aunque es muy posible que esto no contribuya en nada a su fin, a saber, formar lectores mas exigentes. Esto se debe a contingencias históricas y sesgos cognitivos que han llevado a los bibliotecarios Estadounidenses a adoptar esta regla (y a convencerse de su valor) independientemente de su confiabilidad o contribución al fin al que se supone está dirigida. Sin embargo, eso no hace inadecuado, en ausencia de información extra (por ejemplo, de que de hecho dicha práctica no sirve para formar lectores más exigentes), el juzgar negativamente al bibliotecario que no siga dicha regla. Aunque se deba a contingencias históricas y sesgos cognitivos, esta regla sigue siendo parte de lo que es ser un bibliotecario responsable en el contexto actual de los Estados Unidos.

Así pues, a la pregunta de ¿porqué valoramos mas los logros que (el mero alcanzar) las metas?, consecuencialistas y deontologistas dan respuestas radicalmente distintas. Para el consecuencialista radical, como habíamos mencionado, no debemos valorar diferente un logro de un accidente afortunado. Ahora bien, la razón por la cual *parece* que valoramos los medios confiables y los logros es porque en realidad nos dan evidencia, derrotable pero confiable, de que se darán los fines, los cuales son los que realmente tienen valor. En otras palabras, lo único que realmente nos interesa es alcanzar nuestros fines; sin embargo, la mayoría de las veces no sabemos si obtendremos o no lo que deseamos, a lo único que tenemos acceso es a nuestros medios. Por eso buscamos en ellos información, indicios sobre lo que nos interesa en realidad, es decir, si obtendremos nuestros fines o no. En ausencia de conocimiento sobre los fines, apelamos a información sobre la confiabilidad de los medios para hacer nuestros juicios evaluativos tentativos. Sin embargo, esos juicios siguen siendo acerca de los fines, no acerca de la confiabilidad de los medios. La clave aquí es distinguir claramente entre que algo sea bueno o malo y que algo esté conectado evidencialmente con el que algo bueno o malo haya pasado o vaya a pasar. Por ejemplo, si no sabemos si alguien actuará bien o

mal, información sobre sus intenciones nos puede servir para realizar un juicio tentativo sobre sus actos ya que de sus intenciones podemos inferir; derrotablemente, el tipo de actos que podemos esperar de él o ella. Pero al hacerlo, no estamos evaluando sus intenciones, sino que estamos usando sus intenciones como información evidencial para juzgar lo que realmente nos interesa: sus actos. De la misma manera, antes de saber si lo que alguien dice es verdad o falsedad, podemos servirnos de sus razones para adelantarnos en nuestro juicio sobre lo que dirá y anticipar la verdad de lo que dirá, pero otra vez no estaremos valorando sus razones, sino usándolas como evidencia de lo que nos interesa, a saber, la verdad de lo que dice. En consecuencia, una vez que sabemos si lo que dice es verdadero o falso, sus razones se vuelven insustanciales. Igualmente, para el consecuencialista radical, para poner otro ejemplo, parece que valoramos el dinero por sí mismo, cuando en realidad lo que nos interesa es lo que podemos comprar con él y si parece que valoramos el dinero, en realidad, es porque al poseerlo anticipamos el valor de lo que conseguiremos a través de él. En todos los casos, los medios no tienen valor alguno, pero sí pueden darnos información sobre si podemos esperar o no aquello que sí es de valor, es decir, las metas.

Para el consecuencialista moderado, en contraste, los medios sí pueden tener algún valor, pero éste no es completamente independiente del valor de los fines, sino que se deriva de él. Esto significa que los medios confiables son valiosos precisamente porque nos acercan a nuestros fines de manera confiable. Pero este valor, aunque derivado, es genuino, en el sentido de que un medio confiable es valioso aun en esos casos en los que, de hecho, no nos lleva a alcanzar nuestros fines. Así pues, al buen cazador que se le escapa la liebre le podemos reconocer el valor instrumental de su habilidad como cazador, aun en aquellos casos en que dicha habilidad haya sido infructuosa. En consecuencia, el alcanzar nuestras metas por mera suerte tiene la desventaja, sobre el lograr nuestras metas a partir de nuestros esfuerzos, cuidados, etc., de ser un éxito endeble, frágil, uno que no podemos esperar se repita; lo que explica su inferioridad frente al logro, el cual es robusto, repetible y confiable.

Finalmente, el deontologista propone una especie de dualismo en el cual las metas tienen un tipo de valor y los logros otros.<sup>20</sup> Las metas tienen un valor similar al que tienen otras cosas que nos pueden agradar o convenir, como la lluvia u otras cosas buenas de la vida que no siempre podemos adjudicar a algún agente;

---

<sup>20</sup>. En realidad, como hemos visto, no son dos regímenes de valor distintos, sino tres: uno que corresponde a las metas, otro a los medios y uno tercero a los logros.

mientras que los logros tienen un valor eminentemente humano. Alcanzar una de nuestras metas, no a través de la manifestación de nuestras capacidades, habilidades o esfuerzos, sino por buena suerte es bueno en el mismo sentido en que es bueno el buen clima o el haber nacido en una familia de dinero y al igual que ellos no podemos adjudicarlo a nadie; mientras que lograr nuestras metas es bueno en otro sentido, uno que nos atañe directamente como seres humanos, uno que dice algo sobre nosotros, sobre nuestro carácter. Esto explica porque preferimos los logros a (el mero alcanzar) las metas: porque los primeros son virtuosos por partida doble: por un lado, son cosas buenas, y por el otro, son actos buenos, es decir, podemos rastrear su bondad a los actos correctos del agente. En otras palabras, aparecen como valiosos dentro de dos regímenes normativos distintos: el de las cosas y el de los actos.

En resumen, el debate entre deontologistas y consecuencialistas corresponde a dos maneras diferentes de concebir la normatividad de los medios y su relación con la normatividad de los fines. Desde la perspectiva deontológica, una persona está justificada en tomar un riesgo si, ya sea de manera consciente o automática, toma las medidas necesarias para actuar con el debido cuidado. Desde la perspectiva consecuencialista, una persona está justificada en tomar un riesgo si, dadas las probabilidades de obtención de cada resultado, el cálculo de costos y beneficios favorecen la toma del riesgo. Aunque diferentes, suele considerarse que ambos criterios están íntimamente relacionados: que la razón por la cual es bueno tomar cuidado en lo que hacemos, es porque así reducimos objetivamente el riesgo de no lograr nuestros objetivos. Igualmente, se espera que una explicación consecuencialista de lo que hace tomar un riesgo racional o no capture nuestra intuiciones acerca de lo que es racional esperar de un agente cuidadoso. Además, una vez que introducimos la distinción tripartita entre alcanzar un fin, hacer lo correcto para conseguirlo (es decir, estar justificados en esperar el fin), y lograr un fin (es decir, alcanzarlo *porque* hicimos lo correcto para conseguirlo), tiene sentido preguntarse cual de estas tres nociones es fundamental. Tradicionalmente, los consecuencialistas piensan que alcanzar la meta es lo que único tiene valor en sentido fundamental, mientras que el logro y la justificación sólo lo tienen en sentido derivado, a lo más; mientras que los deontológicos

piensan que el valor de la justificación no se deriva del valor de la meta, sino que es valioso en sí mismo.<sup>21</sup>

## B. Un ejemplo matemático

Consideremos ahora un caso presentado por Selim Berker en su artículo de (2013), y que pretende poner en problemas a los consecuencialistas, especialmente a aquellos epistemólogos que sostienen que para que una creencia sea verdadera es suficiente que, además de ser verdadera, sea generada por un proceso que tiene a generar mas creencias verdaderas que falsas:<sup>22</sup>

Supongamos que lo siguiente es verdadero sobre mí: siempre que contemplo si un número natural dado es primo, formo la creencia de que no lo es. “¿Es 25 un número primo? No, no lo es.” “¿Es 604 un número primo? No, no lo es.” “¿Es el 7 un número primo? No, no lo es.” También estipulemos que este es el único proceso cognitivo a través del cual me formo creencias sobre números primos (Soy un tipo ingenuo). Puesto que la proporción de números primos a número compuestos menores que  $n$  se aproxima a 0 cuando  $n$  se acerca al infinito, mi proceso de formación de creencias tiende a producir una razón de creencias verdaderas a falsas que se aproxima a 1. Por lo tanto, los confiabilistas ... se ven obligados a decir que, debido a que mi proceso de formación de creencias es casi perfectamente fiable, cualquier creencia formada sobre su base está justificada. ¡Pero eso es una locura! Cuando creo una creencia de que 7 no es primo, simplemente no es correcto decir que, aunque esa creencia es falsa, es

---

<sup>21</sup>. Por supuesto, cabe la posibilidad que no haya nada detrás del debate entre deontologistas y consecuencialistas mas que preferencias personales de los propios filósofos y filósofas que develan algo de su carácter, como ha sugerido Gilles Deleuze (2001). Según el filósofo francés, el consecuencialista muestra rasgos sádicos, mientras que el deontologista muestra rasgos masoquistas, en tanto el primero concibe a la normatividad como fundamentada en algo externo y el segundo la concibe como emanando de nuestras propias prácticas y situaciones sociales.

<sup>22</sup>. Como he señalado con anterioridad, para los objetivos de este trabajo la distinción entre procesos y mecanismos no es importante, por eso no es importante tampoco la cuestión de si lo que debe ser confiable es el mecanismo o el proceso que genera la creencia.

epistemológicamente redimida por la verdad de las otras creencias que se formarían a través del proceso que la condujo. Así, .... su teoría produce el veredicto epistémico equivocado.<sup>23</sup>

Aún cuando el problema está formulado explícitamente contra un tipo particular de consecuencialista, no es difícil ver cómo se puede generalizar a otro tipo de consecuencialismos epistémicos, y dado que mi modelo pretende recuperar del consecuencialismo epistémico la idea de que es una virtud de nuestros mecanismos el acercarnos a nuestros fines – y que, por lo tanto, una virtud de nuestros mecanismos epistémicos es el acercarnos a nuestros fines epistémicos, a saber, la generación y mantenimiento de creencias verdaderas –, vale la pena explorar cómo trataría un problema como este.

En este respecto, un punto central a destacar del ejemplo de Berker es que es uno de esos confusos experimentos mentales en los que se nos pide suponer que una persona – en este caso, una persona concreta, Berker mismo – es completamente distinto, en aspectos epistémicamente fundamentales, a las personas reales (y en particular, sustancialmente distinto del verdadero Berker de carne y hueso) para luego pedirnos nuestras intuiciones sobre el estatus epistémico de unas creencias que las personas comúnmente no tienen, producidas por un proceso cognitivo que ninguna persona posee. El caso es tan abigarradamente contrafáctico, que es natural cuestionar su relevancia para entender nuestras prácticas de evaluación epistémica reales. Sin embargo, me parece que el ejemplo de Barker **sí** es muy ilustrativo de ciertos aspectos importantes de nuestras prácticas evaluativas, pero no exactamente por las razones que Berker señala. El ejemplo de Barker, me parece, nos puede ayudar a arrojar luz sobre, por ejemplo, por qué nuestras prácticas epistémicas sobre juicios matemáticos – y juicios científicos en general – son tan distintas de aquellas que versan sobre juicios cotidianos; y también sobre la diferencia entre evaluar actos y estados epistémicos humanos y otros estados y

---

<sup>23</sup>. “Suppose the following is true of me: whenever I contemplate whether a given natural number is prime, I form a belief that it is not. “Is 25 prime? No, it is not.” “Is 604 prime? No, it is not.” “Is 7 prime? No, it is not.” Let us also stipulate that this is the only cognitive process by which I form beliefs about the primeness of natural numbers. (I’m a simpleminded kind of guy.) Since the ratio of prime to composite numbers less than  $n$  approaches 0 as  $n$  approaches infinity, my belief-forming process tends to yield a ratio of true to false beliefs that approaches 1. Therefore process reliabilists are forced to say that, because my belief-forming process is almost perfectly reliable, any belief formed on its basis is justified. But that’s crazy! When I form a belief that 7 is not prime, it is simply not correct to say that, although that belief is false, it is epistemically redeemed by the truth of the other beliefs which would be formed via the process that led to it. Thus even process reliabilists countenance cross-propositional trade-offs, and where they do, it seems their theory yields the wrong epistemic verdict.”

procesos representacionales análogos que, por lo tanto, también podrían caer dentro del dominio de aplicación de nuestros valores epistémicos.

Imaginemos el mismo caso aplicado, no a un proceso inferencial humano, sino a uno artificial:

Supongamos que Roberta desarrolla un nuevo algoritmo para realizar una tarea dada; este nuevo algoritmo es más rápido que los algoritmos anteriores gracias a su nuevo sistema de procesamiento de datos, el cual asume que los datos a procesar, al ser codificados numéricamente, nunca corresponden a un número primo. En otras palabras, el algoritmo recibe un paquete de datos codificado en forma de un número, y aun cuando no sabe nada sobre qué número es, asume de entrada que no es primo, es decir, lo registra y procesa como si no fuera primo. En otras palabras, para el algoritmo, el número no es primo. Si el dato de entrada es 2500098405061834, para el algoritmo, no es primo. Si el dato de entrada es 1124098488569631, para el algoritmo, no es primo. Si el dato de entrada es 0000000000000007, para el algoritmo, tampoco es primo. También estipulemos que el algoritmo no cuenta con ningún otro mecanismo o subrutina para verificar si un número es primo o no (es un algoritmo simple). Puesto que la proporción de números primos a número compuestos menores que  $n$  se aproxima a 0 cuando  $n$  se acerca al infinito, el algoritmo tiende a producir una razón de resultados correctos a incorrectos que se aproxima a 1. ¿Cómo debemos evaluar dicho algoritmo?

Aún cuando el nuevo ejemplo es análogo al de Berker en los aspectos relevantes, nuestras intuiciones sobre las virtudes del algoritmo difieren sustancialmente de las que teníamos en el caso anterior. Me parece que en este caso es muy claro que nuestro juicio dependerá, entre otras cosas, de cual es el riesgo real que se corre porque el algoritmo no funcione, ¿estamos hablando de un algoritmo que, de fallar, podría costarle la vida a alguien, por ejemplo, un algoritmo clave dentro de un aparato médico? ¿o estamos hablando de una simple subrutina que a lo más podría ‘crashear’ una aplicación de entretenimiento de nuestro teléfono móvil? (Mackenzie 2001) También dependerá de qué tanto se usará el algoritmo, ¿será un algoritmo que correrá continuamente de tal manera que, podemos esperar casi con certeza que en algún momento fallará? ¿o es un algoritmo de emergencia que sólo se usará de vez en cuando? y, finalmente, dependerá también en lo valioso que sea el tiempo que se ahorrará con el uso del nuevo algoritmo. Si bien es cierto que, por lo general, más rápido suele ser mejor; también es cierto que en muchos casos la velocidad no es importante o hasta puede ser contraproducente. Si el algoritmo no recibe información con suficiente velocidad o si el resto del sistema no es capaz de procesar resultados con la velocidad con la que el mecanismo los produce, por ejemplo, serían dos circunstancias en las que un aumento de velocidad no representa una ventaja sustancial.

Esto significa que, una vez que pasamos de hablar de creencias y procesos cognitivos humanos y tratamos de aplicar el mismo esquema a un mecanismo artificial, las intuiciones consecuencialistas regresan con fuerza. ¿Muestra esto que, por lo tanto, sí hay una diferencia importante entre la manera que evaluamos el funcionamiento de un artefacto que el de un humano, de tal manera que el consecuencialismo puede ser adecuado en el caso de las máquinas, peor no en el de las personas? Creo que no. Más que una diferencia importante entre mecanismos artificiales y humanos, el ejemplo de Berker sirve para reforzar esta semejanza en la que he insistido a lo largo del libro: que nuestras prácticas normativas son sensibles a la disponibilidad de relibros. Y aquí es donde el hecho de que el ejemplo proviene de las matemáticas es importante.

A la hora de evaluar un experimento mental es esencial preguntarnos si la descripción del caso es lo suficientemente completa respecto a los rasgos que son relevantes según las diferentes hipótesis en disputa (Dowell, manuscrito). En tanto mi hipótesis sostiene que nuestras prácticas de evaluación – y, por lo tanto, nuestros reportes de conocimiento – dependen sustancialmente de los relibros disponibles al agente en el contexto (entre otros factores contextuales), ésta sostiene que cualquier descripción de una situación que no incluya esta información estará incompleta y que el que la escuche tratará de llenar esta información con los valores por defecto en el contexto. En el caso del experimento mental de Berker, por ejemplo, no sabemos, entre otras cosas, por qué está Berker pensando en números primos, qué tan urgente o importante es tener información correcta sobre esto, etc. factores que, sostengo, juegan un papel importante en nuestras evaluaciones de conocimiento. Esto hace que los que interpretamos el texto de Berker necesitamos completar la descripción con los valores estándar que se asumen por defecto en nuestro contexto. En este caso, tratamos de identificar la manera *normal* por la cual alguien necesitaría decidir sobre un número natural si debe creer que es primo o que no es primo. En otras palabras, de inmediato nos movemos a un contexto matemático.

La ciencia en general, y las matemáticas en particular, son prácticas epistémicas bastante sui-generis. A diferencia de un gran número de nuestras preocupaciones cotidianas, y con contadas excepciones, las preguntas que se plantean en la ciencia carecen de la menor urgencia. Excepto por ciertos casos de investigación médica, climática, etc. no hay ninguna prisa real en encontrar respuesta a la mayoría de las preguntas fundamentales de la ciencia. Sí, por supuesto, como cualquier quehacer humano, el quehacer científico se realiza en un contexto de fechas límites, proyectos con plazos determinados, evaluaciones de resultados a plazos regulares, etc. Aún cuando hay importantes factores prácticos que determinan la

aceptación de, por ejemplo, ciertos axiomas y conjeturas matemáticas que son valiosos para poder obtener otros resultados importantes, por lo general, no es el caso que necesitemos resolver un problema matemático *ya* y debamos adecuar nuestros estándares de aceptaciones a nuestras limitaciones de tiempo. En otras palabras, cuando se trata de proposiciones matemáticas, el tiempo es un factor desdeñable. Al eliminar restricciones de tiempo, muchos otros relibros materiales dejan también de ser significativos a la hora de determinar estándares epistémicos. Aun si le dedicamos pocos relibros matemáticos a, digamos, la investigación geométrica; los relibros acumulados durante más de veinte siglos de investigación geométrica resultan enormes. Esto significa que las matemáticas, y probablemente esto sea algo que se pueda generalizar al resto de la ciencia, son empresas epistémicas de relibros superlativos. Esto tiene profundas consecuencias epistémicas.

En este libro he sostenido, al igual que muchos otros filósofos, que nuestros estándares epistémicos (y no epistémicos también) son sensible a la disponibilidad de relibros. En particular, cuantas (y cuales) posibilidades debemos tomar en cuenta para determinar si una creencia está justificada o no dependerá de con cuales relibros contemos para realizar nuestra pesquisa. Por lo general, si tenemos más relibros, es nuestra responsabilidad considerar más posibilidades; si tenemos pocos, al contrario, es justificable considerar unas cuantas menos. Esto significa que aquellas investigaciones a las que dediquemos más relibros, tendrán estándares epistémicos más altos. De ahí que las matemáticas y otras ciencias tengan estándares epistémicos tan altos.

Por eso es que, cuando nos preguntamos sobre las circunstancias en las cuales alguien está siquiera justificado en creer una proposición matemática como que 7 es un número primo, automáticamente pensamos en estándares muy altos, casi de certeza, estándares que nuestras prácticas actuales de prueba y aprendizaje sí sobrepasan. No en balde, en nuestro contexto socio-histórico, la manera canónica de obtener conocimiento matemático es a través de las prácticas sociales de aprender y probar resultados matemáticos. Aún tomando en cuenta la falibilidad de estas prácticas – tal y como las hemos detallado en la primera sección de este capítulo –, su confiabilidad es más alta que la altísima confiabilidad de siempre asumir que un número natural es primo. Uno debe ser muy cuidadoso, por lo tanto, cuando usa ejemplos matemáticos en experimentos mentales como los de Berker. Aun cuando nuestras prácticas matemáticas también son el producto de contingencias históricas, sus estándares epistémicos son poco sensibles al contexto. Es por ello

que pensamos, correctamente, que para aceptar una proposición matemática como verdadera debemos exigirle a cualquier persona, en un contexto como el nuestro, nada menos que haberlo probado o aprendido de acuerdo a nuestros estándares de rigor epistémico actuales.<sup>24</sup>

---

<sup>24</sup>. Además, el escoger la pregunta ¿es un número natural dado primo? como ejemplo Berker introduce tres nuevas fuentes de posible confusión: nuestra dificultad de pensar sobre dominios infinitos, el sesgo de representatividad y la asimetría entre el saber que un número es primo y saber que no es primo. Empecemos por esta última. Efectivamente, es mas fácil saber que un número no es primo que saber que es primo, y conforme los números crecen, la diferencia se vuelve mas sustancial. Saber que el número 8 no es primo es muy sencillo, basta darse cuenta de que es par; saber que el número 7 es primo es un poco mas difícil, requiere darse cuenta de que no es par, ni divisible entre 3 y 5. Saber que el número 7918 no es primo es muy sencillo, basta darse cuenta de que es par; saber que el número 7919 es primo es mucho mas difícil. Saber que el número 393050634124102232869567034555427371542904834 no es primo es muy sencillo, saber que el número de Cullen 393050634124102232869567034555427371542904833 es primo ha sido muy difícil de establecer y es muy difícil de aprender. Esto significa que los relibros que necesitamos para determinar si un número natural es primo o no dependen *en gran parte* en de qué número natural estemos hablando. Cuando se nos pide pensar, como hace Berker en su descripción del experimento mental, en los números naturales en general, inmediatamente se nos vienen a la mente número *extremadamente pequeños* dentro del inmenso dominio infinito de los números naturales. A Berker mismo le sucede cuando usa como ejemplos a los números 25, 604 y 7 en vez de números como 567034555427371542904834, 170141183460469231731687303715884105727 o 1317313771737997107113149157167179199311337347359389701709733739743751761769907937941953967971983991. En otras palabras, somos presas de lo que se conoce como el sesgo de disponibilidad, la tendencia que tenemos todos los humanos a considerar como normales el tipo de casos con los que nos hemos encontrado mas comúnmente. Dado que es más probable que nos hayamos topado con números naturales como 7, 25 o 604, es decir, números excesivamente pequeños, que números de cientos o miles de dígitos, cuando pensamos en *un* número natural, el sesgo de disponibilidad nos hace pensar automáticamente en un número de éstos. Esta es otra razón por la cual, cuando se nos piden nuestras intuiciones sobre qué sería necesario para que la creencia de que un número dado, digamos 7, no sea primo, automáticamente

Cabe preguntarse, entonces, si el extraño Berker del experimento mental del Berker real es el tipo de agente al que debe aplicarse este estándar tan alto, o si mas bien es como el algoritmo de Roberta, para el cual podemos considerar aceptable y hasta virtuoso el simplemente asumir que cualquier número natural es primo. Como hemos mencionado, en nuestro contexto socio-histórico, estar justificado en tener una creencia matemática es haber participado de manera responsable en las prácticas epistémicas que constituyen a las matemáticas como una institución, es decir, haberlo aprendido o probado de manera competente. También hemos señalado que, dados los enormes relibros que la humanidad dedica y ha dedicado a esta empresa, sus estándares epistémicos – tanto los que demanda como los que alcanza – son altísimos, mas altos que los del proceso que persigue el extraño Berker del experimento de Berker. Lo que es importante determinar ahora es si podemos demandar de este extraño ser el participar de dichas prácticas.<sup>25</sup> Es decir si tiene disponibles los relibros necesarios para hacerlo. Es aquí donde cobra importancia la estipulación que introduce el Berker real de que éste es el único proceso cognitivo a través del cual el extraño Berker forma creencias sobre números primos. A través de esta estipulación el Berker real le cierra al extraño Berker la puerta que lo podría comunicar con la comunidad epistémica de los matemáticos, pero según contextualistas como yo, también lo exime de la responsabilidad de hacerlo. En otras palabras, dado que nuestras intuiciones sobre los estándares epistémicos apropiados para evaluar creencias matemáticas humanas dependen de consideraciones que no se aplican en el caso del extraño Berker, no nos dicen mucho sobre si la creencia de este extraño ser es conocimiento o no.

Es aquí donde el Berker real nos hace confundir nuestras intuiciones sobre los estándares aplicables a seres humanos como el Berker real, y aplicárselos a seres extraños como el Berker de su ejemplo. Es por ello que lo que nos muestra el experimento mental de Berker es que tenemos buenas razones (que tienen que ver con los enormes relibros dedicados a la matemática) para decir que quien (es decir, cualquier persona que) base su creencia de que 7 es primo en algo menos que un aprendizaje competente o una prueba matemáticas no está justificado en su creencia, lo cual es completamente consistente con el consecuencialismo, o por lo menos con las tesis consecuencialistas básicas que adopto aquí, a saber, que una de las razones por las cuales

---

<sup>25</sup>. A decir verdad, hay casos genuinos en los cuales tiene sentido preguntarse de, por ejemplo, una persona con algunos tipos de discapacidades si efectivamente tiene la capacidad de participar en una institución social epistémica como la de las matemáticas.

preferimos ciertos medios sobre otros, es decir, uno de los criterios bajo las cuales evaluamos nuestros medios es por su capacidad a acercarnos a nuestros fines.

### C. Suerte Semántica y Moral

Los orígenes de este libro se encuentran en un artículo que publiqué en 2012 sobre las fuertes similitudes que encontraba en dos debates aparentemente muy distintos: el debate alrededor de la suerte moral en meta-ética y el debate alrededor del relativismo en la semántica de los modales epistémicos (es decir, expresiones aparentemente modales como “puede que” o “tal vez” cuando se usan para expresar lo que no podemos excluir a partir de lo que sabemos). Después de publicar dicho trabajo, la idea no desapareció de mi mente y poco a poco me fui dando cuenta que las similitudes eran mas profundas y extensas de lo que había pensado originalmente. El propósito central de este libro ha sido precisamente extender los resultados de ese trabajo a otros campos y darles otras aplicaciones. Sin embargo, el ejemplo de los modales epistémicos me sigue pareciendo un muy buen ejemplo para ilustrar el debate entre deontologistas y consecuencialistas, además de dar pie a mi propia propuesta dualista. La propuesta relativista de John MacFarlane que criticó en dicho artículo también me parece una excelente y sofisticada propuesta teórica que se toma en serio el consecuencialismo y criticarla me parece fundamental para fundamentar porqué, después de todo, no soy un consecuencialista.

MacFarlane parte de la intuición básica de los consecuencialistas de que, en muchas ocasiones, lo que nos interesa no está a nuestro alcance cuando lo necesitamos; esto nos obliga a tomar decisiones y hacer juicios en estados de indeterminación y con información limitada. Hasta aquí, MacFarlane sigue la ortodoxia consecuencialista. Sin embargo, lo que distingue la posición de MacFarlane (y autores afines como March Richards) es que en ella, lo correcto incorrecto de estas decisiones y juicios dependerá de la información disponible, no tanto (o no solo) disponible al momento y para la persona que debe tomar la decisión o emitir el juicio, sino disponible al momento y a quién evalúa dicha decisión o juicio. Esto significa que la decisión será correcta o incorrecta dependiendo de la información que dispongamos para evaluarla, es decir, su corrección será sensible al contexto de evaluación. Según este tipo de relativismo, un acto podrá ser correcto desde una perspectiva, pero incorrecto desde otra; conforme conozcamos mas (o menos) sobre sus

consecuencias, el mismo acto podrá pasar de correcto a incorrecto y vice versa. ¿Cuántas veces no nos hemos dado cuenta, conforme aprendemos más sobre una situación, que la decisión que tomamos no era la correcta? MacFarlane quiere hacer justicia de esta intuición y considera que el deontologismo no logra hacerlo. Para el deontologista, lo único importante es lo que uno sabe al momento de tener que tomar la decisión o hacer el juicio. Si más adelante uno aprende más sobre la situación, éste no debe cambiar la corrección o incorrección de nuestros actos pasados. En términos populares, para el deontologista “el hubiera no existe”. Si nos damos cuenta después de que las cosas hubieran sido mejor si hubiéramos actuado de manera distinta, esto no significa que lo que de hecho hicimos (o no) era incorrecto.

Consideremos los siguientes dos casos, aparentemente muy distintos:

- A. La madre de Sally entra en el dormitorio de su hija y la encuentra buscando algo debajo de la cama. "¿Qué pasa?", le pregunta, "¿qué estás buscando?" (1) "Mis gafas – responde Sally – puede que estén por aquí". Después de echar un largo vistazo debajo de la cama, Sally no encuentra ahí sus gafas; así que se dirige a buscar en otros lugares, no sin antes de decir, (2) “Uy, estaba equivocada.”
- B. “El [segundo] ejemplo es el de un camionero; un camionero que va conduciendo su camión y que no se percata de estar atropellando a un niño que jugaba en la calle. Al conocer el suceso, el conductor sentirá una clase de lamentación por la muerte del niño que nadie más sentirá. Fue él mismo, al fin y al cabo, quien causó la muerte del niño. Ante este hecho, consideraríamos una reacción negativa, incluso inmoral, que el camionero no reaccionase con una profunda tristeza y con un fuerte pesar por la acción que ha cometido. Una reacción sana –según la terminología de Williams– consistiría en la sospecha por parte del agente de que había alguna cosa más que él podría haber hecho y que podría haber salvado la vida del niño: si hubiese estado un poco más alerta, o conducido más despacio; si hubiese visto al niño jugando cerca del camino, etc. ...Esperamos, afirma, que la lamentación del agente se dé hasta en casos en los que no pensamos que el agente haya cometido ninguna falta. Si pensamos que el conductor no podía haber hecho nada más para impedir la muerte del niño, intentaremos consolarlo; pero esto no reemplaza que tuviésemos una visión negativa del conductor si él no mostrase ningún tipo de lamentación, y se limitara a afirmar que ‘es un hecho terrible que haya sucedido, pero yo hice todo el que pude por evitarlo’.” (Rosell 2006, 151-2)

Pese a que ambos casos parecen ser radicalmente distintos, desde (2012) he sostenido que, sin embargo, ambos son sólo casos distintos del mismo (tipo de) problema, el problema de conciliar intuiciones consecuencialistas y deontológicas.

Empecemos analizando del primer ejemplo, el de Sally y por qué ha llamado la atención de filósofos del lenguaje como John MacFarlane. La clave es la extraña reacción de Sally tras no encontrar sus anteojos detrás de la cama. Su “Uy, estaba equivocada” nos suena muy normal, pero es difícil identificar exactamente en qué se había equivocado Sally. Sus palabras suenan a una disculpa, pero no es claro de qué error se está disculpando. Si vemos lo que hizo y dijo – juzgar que sus anteojos podrían estar bajo la cama y buscarlos ahí – suena como un comportamiento correcto y racional, nada por lo que disculparse después.

Según MacFarlane, si lo que Sally afirma en (1) fue también de lo que se retractó en (2), debe haber sido algo que Sally ya sabía antes de mirar bajo la cama (algo como lo que (1) parece significar literalmente, es decir, que (3) sus anteojos podrían haber estado debajo de la cama), pero también algo de lo que ella podría retractarse correctamente después de buscarlas allí (presumiblemente, algo que después se dio cuenta era falso, como que (4) sus gafas, de hecho, estaban debajo de la cama o que, (5) tras buscar bajo la cama, llegaría a saber que sus gafas estaban debajo de la cama. Desafortunadamente, no existe ninguna proposición, en el sentido tradicional, que sea verdadera en el contexto de (1) (como (3)) pero falsa en el contexto de (2) (como (4) y (5)). Por lo tanto, concluyen los relativistas, necesitamos un nuevo tipo de proposición que sea sensible al contexto, y en particular, sensible lo que sabemos sobre la situación en el contexto desde el que la evaluamos (el cual, por supuesto, puede cambiar y no ser el mismo que el del hablante al momento de hacer la afirmación). Así, lo que Sally habría afirmado en (1) sería una proposición sensible a lo que Sally sabe o no sobre la ubicación de sus anteojos. Pero lo que Sally sabe sobre la ubicación de sus anteojos cambia cuando mira debajo de la cama – pasa de no saber si están ahí sus anteojos a saber que no están ahí. Por lo tanto, el valor verdad de lo que dijo Sally – de la proposición que expresó al decir “Mis gafas puede que estén por aquí” cambiará de ser algo verdadero cuando Sally lo dijo, a falso cuando se dio cuenta de que las gafas no estaban ahí. Así, ~~dado que lo que Sally sabía antes de mirar bajo la cama era consistente con que sus anteojos estuvieran ahí, esta proposición era verdadera al momento de que Sally lo afirmó y Sally lo sabía, pero se volvió falsa una vez que cambió lo que Sally sabía sobre sus anteojos después~~

~~de buscarlos debajo de la cama.~~ Desde esta nueva perspectiva, ya no era factible que sus anteojos estuvieran ahí, por lo que la misma proposición, desde este nuevo contexto, se volvió falsa.

Tradicionalmente, no pensamos que las proposiciones pueden cambiar de valor de verdad. una vez que algo es cierto, no puedo volverse falso. Si yo hoy digo que es domingo y mañana también digo que es domingo, hoy habré dicho algo verdadero y mañana habré dicho algo falso. Sin embargo, no hay ninguna proposición que haya dicho ayer que haya sido verdadera, peor al decirlo hoy se haya convertido en falsa. No sería correcto describir lo que sucedió diciendo que ayer dije algo que era verdadero ayer, pero falso hoy. En su lugar, lo correcto es decir que ayer dije una proposición verdad y que, usando las mismas palabras, hoy dije una proposición distinta y, esta vez, falsa. Pero esto no es lo que quiere proponer MacFarlane. Lo que el propone es que lo que dijo Sally era primero verdadero y luego, la misma proposición, sin cambiar su sentido, se convirtió en falsa cuando Sally se dio cuenta de que sus anteojos no estaban bajo la cama.

Este tipo de propuestas, tradicionalmente conocidas como *Relativismos*, han sido consideradas absurdas en filosofía por lo menos desde los tiempos de Protágoras, y tratar de defender una propuesta de este tipo va a enfrentar siempre por lo menos dos tipos de problemas: problemas fundacionales y descriptivos. Los problemas fundacionales son determinar su una noción de proposición (y de verdad) de este tipo no es contradictoria, si puede integrarse al resto de nuestra semántica tradicional, y si es manejable de manera estable, etc. El debate descriptivo gira alrededor de la pregunta de qué tan bien el relativista da cuenta del tipo de fenómenos semánticos que putativamente la motivan. Por ejemplo, ¿qué tan bien explica el relativista lo que sucede en el caso de Sally? ¿Hay explicaciones alternativas? Y si las hay, ¿cómo se comparan sus pros y contras con las del relativismo?

El problema fundacional básico que enfrenta todo relativismo es que parece disolver la normatividad. Si lo correcto o incorrecto de un acto depende tanto de la perspectiva de la que se evalúa, vas la pena preguntarse ¿de qué sirven ya nuestros juicios de qué es correcto o incorrecto? Tal parecería que si alguien ha hecho algo incorrecto, le bastaría solamente cambiar de perspectiva para que se vuelva algo correcto. Si mi perspectiva es diferente de la tuya, ¿cómo podríamos debatir de manera constructiva sobre lo qué es correcto o incorrecto?

Gran parte del trabajo de MacFarlane ha tenido como objetivo explicar porqué estas críticas son injustas y se basan en una imagen equivocada de lo que es la normatividad. Para los críticos del relativismo,

un relativismo como el de MacFarlane es inaceptable porque, si lo que afirmamos no es ni verdadero ni falso al momento en que lo afirmamos (como parece ser cada vez que decimos cosas como “Puede que mis gafas estén por aquí”, o en general, cosas sobre el futuro), no podemos comprometernos con su verdad y si no podemos comprometernos con su verdad, tampoco podemos afirmarlas, lo que desemboca en paradoja. Sin embargo, responde MacFarlane, es un prejuicio absurdo el pensar que no nos podemos comprometer con algo indeterminado. Por supuesto que podemos hacerlo y lo hacemos todo el tiempo. Cada vez que prometemos hacer algo a futuro, por ejemplo, nos comprometemos con algo aun indeterminado. A decir verdad, el hecho de que el futuro aun sea indeterminado es lo que da sentido a este tipo de promesas: nos comprometemos precisamente porque aún podemos hacer algo. En este sentido, la moraleja de MacFarlane es que toda afirmación tiene algo de promesa, es decir, su normatividad no se extingue al momento de que se hace la afirmación. En muchas ocasiones, afirmamos algo, es decir, nos comprometemos con la verdad de algo precisamente porque aún podemos hacer algo al respecto.

Según (Barceló 2012), la reticencia que tenemos a aceptar *apriori* un relativismo como el de MacFarlane se debe a que nuestra concepción de la normatividad está guiada aún por principios deontológicos inaceptables y, en particular, por lo que en otros contextos se conoce como el “principio del control”, el principio según el cual no podemos ser responsables de aquello que no está bajo nuestro control, es decir, que la suerte no debe de tener consecuencias normativas. Esto significa que todos los factores relevantes para determinar si un acto es correcto o incorrecto, bueno o malo, deben estar determinados y disponible al agente al momento de actuar o tomar la decisión de actuar. En otras palabras, una vez que hemos actuado, *alea jacta est* y no hay vuelta atrás.

Contra este principio altamente intuitivo, MacFarlane no presenta otro principio opuesta y también intuitivo, sino que va a los casos particulares y muestra que, aunque el principio nos parece intuitivo, de hecho no actuamos en consonancia con él, sino que en nuestras prácticas normativas sí evaluamos negativamente actos en los cuales lo negativo no se debe a nada que estuviera bajo el control o acceso del agente al momento de actuar, sino a consecuencias que reconocemos como imprevisibles. El ejemplo de Sally es uno, pero el de Williams es otro. Efectivamente consideraríamos una reacción negativa, incluso inmoral, que el camionero no reaccionase con una profunda tristeza y con un fuerte pesar por la acción que ha cometido aun en casos en los que no pensamos que el camionera haya cometida ninguna falta, sino que ha

conducido con suficiente cuidado y responsabilidad. En otras palabras, pese a nuestras intuiciones deontológicas, no podemos imaginar ninguna circunstancia en la cual no valoráramos negativamente a quien, tras causar la muerte de una o más personas, no se sintiera mal, disculpara o aceptará indemnizar a los afectados, sino que respondiera con un “ni modo, estas cosas pasan” o un “qué pena”. Aun más, es también fuerte nuestra intuición de que la manera en que debería de actuar el causante de dichas muertes es sustancialmente diferente de la manera en que debería comportarse cualquier persona, lo mínimamente compasiva y empática, ante una desgracia de esta envergadura. En otras palabras no basta un acto de *condolencia* sino que se necesita un acto de *arrepentimiento*, un *mea culpa*.

Sin embargo, según el deontologismo, de hecho es cierto que “estas cosas pasan” y que hay circunstancias en las cuales el agente no tiene responsabilidad alguna sobre ciertas consecuencias imprevisibles de sus actos, no importa qué tan nefastas y negativas. Recordemos que, según el deontologismo, estas consecuencias no podrían realmente atribuirse al agente, sino que sólo deberían considerarse efectos inconvenientes de actos irreprochables. Sin embargo, tenemos intuiciones fuertes de que nuestros actos pueden llegar a tener consecuencias tan negativas que no pueden considerarse meros efectos inconvenientes. Los casos que menciona Williams y Macfarlane son precisamente de este tipo.

En (2001) Andrew Latus ya había reconocido que el problema de la suerte moral puede presentarse como la paradójica conjunción de un principio general intuitivo y ciertos hechos morales particulares que lo contradicen:

- 1) La intuición de que la suerte no debe hacer diferencias morales (por ejemplo, que la suerte no debe afectar el valor moral de una persona ni aquello de lo que una persona es moralmente responsable).
- 2) el hecho (documentado empíricamente por Berg-Cross 1975, Cushman et.al. 2009, Cushman 2008, etc.) de que la suerte sí parece hacer diferencias morales (por ejemplo, culpamos al conductor desafortunado que causa un daño más que al afortunado conductor que pese a ser diligente llega a su destino sin mayores consecuencias).

La paradoja surge precisamente en los casos de suerte, ya sea buena o mala. En este tipo de casos, la intuición deontológica nos dice que la suerte no debe afectar nuestros juicios normativos. En los casos en los que el agente actúa con cuidado, por ejemplo, sigue siendo irreprochable su conducta aun cuando, por mala suerte,

termine causando algún daño cualquier otro tipo de resultado negativo de sus actos. Sin embargo, es un hecho ampliamente documentado que cuando nos enfrentamos con este tipo de casos, nuestras intuiciones se vuelven consecuencialistas y consideramos que merece algún tipo de castigo – comúnmente una indemnización monetario a los afectados – el que causó daño, y que el cuidado con el que actuó tan sólo sirve, a lo más, para mitigar dicho castigo. En otras palabras, lo que tenemos aquí son actos para los cuales tenemos intuiciones fuertes en sentidos opuesto e incompatibles. Para el deontologista, el agente actuó bien; para el consecuencialista, actuó mal. El trabajo del filósofo es tratar de disolver la paradoja.

Como en todos los casos en los cuales nos enfrentamos a intuiciones en conflicto – una que dice que un objeto A es P (en este caso que un acto es bueno o mal) y otra que dice que A no es P – al filósofo se le abren varias opciones. Las dos primeras son las más obvias: defender una de las intuiciones y criticar la otra. De hecho, en la literatura sobre normatividad, este tipo de respuestas suelen ser las más comunes. Para los consecuencialistas, la regla deontológica general de que la suerte no debe tener relevancia normativa es falsa y debemos respetar nuestras prácticas normativas usuales; mientras que para los deontologistas, dichas prácticas son erróneas, pues lo correcto es respetar la regla general de que sólo somos responsables de aquello que está bajo nuestro control y disposición.

Nótese que la misma paradoja que Latus identificó en el caso de la suerte moral se da también en casos como el de Sally, es decir, también ahí tenemos dos intuiciones en conflicto:

- . 1) La intuición de que la suerte no debe hacer diferencias semánticas, por ejemplo, que la suerte no debe afectar de que es responsable una persona cuando hace una afirmación
- . 2) El hecho de que la suerte sí parece hacer diferencias semánticas (por ejemplo, pedimos disculpas por haber dicho algo falso, pese a que no teníamos manera de saberlo).

Al igual que en el caso de la suerte moral, en mi artículo de 2012, afirmé que las respuestas al problema que llamé “suerte semántica” suelen ser también de dos tipos. Algunos toman la perspectiva consecuencialista y afirman que la regla general está equivocada, que no hay nada malo en que la suerte afecte las condiciones de verdad de lo que decimos; mientras que otros defienden la posición deontológica de que la suerte nunca marca una diferencia en las condiciones de verdad de una aserción. Según William, por ejemplo, (i) el conductor del camión es (i) inocente, (ii) desafortunado, y (iii) justificadamente se siente mal después del

hecho porque (iv) hizo algo horrible. Es (i) inocente porque el accidente no se debió a ninguna negligencia de su parte, sin embargo fué lo suficiente desafortunado (ii) para atropellar al niño. Además, (iii) está justificado en sentirse mal porque (iv) realmente hizo algo horrible: mató a un niño. Una posición análoga a la de Williams para el ejemplo de MacFarlane sería decir que Sally es (i) inocente (al afirmar que las gafas pueden estar debajo del lecho no dijo nada malo), (ii) desafortunada (ya que las gafas, de hecho, no estaban allí), y (iii) justificada al retractarse *post facto* porque dijo algo que era, de hecho, (iv) falso.

Sin embargo, en 2015 me dí cuenta de que el menú de opciones era en realidad mas amplio del que pensaba en (2012) y publiqué en mi blog una nueva clasificación de posibles posiciones que se pueden tomar, en general, en filosofía cuando nos enfrentamos a dos intuiciones en tensión. En otras palabras, tenemos la intuición fuerte de que A es B y de que A no es B. Como ya he mencionado, existen las opciones obvias de defender una intuición y criticar la otra. Esto es lo que suelen hacer los deontologistas y consecuencialistas ya mencionados. Sin embargo, también existen por lo menos tres otras opciones para tratar de incorporar las *dos* intuiciones en una simple teoría sin caer en contradicción (además, por supuesto, también sigue existiendo la opción dialetheista de aceptar la contradicción, no como un problema que reosolver, sino como una característica del fenómeno, como lo hace, por ejemplo, Graham Priest 1985):

1. Dualismo de hechos/objetos
2. Dualismo de valores
3. Relativismo

# VI. Consecuencia

En todos los dominios de dislibro y prácticas normativas la noción de consecuencia juega un papel fundamental – aun más si uno es consecuencialista, pero también para un deontologista. Sin embargo, capturar exactamente qué significa que algo sea consecuencia de otra cosa y, en particular, de una acción humana ha sido un problema gigantesco en la historia de la filosofía. En las siguientes páginas, no pretendo haber encontrado la tan buscada definición, pero sí busco explorar por lo menos algunas características que, sin duda, son fundamentales para entender la elusiva noción de consecuencia y que, por lo tanto, contribuyen a dibujar una imagen más completa de la estructura de la normatividad:

1. No hay una sola noción de consecuencia que opere a lo largo de todos los dominios normativos y se aplique a todos los valores. Por lo tanto, la noción de consecuencia no es la misma que la de causalidad.
2. Para que X sea consecuencia de Y, es necesario (pero no suficiente) que no haya un Z tal que Z sea (i) consecuencia de X y (ii) derrote la justificación de Y.

En este capítulo desarrollaré cada una de estas características en turno. Respecto a la primera, usaré como ejemplo el caso epistémico en el cual la noción de consecuencia relevante no será el de causalidad, sino el de inferencia competente, aunque espero que lo ahí dicho se extienda fácilmente a otros sistemas normativos.

Respecto al segundo, introduciré primero la noción de derrotador y luego la usaré para diagnosticar algunas paradojas famosas, como la del prólogo y la del sorites, mostrando que surgen de considerar como consecuencias relevantes cosas que no lo son realmente.

## A. Derrotadores

Otra noción que podemos adoptar de la epistemología y adaptar a casos más generales es la de derrotador. La idea básica es sencilla, en una sección anterior hacíamos la diferencia entre aquellas situaciones en la que una medida es exitosa y aquella en la que no. En un texto anterior (2017) había adoptado el término de presuposición para hablar de las condiciones necesarias para que una medida sea exitosa. Desafortunadamente, el término “presuposición” en lingüística y filosofía del lenguaje tiene ciertas connotaciones que ahora me parecen desafortunadas para hablar del fenómeno que me interesa. En

particular, la noción de “presuposición” tiene la connotación de ser convencional y en ciertos debates se suele contrastar con cálculos de costo beneficio (por ejemplo, en Barceló y Stainton 2018). Afortunadamente, la noción que trataba de capturar con mi uso de “presuposición” puede capturarse de manera perfecta con otra noción dentro del cajón de herramientas de la filosofía actual de la cual es su opuesto perfecto: la de “derrotador”. Un derrotador es una condición suficiente para que una medida falle. De ahí que lo que antes solía llamar una “presuposición” sea lo opuesto a lo que tradicionalmente, y de ahora en adelante, llamaré un “derrotador”: una medida presupone que se satisfecerán ciertas condiciones bajo las cuales dicha medida puede garantizar su éxito; si dichas condiciones no se satisfacen, es decir, si aparece un derrotador, desaparece dicha garantía – la relación entre la medida y su éxito garantizado es derrotada – y en su lugar, nos queda la aleatoriedad y la suerte: si tenemos buena suerte, tendremos éxito, si no, fallaremos. Así es como se relacionan las nociones de suerte, presuposición, racionalidad, aleatoriedad, éxito, falla y derrota (de tal manera que cualquiera de ellas puede definirse a partir de las otras).

¿Cómo se aplica esta noción a los ejemplos que hemos ya mencionado? Muy sencillo, si estamos preparados para una inundación que nos tenga aislados por una semana, el que la inundación no nos tenga aislados por más de una semana es la presuposición bajo la cual opera nuestra preparación, y si la inundación dura más, esto derrotaría la relación entre nuestra supervivencia y nuestra preparación. Una vez derrotada esa relación, nuestra supervivencia ya no depende de nuestra preparación. Igualmente, Mario Gómez-Torrente ha sostenido que el uso de términos vagos presupone que el contexto de uso es uno en el que no hay series soríticas, lo que significa que las series soríticas son derrotadores de la referencia de los términos vagos. En otras palabras, poseer un concepto vago no debe entenderse en el sentido infalibilista de poseer la capacidad de, en cualquier situación, ser capaz de distinguir entre aquellos objetos que caen sobre él y aquellos que no lo hacen. En su lugar, debe pensarse como la capacidad falible de, en situación previsible, ser capaces de hacer dicha distinción. Por ejemplo, poseer el concepto de “montón de arena” no requiere tener la capacidad de, en cualquier situación ser capaz de distinguir un montón de arena de algo que no lo es, sino de ser capaz de hacerlo en situaciones normales, es decir, racionalmente previsible. En este sentido, el que no seamos capaces de hacer tal distinción en, por ejemplo, una serie de sorteos no cuenta como evidencia de que no poseemos el concepto – o peor aún, de que el concepto sea deficiente en algún sentido – sino que

sólo demuestra que somos falibles y que, por lo tanto, las herramientas que usamos para dar cuenta de la realidad también lo son.

Pensemos en los conceptos, entonces, como herramientas cuyo objetivo es la definición de una extensión a la manera en que hablamos anteriormente, por ejemplo, de las medidas en contra de los ataques a civiles a través de camiones pesados o de medidas para prepararse en caso de riesgo de inundación. Así como es racional tomar medidas como prohibir la circulación de automóviles y colocar bloques de hormigón alrededor de las calles donde se esperan se junten multitudes, aunque sabemos que las medidas podrían evitar ciertos ataques, pero no todos, así también es racional usar conceptos vagos para hacer diferencias en una dimensión a sabiendas que podemos usarlos para definir una extensión en ciertos contextos, pero no en todos.

### C. Como pensar sobre otra cosa

Permítaseme empezar con una paradoja que, hasta lo que yo sé, no ha sido estudiada en la filosofía y que llamaré la paradoja de las tres monedas:

Cuenta la historia que era muy difícil hacerle pagar Proclo sus deudas, pese a ser un sofista muy rico. En una ocasión, Glaucón le había prestado un dracma y no había logrado cobrar. Una tarde, Glaucón se enteró de que Proclo acababa de cobrar tres dracmas por uno de sus trabajos, así que lo abordó en la plaza para cobrar su dinero. Pero, una vez más, Proclo no le haría fácil su trabajo a Glaucón.

- Mira, Glaucón - le respondió sacando de su bolsillo tres monedas - éstos son los tres dracmas que tengo. Me dices que te debo uno ¿podrías decirme cual?

- No entiendo tu pregunta.

- Quiero que me digas cual de estas monedas es la que te debo. ¿Cómo es posible que digas que te debo una moneda, pero no puedas decirme cuál?

- Ninguna moneda en particular, Proclo, ... cualquiera.

- Entonces - respondió Proclo señalando dos de las monedas en su mano - puedo darte una de estas dos y - ahora señalando a la tercera - no tengo que darte ésta otra ¿verdad?

- Tienes razón Proclo.

- Entonces no tengo que darte esta moneda.
  - Así es.
  - Y lo mismo sucede con esta otra moneda – continua Proclo señalando a una de las otras dos monedas –. Tampoco tengo que dártela ¿verdad? En tanto que puedo darte una de las otras dos.
  - Creo que tengo que afirmar que tienes razón Proclo – respondió Glaucón.
  - Y esta última moneda – arguyó Proclo, señalando la tercera moneda en su mano – no es diferente a las otras dos, ¿verdad? Tampoco tengo que dártela, porque puedo darte cualquiera de las otras.
  - Así parece.
  - Entonces, no tengo que darte esta moneda, pero – continuó, señalando una de las otras monedas en su mano – tampoco tengo que darte esta, ni – señalando la tercera y última moneda en su mano – esta otra tampoco. Por lo tanto, reconoces que no debo darte ninguna de estas monedas que traigo ¿verdad?
- Glaucón se quedó callado, sin saber que responder, por lo que Proclo continuó hablando
- Bueno, Glaucón, ya que no debo darte ninguna de estas monedas que traigo conmigo, podemos dar por concluido este asunto y yo seguiré por mi camino. Hasta luego.

El objetivo de este texto es ofrecer una manera de resolver ciertas paradojas clásicas, motivando una revisión de nuestra manera de entender la relación entre aceptación e inferencia. Para lograr esto, he introducido una paradoja novedosa pero de muy fácil solución. Lo que pretendo hacer después es extender la solución de esta paradoja a otras paradojas más interesantes, como la de la lotería.

Hay tres maneras de resolver la paradoja: mostrando o bien (i) que la conclusión (S) no es realmente contradictoria, o (ii) que (S) no se sigue de las premisas (P), (Q) y (R), o que (iii) alguna de dichas premisas no es realmente aceptable. En cada caso, además, es necesario explicar o bien (i) porqué la conclusión, pese a no ser contradictoria, nos *parece* serlo, o (ii) porqué, aunque no se sigue, *parece* seguirse de las premisas, o bien (iii) porqué, aunque alguna de las premisas no es aceptable, todas *nos parecen* aceptables. Soluciones de cada uno de estos tipos existen en la literatura. Mi solución es del segundo tipo, trataré de mostrar que, en todos los casos, la conclusión no se sigue de las premisas.

## I. Resolviendo la paradoja de las tres monedas

Empecemos entonces con la paradoja de las tres monedas. Cualquiera que tenga por lo menos el mínimo de conocimiento de alguna lógica intensional habrá reconocido de inmediato el paso falaz en la argumentación de Proclo: pasar de  $\diamond P$ ,  $\diamond Q$  y  $\diamond R$  a  $\diamond(P \wedge Q \wedge R)$ . Como es bien sabido, el operador  $\diamond$  es débil y por lo tanto esta inferencia es inválida: en general, de que esté permitido  $P$  y este permitido  $Q$ , no se sigue que este permitido  $P$  y  $Q$ . Por ejemplo, yo me puedo casar con Ana y me puedo casar con María, pero de ello no se sigue que me puedo casar con Ana y María. Un paso falaz similar se da al principio de la paradoja, cuando Proclo trata de pasar de  $\forall \diamond F$  a  $\diamond \forall F$ , es decir de que no hay ninguna moneda que deba entregarle Glaucón a que no debe entregarle ninguna moneda a Glaucón.

Si bien creo que esta manera de diagnosticar y resolver la paradoja es correcta, prefiero no usar explícitamente las herramientas de la lógica intensional en este texto por dos razones: en primer lugar, porque creo que la solución puede presentarse también en términos más intuitivos y no quiero dirigir este trabajo sólo a aquellos con conocimientos o inclinación por la lógica formal, y segundo – aunque esto puede que no quede claro más que a los que tengan estos antecedentes lógico-formales – porque no quiero dar la impresión de que la solución depende esencialmente en la modalidad involucrada, en este caso, la de permisibilidad.<sup>26</sup> Por el contrario, y como trataré de mostrar en el resto de mi plática, creo que el diagnóstico y solución que daré generalizan a otros casos en los que no hay una modalidad explícita.

Para facilitar la exposición de la solución demosle nombre a las tres monedas en la mano de Proclo. Llamemos a las monedas “Hugo”, “Paco” y “Luis”. En un primer momento, Proclo, señalando dos de las monedas en su mano – supongamos que señala a Paco y a Luis – le dice

---

<sup>26</sup>. Una reciente propuesta para resolver estas paradojas apelando a la modalidad de permisibilidad se encuentra en (Kroedel 2013). Una crítica muy clara a los problemas que este tipo de propuestas involucran se encuentra en (Littlejohn 2013). Mi teoría de la relación entre aceptación, e inferencia pretende llenar el hueco en una propuesta como la de Kroedel.

“puedo darte una de estas dos y – ahora señalando a la tercera, es decir, a Hugo, continua – no tengo que darte ésta otra ¿verdad?” Al igual que Glaucón, no podemos sino asentir a la pregunta de Proclo. Ahora bien, reflexionemos sobre lo que sucede cuando aceptamos que, efectivamente, Proclo no debe entregarle Hugo a Glaucón, es fácil darse cuenta de que aceptamos precisamente porque reconocemos que hay una manera en la cual Proclo puede pagar su deuda y así cumplir su compromiso con Glaucón, sin entregar a Hugo, a saber, entregándole a Paco o a Luis. En otras palabras, aceptamos que Proclo no tiene que entregarle Hugo a Glaucón, pero siempre y cuando le entregue a Paco o a Luis. Es decir, reconocemos que Proclo puede no entregarle una de las monedas, siempre y cuando le de una de las otras. Ahora bien, nótese que aunque no hagamos explícita esta condición en nuestra respuesta afirmativa a la pregunta de Proclo, ésta es esencial para nuestra aceptación de la permisibilidad de que Proclo no le entregue Hugo a Glaucón. Esto es así porque, dadas las circunstancias, si Proclo no le entrega Hugo a Glaucón, lo normal sería que le entregara a Paco o a Luis. A esta condición es lo que de ahora en adelante voy a llamar la **condición de normalidad** detrás de nuestra aceptación: aceptamos que Proclo no tiene que entregarle a Hugo a Glaucón, bajo la condición (de normalidad) de que le entregue a Paco o a Luis.<sup>27</sup>

Nótese que el mismo proceso tiene lugar con cada una de las otras dos monedas. Aceptamos que Proclo no tiene que entregarle a Paco a Glaucón, bajo la condición de normalidad de que le entregue a Hugo o a Luis, y aceptamos que Proclo no tiene que entregarle a Luis a Glaucón, bajo la condición de normalidad de que le entregue a Hugo o a Paco. En otras palabras, cada vez que juzgamos si es obligatorio o no que Proclo haga o no alguna cosa, lo que en realidad evaluamos es todo un libro de acción. Cuando aceptamos que está permitido que Proclo no entregue a Hugo, lo que juzgamos como permitido es un libro de acción complejo que no se reduce a Proclo no entregando a Hugo, sino que incluye también y de manera esencial el que le entregue a Paco. Es

---

<sup>27</sup>. En este ejemplo, la condición de normalidad es fácilmente accesible a la conciencia. Sin embargo, no quiero que quede la impresión de que así debe ser. por le contrario, creo que una de las razones por las cuales las condiciones de normalidad son implícitas es porque muchas de ellas son de difícil acceso a la conciencia.

decir, aceptamos que está permitido que no entregue a Hugo, sólo cómo parte de un libro de acción permitido donde, aunque no entrega a Hugo, sí entrega a Paco o a Luis. En otras palabras, en nuestra aceptación, hay algo explícito – aceptamos que Proclo no tiene que entregarle a Hugo a Glaucón – y algo implícito – el que le entregue a Paco o a Luis.

La idea básica detrás de la noción de condición de normalidad es que, cuando evaluamos un enunciado, no consideramos *todas* las maneras en que la proposición expresada podría ser verdadera, sino sólo las que nos parecen mas naturales. Por ejemplo, si nos preguntamos si debemos golpear al desconocido con el que nos cruzamos en la calle, fácilmente descartamos como poco relevante la posibilidad de que el desconocido en cuestión se dirija a cometer un crimen horrible. En otras palabras, cuando pensamos en un desconocido, racionalmente pensamos que no es un criminal peligroso que debe ser detenido por los medios que sean necesarios. Cuando consideramos el enunciado “En la florería de mi tía venden lirios”, por poner otro ejemplo, no sólo pensamos que tenemos una tía que tiene una florería, sino también que en dicha florería es una florería *normal*, es decir, una en la que no se venden sólo lirios. En sentido estricto, una florería que vende lirios podría vender lirios solamente, pero lo racional es considerar la posibilidad mas natural, lo que los lógicos formales llaman los mundos posibles mas cercanos donde la proposición es verdadera.<sup>28</sup> Así pues, cuando Glaucón considera la posibilidad de que Proclo no le de una moneda, estima como lo más natural el que no le de dicha moneda porque le da otra, no porque no le da ninguna moneda.

---

<sup>28</sup>. Algunos teóricos han propuesto una formalización de esta idea apelando a mundos posibles, una relación de cercanía entre dichos mundos posibles y una función que ordena los mundos posibles por qué tan ideales son de acuerdo a la normatividad relevante. Por ejemplo, si estamos modelando la obligatoriedad que emana de un código legal dado, un mundo W1 sería más ideal que otro W2 si W1 obedece mas provisiones de dicho código que W2. Así, una proposición p es obligatoria si y sólo si los mundos posibles donde p es verdadera son mas cercanos al mundo real son mas ideales que los mundos mas cercanos donde p es falsa. La idea es ir mas allá del modelo tradicional de obligatoriedad, donde una proposición p es obligatoria es verdadera en todo mundo donde se cumplen con la normatividad relevante. Este nuevo modelo encaja con mi propuesta en que, cuando evaluamos si una proposición dada es obligatoria, no consideramos todas las maneras en que la proposición podría ser verdadera, sino sólo las que nos parecen mas naturales, es decir, aquellas donde se cumplen sus condiciones de normalidad.

Proclo fuerza dicha estimación sobre Glaucón al decirle “puedo darte una de estas dos y – ahora señalando a la tercera – no tengo que darte ésta otra ¿verdad?. Es por ello que Glaucón considera, al considerar si Proclo debe entregarle a Hugo, que lo normal sería que le entrega a Paco o a Luis en su lugar.

Ahora bien, si aceptamos que Proclo no tiene que entregarle a Hugo a Glaucón, que no tiene que entregarle a Paco, y que no tiene que entregarle a Luis ¿porqué no hemos de aceptar que no tiene que entregarle a Hugo ni a Paco ni a Luis, es decir, que no tiene que entregarle ninguna de las monedas en su mano? Creo que la respuesta debe ser ya obvia. Si bien podemos reconocer fácilmente un posible libro de acción en el que Proclo puede pagar su deuda sin entregar a Hugo (porque, en su lugar, le entrega a Paco o a Luis), podemos reconocer un posible libro de acción en el que Proclo puede pagar su deuda sin entregar a Paco (porque, en su lugar, le entrega a Hugo o a Luis) y podemos reconocer un posible libro de acción en el que Proclo paga su deuda sin entregar a Luis (porque, en su lugar, le entrega a Hugo o a Paco), no podemos reconocer ningún posible libro de acción en el que Proclo paga su deuda sin entregar una de las monedas, ya sea a Hugo o a Paco o a Luis.

Para entender mejor como funcionan las condiciones de normalidad en la inferencia, analicemos otro caso posible. Supongamos que después de que Glaucón reconoce que Proclo no tiene que entregarle a Hugo, ni tiene que entregarle a Paco, Proclo le hubiera preguntado si de ello se sigue que no tiene que entregarle ni a Hugo ni a Paco. La respuesta de Glaucón seguramente sería positiva. Efectivamente, Proclo puede no entregarle a Hugo ni a Paco a Glaucón y aun así pagar su deuda, siempre y cuando le entregue a Luis. En otras palabras, podemos bien aceptar que Proclo no tiene que entregarle ni a Hugo ni a Paco, pero sólo bajo la condición de normalidad de que le entregue a Luis. En este caso, la inferencia de que Proclo no tiene que entregarle a Hugo, ni tiene que entregarle a Paco, a que no tiene que entregarle ni a Hugo ni a Paco, no es falaz. La razón se encuentra en las condiciones de normalidad: Recordemos que la condición de normalidad de que Proclo no entregue a Hugo es que entregue a Paco o a Luis, y la condición de normalidad de que

Proclo no entregue a Paco es que entregue a Hugo o a Luis. Ahora bien, para que aceptamos que Proclo no entregue ni a Hugo ni a Paco, deben cumplirse las condiciones de normalidad detrás de cada una de las premisas del argumento, es decir, tiene que entregar a Paco o a Luis y a Hugo o Luis. La única manera en que, tanto la conclusión como ambas condiciones de normalidad se cumplan es que Proclo entregue a Luis. En otras palabras, en el contexto de este argumento, la conclusión tiene como condición de normalidad que Proclo le entregue a Luis a Glaucón. En otras palabras, cuando ejecutamos la inferencia de que Proclo no tiene que entregarle a Hugo, ni tiene que entregarle a Paco, a que no tiene que entregarle ni a Hugo ni a Paco, lo que conjuntamos no es sólo la parte implícita de las premisas, sino también sus condiciones de normalidad. Esquemáticamente:

1. No tengo que entregar a Hugo [bajo la condición de normalidad de que entrego a Paco o a Luis]
2. No tengo que entregar a Paco [bajo la condición de normalidad de que entrego a Hugo o a Luis]

Por lo tanto, no tengo que entregar ni a Hugo ni a Paco [bajo la condición de normalidad de que entrego a (Paco o a Luis) y entrego a (Hugo o a Luis) sin entregar ni a Hugo ni a Paco, es decir, bajo la condición de normalidad de que entrego a Luis]<sup>29</sup>

Nótese como la conclusión hereda la parte que es consistente con la conclusión de la conjunción de las condiciones de normalidad de las premisas. En otras palabras, la condición de normalidad de no entregar ni a Hugo ni a Paco es entregar a Luis porque esa es la única manera de que se cumplan la conclusión y las condiciones de normalidad de las premisas. Sin embargo, no podemos hacer lo mismo para obtener la conclusión que quiere Proclo, precisamente porque las condiciones de normalidad de las premisas son inconsistentes con la conclusión, es decir, no hay manera que se cumplan las condiciones de normalidad de todas las premisas y la conclusión. No es posible entregar

---

<sup>29</sup>. Para seguir este tipo de inferencias al lector le puede servir pensar en la relación entre una proposición y su condición de normalidad como una implicación material cuyo antecedente es la conjunción de las condiciones de normalidad y cuyo consecuente es la proposición en cuestión.

a Hugo o a Paco, a Paco o a Luis y a Hugo o a Luis, y al mismo tiempo no entregar ni a Hugo ni a Paco ni a Luis. Por eso, no podemos inferir de que Proclo puede no entregar a Hugo, puede no entregar a Paco y puede no entregar a Luis, a que puede no entregar a ninguno de los tres.<sup>30</sup>

1. No tengo que entregar a Hugo [bajo la condición de normalidad de que entrego a Paco o a Luis]
2. No tengo que entregar a Paco [bajo la condición de normalidad de que entrego a Hugo o a Luis]
3. No tengo que entregar a Luis [bajo la condición de normalidad de que entrego a Hugo o a Paco]

Por lo tanto, no tengo que entregar ni a Hugo ni a Paco ni a Luis [bajo la condición de normalidad de que entrego a Paco o a Luis, a Hugo o a Luis, a Hugo o Paco y ni a Hugo ni a Paco ni a Luis]

En general, uno puede asentir a una serie de enunciados  $P_1, P_2, P_3, \dots P_n$  sin comprometerse con una de sus consecuencias lógicas  $C$ , si (la conjunción de) las condiciones de normalidad de  $P_1, P_2, P_3, \dots P_n$  son inconsistentes con  $C$ .<sup>31</sup>

Nótese como apelar a las condiciones de normalidad es esencial para explicar así la falacia detrás de la paradoja de las tres monedas. Nótese además que el número de monedas en la mano de

---

<sup>30</sup>. Hay otra manera, completamente equivalente de diagnosticar y resolver la paradoja. Podemos decir que ésta comete una falacia de equivocación, pues hay dos lecturas del enunciado “Proclo no tiene que darle ninguna de las monedas en su mano a Glaucón”, una (en la que significa que ninguna de las monedas que Proclo tiene en su mano es la que tiene que darle a Glaucón) que efectivamente se sigue de lo que Glaucón le aceptó a Proclo, pero no tiene la consecuencia indeseable de llevar a Glaucón a inconsistencia y otra (en la que significa que es falso que Proclo tiene que darle a Glaucón alguna de las monedas en su mano) que sí tiene esta consecuencia, pero no se sigue de lo que Glaucón aceptó. Por cuestiones de espacio, no desarrollaré esta manera de resolver la paradoja.

<sup>31</sup>. Nótese que esto es compatible con que un sujeto acepte  $C$  tanto como  $P_1, P_2, P_3, \dots P_n$ . Esto es así porque el sujeto puede asentir a  $C$  bajo otras condiciones de normalidad que no tengan nada que ver con las condiciones de normalidad de  $P_1, P_2, P_3, \dots P_n$ . Sin embargo, en ese caso, no habrá inferido  $C$  de dichas creencias.

Proclo es completamente irrelevante para la paradoja y su solución. Pudimos haberla formulado con dos, cinco, veinticinco o doce mil monedas. Es más, pudimos haberla formulado incluyendo todos los dracmas de la Grecia antigua. (¿Podríamos haberla formulado con un número infinito de monedas? Creo que sí, pero este punto no es relevante ahora, así que no nos detendremos en él) Es más, me parece que el fenómeno va mas allá de asuntos de dinero<sup>32</sup> y deudas o de permisos y obligaciones. En realidad, tiene más que ver con qué significa aceptar un enunciado, como trataré de mostrar en lo que queda del capítulo.

En el caso de la paradoja de las tres monedas, creer que Proclo le debe entregar a Galucón una moneda para cumplir con su deuda no significa que haya tal cosa como la moneda que uno crea es la que se le deba entregar para cumplir con dicha deuda. En general, uno puede creer que hay un  $X$  que es  $F$  sin que haya tal cosa como el  $X$  que uno crea que es  $F$ . Uno puede bien creer que un vaso de agua calmará su sed sin que haya tal cosa como el vaso del cual uno crea que es el que calmará su sed. Uno puede saber que uno de los sospechosos es el asesino sin que haya tal cosa como el sospechoso que uno sabe es el asesino; etc.

Algo análogo sucede en el caso de las aceptaciones. El que uno acepte que Proclo le debe entregar a Galucón una moneda para cumplir con su deuda no significa que haya alguna moneda que uno deba aceptar sea la que le deba entregar para cumplir con dicha deuda. En general, uno puede aceptar que hay un  $X$  que es  $F$  sin que haya tal cosa como el  $X$  tal que uno acepte que dicho  $X$  sea  $F$ . Esto se debe a que las condiciones de normalidad de los enunciados existenciales son muy distintos de las de los enunciados particulares. Cuando uno hace una afirmación existencial, como “Me debes un dracma” o “nacé en un pueblo muy pequeño”, normalmente no existe un objeto en particular del que uno esté hablando o su identidad es considerada poco relevante. En contraste, cuando alguien dice de un objeto o grupo de objetos  $X$  que tienen cierta propiedad  $F$ , normalmente lo hace para contrastarlo con otros objetos que no son  $F$  (o lo son de una manera más débil, obvia o

---

<sup>32</sup>. Después de todo, la paradoja de las tres monedas se podría haber resultado también haciendo una distinción entre dinero – que es lo que Proclo le debe realmente a Galucón – y monedas – que es lo que usamos para representar el dinero.

menos importante). Igualmente, cuando uno considera un enunciado de la forma “ $X$  es  $F$ ” (o “Los  $X$  son  $F$ ”) no piensa en todas las posibilidades que lo harían verdadero, sino en las más naturales, y ellas usualmente excluyen casos en los que es verdadero de manera vacua, por ejemplo, porque no hay ningún  $F$ .

Ilustremos esta tesis con un ejemplo. Supongamos que nuestro amigo Víctor es un activista comprometido con la lucha por los derechos de las minorías sexuales, y que ha dedicado a dicha lucha mucho de su vida. Además, aunque en menor grado, le preocupan los problemas que genera la explotación demográfica y por ello ha llegado a la conclusión de que ninguna familia debería tener más de un hijo, a lo más. Si a Víctor se le preguntara si cree que las parejas gay debieran no tener más de un sólo hijo (en un contexto en el cual su compromiso con los derechos de las minorías sexuales no fueran conocido y asumido), tenemos la fuerte intuición de que su respuesta no podría ser “sí”, pese a que de que ninguna familia debería tener más de un hijo se sigue que las parejas gay – en tanto familias – tampoco deberían tener más de un hijo. En general, si creemos que todos los  $X$  son  $Y$  (y que tal información es relevante para la conversación) no solemos afirmar enunciados más débiles como que un  $X$  en particular es  $Y$  o que un grupo particular de  $X$  es  $Y$ . Es por ello que uno puede creer que todos los  $X$  son  $Y$ , y rechazar que un  $X$  en particular o un grupo particular de  $X$  sean  $Y$  sin caer en inconsistencia.

Nótese además que si, en vez de preguntarle sobre parejas gays, le hubiéramos preguntado sobre cualquier otro tipo *particular* de familia si creía que no debía tener más de un hijo, pasaría exactamente lo mismo. Dado que Víctor no cree que ningún tipo *particular* de familia debería tener a lo más un hijo, sino que *ninguna familia en general* debería tener más de un hijo, lo correcto sería que rechazara la afirmación de que, por ejemplo, las familias monopaterales no deberían tener más de un hijo, o las familias con padres de escasos recursos, etcétera. Esta tensión entre la aceptación de enunciados particulares y una creencia general que parecía ser inconsistente con ellas es lo que está detrás de la paradoja de las tres monedas – y de otras paradojas conocidas, como trataré de mostrar en este texto: Así como Víctor puede tener la creencia general de que las familias no deberían tener

más de un hijo, sin aceptar de ningún tipo particular de familia que no debería tener más de un hijo, en la paradoja de las tres monedas, Glaucón puede creer racionalmente que se le debe entregar una moneda pero no aceptar de ninguna moneda que es la que se le debe entregar (pues no hay tal cosa como *la* moneda que crea que sea la que Proclo le debe entregar).

Reconozco que pocos tomarían la paradoja de las tres monedas como una paradoja genuina o especialmente difícil de resolver. Sin embargo, como trataré de mostrar en el resto de la plática, me parece que un fenómeno del mismo tipo está detrás de algunas de las paradojas más famosas de la filosofía analítica reciente, en particular, la paradoja del prefacio, la de la lotería y la del *sorites*. Esta simple paradoja me ha servido, mas que nada, para introducir la noción de condición de normalidad en un contexto poco problemático. Es por ello que me he detenido tanto tiempo en ella, antes de pasar a paradojas mas interesantes y genuinas.

## **II. Extendiendo la propuesta: Las paradojas del prefacio y de la lotería.**

Admito que la paradoja de las tres monedas por si misma es poco interesante. Sin embargo, creo que fijarnos qué hay detrás de ella sirve de muy buen propedéutico para tratar luego otras paradojas más interesantes. Empecemos por una muy simple: la famosa paradoja del prefacio, la cual es muy sencilla de formular. Supongamos que un autor –llamémosle “Miguel”– escribe un libro de no-ficción. Miguel es un investigador concienzudo y ha hecho una labor formidable investigando su tema, pero también es sensato y sabe que pese a que lo que ha escrito es el resultado de una investigación seria y profunda, lo más probable sea que haya cometido por lo menos algún error y, en consecuencia, en su prefacio reconoce que por lo menos algo de lo que sostiene en el libro muy probablemente sea falso. No hay nada extraño, ni insensato en esto. Sin embargo, la actitud de Miguel frente a su propio libro parece ser paradójica. Por un lado, tiene la creencia existencial de que por lo menos una de las oraciones que sostiene en su libro es falsa, pero si le preguntáramos, de cualquier oración en su libro si es verdadera, sin duda el autor afirmaría su verdad sin chistar. En otras palabras, sean  $P_2, P_3, \dots, P_n$  las oraciones contenidas en su libro, el autor aceptaría que " $P_1$  es

verdadera", " $P_2$  es verdadera"... o " $P_n$  verdadera". Esto parece significar que piensa que todas y cada una de las afirmaciones de su libro son verdaderas, lo cual es inconsistente con su propia afirmación en el prefacio de su obra de que por lo menos alguna de dichas afirmaciones es falsa.

Esquemáticamente, el argumento tiene la siguiente forma:

1. La primera afirmación del libro es verdadera. [Premisa]
2. La segunda afirmación del libro es verdadera. [Premisa]
3. La tercera afirmación del libro es verdadera. [Premisa]
- ...
- N. La  $n$  afirmación del libro es verdadera. [Premisa]
- N+1. Hay  $n$  afirmaciones en el libro. [Premisa]
- N+2. Por lo menos una de las afirmaciones del libro es falsa. [Premisa]
- N+3. De 1 a N+1 se sigue que todas las afirmaciones del libro son verdaderas.
- N+4. Pero N+2 y N+3 se contradicen entre sí.

La paradoja del prefacio nos enfrenta con un dilema vergonzoso ambos de cuyos cuernos son inaceptables. O bien el autor de ningún libro puede realmente reconocer que debe haber por lo menos un error en su libro (y decirlo sólo es falsa molestia) o bien todo autor introduce en sus libros algo que no cree que es verdadero. Sin embargo, nos parece muy claro que es posible que uno puede racionalmente llenar un libro de afirmaciones que considera verdaderas y, sin embargo, reconocer que por lo menos una de ellas puede ser falsa, sin caer en una contradicción. ¿Qué sucede en esta situación entonces?

Mi propuesta es ver la paradoja del prefacio como completamente análoga a la paradoja de las tres monedas. En ambas, me parece, hay una aparente tensión entre un juicio existencial positivo y una serie de juicios particulares negativos que parecen excluir la instanciación del existencial. En el caso de la paradoja de las tres monedas, la tensión se daba entre el juicio existencial "Proclo le debe una moneda a Galucón" y la serie de juicios particulares de la forma "Proclo no le debe dar

esta moneda a Galucón”, “Proclo no le debe dar esta otra moneda a Galucón”, etc., que parecen excluir la posibilidad de que el existencial se instancie (ya que si se instanciara, debería de ser en esta moneda, o en esta otra, etc.) Como burlescamente preguntaba Proclo de manera retórica, ¿cómo es posible que digas que te debo una moneda, pero no puedas decirme cuál?

La paradoja del prefacio tiene exactamente la misma estructura. Tenemos un juicio existencial – hay por lo menos una afirmación falsa en este libro – en aparente tensión con una serie de juicios particulares – esta afirmación es verdadera, esta también, y esta también... – que parecen excluir la instanciación del existencial. Es como si Julián le preguntará a Miguel: ¿Cómo es posible que digas que hay por lo menos una falsedad en el libro, pero no puedas decirme cuál? Y así como tiene una estructura similar, su diagnóstico y solución también es completamente análogo. Basta reconocer que cada vez que el autor del libro acepta una de sus propias afirmaciones como verdadera, lo hace bajo la condición de normalidad de que no es infalible, es decir, que puede, ha y volverá a cometer errores y afirmar cosas que son falsas. En otras palabras, acepta cada cosa que ha dicho bajo la condición de normalidad de que en otras ocasiones habrá dicho o dirá cosas falsas. En particular, en el caso del libro, el autor acepta de cada afirmación que hace que es verdadera, bajo la condición de normalidad de que por lo menos *alguna otra* de las afirmaciones que hace en él es falsa.

Nótese como este caso es completamente análogo al caso de las tres monedas. Así como Glaucón aceptaba de cada una de las monedas en la mano de Proclo que no tenía que dársela, bajo la condición de normalidad de que le diera *alguna otra* moneda, así también el autor del libro acepta de cada una de las afirmaciones que hace en el libro que es correcta, bajo la condición de normalidad de que por lo menos *alguna otra* no lo esté. Es por ello que, de que el autor acepte de cualquiera de las afirmaciones en el libro que cree que es verdadera no se sigue que el autor crea (o esté comprometido a aceptar) que todas ellas son verdaderas. Esto se debe a que, si juntamos las condiciones de normalidad de cada una de las premisas en el argumento, veremos que son inconsistentes con la conclusión de éste. Imaginemos que el autor hace 300 afirmaciones en su libro.

Es fácil ver que el autor acepta que su primera afirmación es verdadera sólo bajo el supuesto de que por lo menos una de las otras 299 afirmaciones no lo sea. Análogamente, acepta que su segunda afirmación es verdadera, sólo bajo el supuesto de que alguna de las otras no lo sea. Así sucesivamente con la tercera, la cuarta, la enésima afirmación. Si acepta una es porque piensa que debe haber *otra* que sea falsa. En otras palabras, detrás de la aceptación de que la afirmación  $n$  es verdadera, está la condición de normalidad de que hay una afirmación distinta a  $n$  en el libro que no lo es. Sin embargo, no hay ninguna situación aceptable para el autor en la cual se cumplen el conjunto de las condiciones de normalidad de las afirmaciones particulares – es decir, que alguna de las otras afirmaciones es falsa – y la conclusión del argumento – que todas las afirmaciones en el libro son verdaderas. Por ello es que, aunque la conclusión se sigue deductivamente de las premisas, quien acepte las premisas no está racionalmente comprometido a aceptar la conclusión.

Aplicando lo visto en la sección anterior sobre obligatoriedad al caso de la aceptación o creencia, mi propuesta es que cuando evaluamos una proposición para aceptarla o rechazarla, lo que hacemos es evaluar una situación más compleja, que consideramos es la manera más natural en la que se cumple la proposición.<sup>33</sup> Continuando con uno de los ejemplos ya mencionados. Si se nos pide evaluar el enunciado “La florería de mi tía vende lirios”, lo que evaluamos no es solamente la verdad de la proposición “La florería de mi tía vende lirios”, sino mas bien una situación mas compleja donde tenemos una tía que tiene una florería con lirios y otras flores, los precios que da son normales, no es el frente de una operación internacional de lavado de dinero o contrabando de divisas escondidas en macetas, etc. Cuando asentimos al enunciado de que la florería de nuestra tía vende lirios, lo que está detrás de nuestra afirmación es la consideración de que el mundo es muy probablemente tal y como lo describe dicha situación. Si dicha situación nos parece real o por lo menos muy probable, aceptamos el enunciado. En otras palabras, la aceptación de un enunciado es

---

<sup>33</sup>. A decir verdad, no es suficiente que la proposición en cuestión sea verdadera en la situación, sino que se requiere algo más fuerte, algo como que la proposición describa bien la situación. Evito esta complejidad a lo largo de mi texto porque la diferencia no es muy importante para los casos que aquí discuto. Una idea muy similar es desarrollada por Mitchell S. Green en (1999).

derivada, no de la aceptación de una proposición ligada semánticamente al enunciado, sino de la aceptación de una situación más compleja; dicha situación es más compleja que la proposición en tanto contiene mayores compromisos. Estos compromisos extras que la situación añade a la proposición es lo que he llamado condiciones de normalidad. Es por ello que para poder combinar dos proposiciones en una inferencia, necesitamos considerarlas en una sola situación (que comúnmente es la suma de las dos situaciones en las que las aceptábamos por separado), y al reunir las en dicha situación, sumamos sus condiciones de normalidad.

En consecuencia, el que exista una situación aceptable en la que P es verdadera y otra en la que Q es verdadera, no garantiza que exista una situación aceptable en la que ambas P y Q sean verdaderas. Sólo si ambas son verdaderas en una misma situación (es decir, si sus condiciones de normalidad son compatibles) podemos combinarlas en una inferencia. Es por ello que la regla de conjunción no se cumple siempre, sino solamente cuando se satisfacen ciertas condiciones (entre ellas, el que las condiciones de normalidad de los conjuntos sean consistentes entre sí y con la conjunción). Es por ello que, en el argumento de la paradoja, que el autor acepte las premisas de 1 a N, no lo compromete a que acepte también la conjunción total N+3 (todas las afirmaciones del libro son verdaderas). Así se evita comprometerse a la contradicción de aceptar tanto N+3 como N+2 (por lo menos una de las afirmaciones del libro es falsa), y se resuelve la aparente paradoja.<sup>34</sup>

Una paradoja de estructura muy similar es la paradoja de la lotería, según la cual hay una tensión entre creer que un boleto de la lotería ganará, y rechazar cualquier afirmación de que algún boleto particular será el ganador. En palabras de Mauricio Zululaga (2005):

Supongamos que hemos comprado un billete de lotería, alguno de los cuales deberá ganar. Supongamos además que la hay 100 billetes de lotería. La probabilidad de que uno de los billetes gane es muy baja: 0,01. Supongamos que usted ha comprado un

---

<sup>34</sup>. Es importante notar también que dado que las condiciones de normalidad nos dan condiciones necesarias mas no suficientes para la inferencia, de mi propuesta no se sigue que sería correcto para el autor llegar al enunciado 299 y decir “Dado que sé que por lo menos un enunciado de los que contiene mi libro es falso, y he pasado por todos menos este último y de todos estoy seguro es correcto, debo concluir por razón que este último debe ser falso.”

billete de lotería. De acuerdo con la baja probabilidad de que su billete gane, usted está justificado para afirmar que creer que perderá —suponiendo que uno está justificado para creer que algo será el caso, si la probabilidad de que ocurra es mayor a 0,5—. Pero bajo estas condiciones usted también está justificado para creer que cada uno de los billetes perderá. Así es racional creer que ninguno de los billetes ganará, porque para cada uno de ellos usted ha podido establecer que la probabilidad de que gane es menor a 0,5. Pero, *ex hypothesi*, ha de haber un billete de lotería que gane, así que usted cree algo contradictorio. Usted está justificado para creer que ningún billete ganará y, al mismo tiempo, cree que uno lo hará.

Una vez más, cada vez que aceptamos de un boleto que no será el ganador lo hacemos bajo la condición de normalidad de que *otro* boleto ganará, es decir, lo que aceptamos es una situación en la que el boleto en cuestión no gana, pero otro sí. Por ello, la conjunción de todas las condiciones de normalidad de todas las afirmaciones sobre cada boleto son incompatibles con el que ningún boleto gane (a decir verdad, cada una de las proposiciones de cada una de las creencias particulares es ya incompatible con el que ningún boleto gane).

Las tres paradojas que hemos mencionado hasta ahora tienen la misma estructura: pretenden que haya una inconsistencia entre aceptar una proposición existencial (1) y una serie exhaustiva de aceptaciones singulares sobre el dominio del existencial de dicha proposición (2):

1. *S* acepta que hay (por lo menos) un *X* que es *F*.
2. Para todo *x* en *X*, *S* acepta que *x* no es el/uno de los *F*.

En el caso de las tres monedas, éstas son:<sup>35</sup>

1. Glaucón acepta que Proclo le debe entregar una de las monedas en su mano.
2. Para toda moneda *x* en la mano de Proclo, Glaucón acepta que esa moneda *x* no es la que Glaucón le debe entregar.

En el caso del prefacio, éstas son:<sup>36</sup>

1. El autor acepta que por lo menos una de las afirmaciones en su libro es falsa.

---

<sup>35</sup>. Las *X* son las monedas en la mano de Proclo, y *F* es la propiedad de deber ser entregada a Glaucón.

<sup>36</sup>. Las *X* son las afirmaciones en el libro, y *F* es la propiedad de ser falsa.

2. Para toda afirmación  $x$  en su libro, el autor acepta que esa afirmación  $x$  no es una de las las que son falsas.

Finalmente, en el caso de la lotería, éstas son:<sup>37</sup>

1. Aceptamos que por lo menos uno de los boletos de la lotería ganará.
2. Para todo boleto  $x$  de la lotería, aceptamos que ese boleto  $x$  no es el que ganará.

Gracias a que tienen la misma estructura, las tres paradojas se resuelven de la misma manera, mostrando que de (1) y (2) no se sigue que el agente está siendo irracional al aceptar o comprometerse a aceptar una contradicción. Para ello, basta darse cuenta de que lo que el agente acepta en (1) no es inconsistente con lo que acepta en (2). En el caso de la paradoja de las tres monedas, habíamos mostrado que aunque Glaucón acepta de todas las monedas en la mano de Proclo que éste no debe entregársela, esto no lo compromete a que deba aceptar que Proclo no debe entregarle ninguna moneda. En otras palabras, de (2) no se sigue que

- 3, Glaucón está comprometido a aceptar que Proclo no le debe entregar ninguna de las monedas en su mano.

De (1) y (3) sí se sigue que Glaucón está siendo irracional en su aceptación de una contradicción explícita (pues en (1) acepta que Proclo le debe entregar una de las monedas en su mano y en (3) debe aceptar que no le debe entregar ninguna de las monedas en su mano). Afortunadamente, (3) no es una descripción adecuada de lo que Proclo ha aceptado; (2) lo es, y de (2) no se sigue (3).

De manera similar, en el caso del prefacio, hay una tesis (3) similar:

3. El autor está comprometido a aceptar que ninguna de las afirmaciones en su libro es falsa.

Una vez más, de (1) y (3) sí se sigue que el autor está siendo irracional en su aceptación de una contradicción explícita, y también en este caso (3) no es una descripción adecuada de lo que el autor ha aceptado; (2) lo es, y de (2) no se sigue (3).

Igualmente, en el caso de la lotería, tenemos la tesis

---

<sup>37</sup>. Los  $X$  son los boletos de lotería, y  $F$  es la propiedad de ganar la lotería.

3. Aceptamos que ninguno de los boletos de la lotería ganará.

tal que, aunque de (1) y (3) sí se seguiría que estamos siendo irracionales en nuestra aceptación de una contradicción explícita, (3) no es una descripción adecuada de lo que el autor ha aceptado; (2) sí lo es, pero de (2) no se sigue (3).

En otras palabras, en general, de (2) no se sigue que

3.  $S$  está comprometido a aceptar que ningún  $X$  es  $F$ .

que sería lo necesario para derivar la irracionalidad del agente involucrado en cada paradoja. En otras palabras, en todos los casos la paradoja se resuelve rechazando la inferencia de, para todos los  $X$ ,  $S$  cree que  $X$  es  $F$ , a  $S$  cree que todos los  $X$  son  $F$ .

### III. La Paradoja Sorites

Me parece que apelar así a las condiciones de normalidad tiene muchas ventajas, entre ellas la de poder dar un diagnóstico y solución similar para un gran número de paradojas y acertijos filosóficos, desde la paradoja de Moore a las presuntas fallas de clausura epistémica, etc. Por supuesto, no puedo mostrar esto aquí, aunque vale la pena mencionar que Nancy Nuñez, está explorando esta manera de explicar las aparentes fallas de clausura epistémica. Sin embargo, no quiero dejar de pasar la oportunidad de mostrar como apelar a condiciones de normalidad podría usarse para resolver otra famosa paradoja, aparentemente muy distinta a las antes presentadas: la paradoja de sorites.<sup>38</sup>

En una de sus versiones más conocida, se dice que la paradoja se produce porque mientras el sentido común sugiere que los montones de arena tienen las siguientes propiedades, éstas son inconsistentes entre sí:

P. Dos o tres granos de arena no son un montón.

Q. Un millón de granos de arena juntos sí son un montón.

R. Si  $n$  granos de arena no forman un montón, tampoco lo serán  $(n+1)$  granos.

---

<sup>38</sup>. Mi tratamiento de la paradoja sorites está inspirado en el trabajo de Delia Graff Fara (2000).

En otras palabras, si bien el sentido común sugiere aceptar como verdaderas estas tres proposiciones, de su conjunción se sigue una contradicción, a saber, (S) que un millón de granos de arena juntos son y no son un montón. Esto es un problema porque, idealmente, quisiéramos una teoría de los montones (y términos similares) que respetara nuestras intuiciones de sentido común sin comprometernos a aceptar una contradicción.

Para hacer aún más obvio el paralelismo entre el sorites y las paradojas antes vistas, pensemos en él como proponiendo una tensión entre una tesis existencial – que hay una diferencia entre lo que es un montón y lo que no lo es – y la conjunción de una serie de juicios particulares – que la diferencia no es entre  $n$  y  $n+1$  granos – que parecen excluir la instanciación del existencial. En otras palabras, detrás del sorites escuchamos la voz de Proclo preguntando ¿cómo es posible que digamos que hay un límite a lo que es un montón, pero no podamos decirme cuál? Aplicando el mismo esquema que aplicamos a las paradojas anteriores, tenemos que la paradoja sorítica<sup>39</sup> se basa en el error de creer que es inconsistente (1) sostener una creencia existencial, a saber, que hay un límite al número de granos de arena que forman un montón, y (2) una serie exhaustiva de creencias singulares sobre el dominio del existencial de dicha creencia, es decir, que para todo número  $n$  entre 2 y un millón, creemos que  $n+1$  no es el número límite de granos de arena que forman un montón; y al igual que ellas, se resuelve reconociendo que hay una premisa intermedia (3) de la forma ‘Hay un/ unos  $X$  que  $S$  cree es el/los  $F$ ’ tal que (1) no implica a (3), así que aunque (2) y (3) implican cierta inconsistencia por parte del sujeto, (1) y (2) no. En este caso, la falsa premisa intermedia es:

3. Hay un número  $n$  de granos de arena que creemos es el número límite de granos de arena que forman un montón.

---

<sup>39</sup>. Los  $X$  son números naturales (entre 2 y un millón), y  $F$  es la propiedad de ser el número límite de granos de arena que forman un montón.

En otras palabras, el error detrás de la paradoja sorítica es pensar que quién cree que hay una línea divisoria entre lo que cae dentro de la extensión de un término y lo que cae fuera de él debe creer también de algún lugar que ahí se encuentra dicha línea divisoria. Espero haberlos convencido de que no es así, y que por lo tanto, nuestros juicios de sentido común no nos llevan a ninguna contradicción.

#### **IV. Observaciones finales a forma de moraleja**

En esta plática he propuesto una nueva manera de concebir la aceptación de proposiciones y su relación con la creencia y la inferencia. He mostrado además que adoptar cierta propuesta nos permite resolver de manera relativamente sencilla algunas paradojas aparentemente tan disímbricas como la del prefacio, la lotería y el *sorites*. Reconozco que aún falta mucho para desarrollar mi propuesta y mostrar otras ventajas que justifique su adopción. Sin embargo, no quiero terminar sin decir algo más sobre la dirección hacia la que quiero llevar mi propuesta.

Mi objetivo es criticar cierta manera tradicional de entender la inferencia y las actitudes proposicionales – incluyendo la creencia y la aceptación –, a saber, como una relación entre un agente y una proposición. En este sentido, hay cierta continuidad entre lo que aquí propongo y lo que mi colega Ángeles Eraña ha propuesto esta mañana: ambos pensamos que la epistemología tradicional se ha equivocado al colocar a los individuos y sus actitudes hacia proposiciones al centro de su análisis, tratando de reducir el resto de los fenómenos epistémicos (el estado epistémico total de un individuo, la vida epistémica de los grupos sociales, el aprendizaje, la inferencia, etc.) a actitudes proposicionales individuales y relaciones entre ellas. Ella ha puesto el acento en la irreducible dimensión social de lo epistémico, y aquí he tratado de argumentar a favor de la subdeterminación de los contenidos de nuestras actitudes epistémicas por las proposiciones.

En el modelo tradicional, a la base están las actitudes proposicionales de los agentes individuales. El estado epistémico de un individuo en un momento dado es reducible al conjunto de actitudes proposicionales que tiene. La vida epistémica de un grupo a su vez se reduce, en este

modelo tradicional, a la vida epistémica de sus miembros individuales. En su lugar, proponemos una visión de la vida mental, dónde la actividad epistémica del grupo es primaria, y sólo en ella podemos entender la del individuo; donde las actitudes proposicionales no son básicas, sino derivadas; y en su lugar la unidad básica de análisis de nuestras representaciones mentales es mas amplia – lo que he llamado, siguiendo a Jon Barwise y John Perry (1983), una situación. En esta plática solamente he empezado a bosquejar esta última parte, pero no quería desaprovechar la oportunidad para mencionar un poco sobre el proyecto más amplio en el que quiero enmarcar mi trabajo.

## REFERENCIAS

- Alston, J. (2000), *Illocutionary Acts and Sentence Meaning*. Ithaca: Cornell University Press.
- Austin, J. L. (1956, 1957) "Un alegato en pro de las excusas" en *Ensayos filosóficos*, compilados por J. O. Urmson y G. J. Warnock; traducción y presentación de Alfonso García Suárez, Madrid: Revista de Occidente, 1975
- Barceló, Axel (2017) "Inferencia y Presuposición", manuscrito bajo dictamen.
- (2015) "Making Quantified Truths True", en Torza, Alessandro (Ed.) *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library, 323-338
- Barceló, Axel y Robert Stainton (2018) "Quasi-Factives and Cognitive Efficiency", en *Relevance: Pragmatics and Interpretation*, editado por Kate Scott, Billy Clark y Robyn Carston, Cambridge University Press.
- Barwise, Jon & John Perry, (1983) *Situations and Attitudes*, MIT Press.
- Berker, Selim. (2013) "The Rejection of Epistemic Consequentialism." *Philosophical Issues* 23 (1) (October): 363–387
- Brandom, Robert (1994), *Making it Explicit*. Cambridge: Harvard University Press.
- Bortolotti, Lisa (2016) "The Ethics of Delusional Belief" *Erkenntnis*, April 2016, Volume 81, Issue 2, pp 275–296
- Deleuze, Gilles (2001) "Presentación de Sacher-Masoch." *Lo frío y lo cruel*, Harvard: 84-93.
- Descartes, René (1641) *Meditaciones metafísicas con objeciones y respuestas*, trad. de Peña García, Vidal. Alfaguara, Madrid, 1993
- (1637) *Dislibro del método para bien dirigir la razón y buscar la verdad en las ciencias*, introducción de Miguel Angel Granada. Traducción y notas de Eugenio Frutos. Ed. Planeta, Barcelona, 1984, 205 págs.
- Douven, I. (2006), "Assertion, knowledge, and rational credibility", *Philosophical Review* 115: 449–485.
- Dowell, Janice (manuscrito) "When are Truth-assessments Data for Semantic Theorizing about Epistemic Modals?"
- Dutant, Julien (2015), "The legend of the justified true belief analysis." *Philosophical Perspectives*, 29: 95–145.  
doi:10.1111/phpe.12061

- Edmonds, M. Leslie (1987) “Superstition to Science: The Role of Research in Strengthening Public Library Service to Children”, *Library Trends*, 35 (3): 509-520.
- Earl, Dennis (2017) “Concepts” en *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, <http://www.iep.utm.edu/>, retrieved on January 8th, 2017.
- Eraña, Ángeles (2003) “Normatividad epistémica y estructura heurística del razonamiento”, *Crítica, Revista Hispanoamericana de Filosofía*. Vol. 35, No. 104 (agosto 2003): 69–108
- Fara, Delia Graff (2000) “Shifting Sands: An Interest-Relative Theory of Vagueness”, *Philosophical Topics* 28(1): 45–81.
- Franssen, Maarten (2006) “The Normativity of Artefacts”, *Studies in History and Philosophy of Science*, 37: 42-57.
- García-Aguilar, Claudia Lorena (2004) “Especificidad de dominio y sesgo en la psicología del razonamiento humano”, *Signos Filosóficos*, vol. VI, núm. 12, julio-diciembre, 2004, pp. 63-91
- Gardner, John (2007) *Offences and Defences*, Oxford University Press.
- Gómez-Torrente, Mario, (2010), “The Sorites, Linguistic Preconceptions, and the Dual Picture of Vagueness”, en R. Dietz y S. Moruzzi (eds.), *Cuts and Clouds. Essays in the Nature and Logic of Vagueness*, Oxford University Press, Oxford, 228-253.
- Greco, John (2010) *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*, Cambridge University Press.
- Green, Mitchell S. (1999) “Attitude Ascription’s Affinity to Measurement”, *International Journal of Philosophical Studies* 7(3): 323-348.
- Hansson, Ove; Karin Edvarsson Björnberg y John Cantwell (2016) “Self-Defeating Goals”, *Dialéctica*, Volume 70, Issue 4, December 2016, Pages 491–512
- Hernández, Héctor (2012) *Las Fascinantes Paradojas de la Probabilidad: Concepto, resolución e importancia en las ciencias y en la filosofía*, Editorial Académica Española.
- Hetherington, Stephen (2017) “Fallibilism”, *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, <http://www.iep.utm.edu/>, consultada el 7 de enero.
- Korsgaard, Christine M. et.al. (2000) *Las fuentes de la normatividad*, traducida por Laura Lecuona y E. Manríquez, traducción revisada por Faviola Rivera, Instituto de Investigaciones Filosóficas, Colección:Filosofía Contemporánea, Ciudad de México.

- Kroedel, Tomas (2013) “The Lottery Paradox, Epistemic Justification, and Permissibility”, *Logos & Episteme* 4(1): 103-111.
- Langton, Rae & Caroline West (1999) “Scorekeeping in a pornographic language game”, *Australasian Journal of Philosophy*, 77:3, 303-319
- Littlejohn, Clayton (2013) “Don’t know, don’t believe: reply to Kroedel”, *Logos & Episteme* 4(2): 231-238.
- Littlejohn, Clayton (*en prensa*) “The Unity of Reason”, en John Turri (ed.), *Epistemic Norms: New Essays on Action, Belief, and Assertion*.
- MacFarlane, John (2014). *Assessment Sensitivity: Relative Truth and its Applications*. Oxford University Press.
- Mackenzie, Donald (2001) *Mechanizing Proof: Computing, Risk, and Trust*, The MIT Press.
- Martin, Justin W. & Fiery Cushman (2015) “The Adaptive Logic of Moral Luck”, en Justin Sytsma y Joshua Knobe (eds.) *Blackwell Companion to Experimental Philosophy*, Blackwell.
- Morado, Raymundo (2000) “Construcción, Reconstrucción y Evaluación de Argumentos”, apuntes del 26 de mayo del 2000 para el libro *Filosofía del Lenguaje (Análisis Lógico de Argumentos)*, Colegio de Filosofía, Facultad de Filosofía y Letras. Julio a Noviembre del 2000, Grupo 0021, clave 0841 (3416 para estudiantes del plan 1999). Consultado en línea el 13 de febrero de 2017 en <http://www.filosoficas.unam.mx/~morado/Libros/0Argumentos/cons.htm>
- Morton, Adam (2006) Review of Alan Miller’s *Understanding People*, *Mind* 115 (459) 780
- Nagel, Thomas, (1979), “Moral Luck”, en *Mortal Questions*, Nueva York, Cambridge University Press.  
Traducción al castellano: La muerte en cuestión, México, FCE, 1981. Se ha reeditado con el título Ensayos sobre la vida humana, 2000.
- Nozick, R. (1993), *The Nature of Rationality*, Princeton, NJ: Princeton University Press.
- Padilla, Hugo (1971) “Condiciones para saber. (Sobre un artículo de Edmund L. Gettier)”, *Crítica, Revista Hispanoamericana de Filosofía*, 5(13): 129-134.
- Peacocke, C. (2005). Rationale and Maxims in the Study of Concepts, *Noûs*, 39 (1): 167–78.
- Peirce, Charles Saunders (1868) “Some Consequences of Four Incapacities.” *The Journal of Speculative Philosophy* 2(3):140-157.
- Priest, Graham (1985) “Inconsistencies in motion”, *American Philosophical Quarterly* 22, no. 4: 339-346.

- Rescorla, Michael (2009), "Assertion and its Constitutive Norms", *Philosophy and Phenomenological Research*, Vol. LXXIX No. 1, July 2009, pp. 98 - 130.
- Rosell, Sergi (2006) "Nagel y Williams acerca de la suerte moral", *Revista de Filosofía* 1(1): 143-165
- Rudy, Fernando (manuscrito)
- Ryle, Gilbert (1953) "Ordinary Language", *The Philosophical Review*, Vol. 62, No. 2 (Apr., 1953), pp. 167-186
- Sartorio, Carolina, (2015) "Resultant Luck and the Thirsty Traveler", *Methodes* 6, 153-171.
- Searle, John (1969), *Speech Acts*. Cambridge: Cambridge University Press.
- (2001), *Rationality in Action*. Cambridge: MIT Press.
- Sellars, Wilfred (1963), "Some reflections on language games", reimpresso en *Science, Perception, and Reality*. New York: Routledge and Kegan.
- Steglich-Petersen, Asbjørn (2006) "No Norm Needed: on the Aim of Belief", *The Philosophical Quarterly* 56 (225), 499–516.
- Strawson, Peter Frederick (1959) *Individuals: An Essay in Descriptive Metaphysics*. Londres, Methuen.
- Travis, Charles (1985) "On What Is Strictly Speaking True", *Canadian Journal of Philosophy*, 15 (2):187 - 229.
- van Eemeren, Frans H. and Houtlosser, Peter, (2003) "More about Fallacies as Derailments of Strategic Maneuvering: The Case of Tu Quoque" OSSA Conference Archive. 93.
- Velleman, D. (2000) *The Possibility of Practical Reason*, Oxford University Press.
- Williams, Bernard (1981): "Moral Luck", en *Moral Luck*, Cambridge, Cambridge University Press.
- Traducción al castellano: La fortuna moral, México, UNAM, 1993.
- Williamson, Timothy (2000) *Knowledge and its Limits*. Oxford: Oxford University Press.
- Zheng, Robin y Nils-Hennes Stear (manuscrito) "Imagining in oppressive contexts, or, what's (intrinsically) wrong with "Blacking Up"?"
- Zuluaga, Mauricio (2005) "El Problema de Agripa", *Ideas y Valores* 54 (28): 61-88.