

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
POSGRADO EN FILOSOFÍA DE LA CIENCIA**

**Racionalidad, reglas de razonamiento  
y equilibrio reflexivo**

**TESIS QUE PARA OPTAR POR EL GRADO DE  
MAESTRO EN FILOSOFÍA DE LA CIENCIA**

**PRESENTA**

**Alejandro Vázquez del Mercado Hernández**

**Tutora: Dra. María de los Ángeles Eraña Lagos**  
Instituto de Investigaciones Filosóficas

**MÉXICO D.F., MAYO DE 2013**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

# Contenidos

1. INTRODUCCIÓN .....	3
2. RACIONALIDAD Y REGLAS DE RAZONAMIENTO .....	8
La racionalidad como objeto de estudio .....	8
La racionalidad y las relaciones inferenciales .....	12
El razonamiento en un marco de reglas .....	25
3. ESTRATEGIA PARA VALIDAR REGLAS DE RAZONAMIENTO .....	34
Criterio de corrección.....	34
Métodos basados en Intuiciones.....	37
Las fuentes de la normatividad epistémica y el problema normativo .....	41
4. EL EQUILIBRIO REFLEXIVO COMO MÉTODO .....	46
Equilibrio reflexivo histórico y clásico .....	46
El equilibrio reflexivo naturalizado .....	58
Equilibrio reflexivo naturalizado informado y experimentalismo moderado .....	66
5. CONCLUSIÓN .....	73
6. BIBLIOGRAFÍA .....	76

## 1. INTRODUCCIÓN

La racionalidad epistémica no es solamente una noción filosófica de gran importancia, sino que compete a muchas otras disciplinas e incluso forma parte de la vida cotidiana. Regularmente apelamos a la racionalidad de un modo normativo: para guiar nuestras acciones y decisiones. Consideramos que ciertas conductas son racionales y otras, por el contrario, irracionales. Cuando alguien llega a una conclusión que no parece estar garantizada por sus premisas, le indicamos que está siendo irracional; al realizar planes educativos, pensamos que deben fomentar la racionalidad de los estudiantes; y frecuentemente atribuimos muchas de las sorprendentes capacidades que tiene la especie humana a un mayor grado de racionalidad, o incluso —de manera un tanto chauvinista— a que es la única racional.

Realizar inferencias correctas es crucial no sólo en el ámbito privado, sino en la vida social. De hecho, muchos de los métodos—tradicionalmente considerados como adecuados— que tenemos para llegar a conclusiones médicas han sido fuertemente cuestionados (Gigerenzer: 2008, p. 17), lo mismo que en ámbitos legales, económicos (Sen: 1977) o políticos (Pettit: 2001). Tomando esto en consideración, no parece nada ingenuo pensar que el estudio de la racionalidad sea relevante para el avance social.

Por todo lo anterior, no resulta sorprendente que dentro de las ciencias cognitivas la racionalidad sea un área de interés. Usualmente la psicología cognitiva es la disciplina que se ocupa mayoritariamente del tema, pero no es la única. La neuroeconomía, por ejemplo, busca entender el modo como tomamos decisiones utilizando imágenes cerebrales. Asimismo, en el área de la inteligencia artificial se buscan modos de crear agentes que puedan replicar al menos algunas de nuestras capacidades de pensar. Incluso en las partes más abstractas de la computología, tales como la teoría de autómatas, se puede pensar que parte de lo que se hace es un tratamiento formal de la noción de racionalidad.

En el caso de las ciencias sociales, la racionalidad también es una noción muy importante, quizá de manera más prominente (aunque no exclusiva) en la economía, donde la conducta de los agentes económicos se modelan a partir de ciertos presupuestos de maximización,

ayudándose de disciplinas formales como la teoría de la decisión o la teoría de juegos. Esto puede dar la impresión de que en este tipo de disciplinas la noción relevante de racionalidad no es la epistémica, sino la práctica, aquella que tiene que ver con obtener las mejores cosas para satisfacer los deseos propios. Empero, en la economía también existen nociones epistémicas, tales como las creencias de un agente sobre lo que realizarán los otros agentes. Si bien en la economía estándar predomina una actitud instrumentalista —según la cual lo único que importa es que las predicciones sean exitosas— en ocasiones se dice que la validez de la teoría económica depende de que sean razonables estos presupuestos.

Lo anterior muestra la importancia de estudiar con precisión el concepto de racionalidad epistémica. No es extraño que una noción que aparece en diversas disciplinas, y que no ha sido satisfactoriamente definida por ninguna de ellas, sea de interés para la filosofía. Mientras que otras disciplinas como las mencionadas con anterioridad estudian la racionalidad presente en agentes naturales o la que puede ser construida en agentes artificiales, la filosofía estudia el concepto mismo de racionalidad. Dentro de la propia filosofía existen diversas maneras de analizar la racionalidad epistémica según el aspecto que se quiera estudiar: en la filosofía de la mente se examinan los aspectos de la racionalidad como parte de la cognición; en la epistemología se estudia como un concepto normativo acerca de cómo razonar adecuadamente. ¿Qué clase de creencias tenemos la obligación de adquirir o abandonar?, ¿qué hace que un razonamiento sea correcto? Este tipo de preguntas motivan una investigación como la presente.

Como señalé anteriormente, el objeto de estudio de esta tesis es la racionalidad epistémica. Por lo general se contrastan dos tipos distintos de racionalidad: práctica y epistémica. La primera trata de esclarecer cuáles son las acciones más apropiadas para obtener un fin, mientras que la segunda tiene que ver con establecer cuándo es adecuado adquirir una creencia o sostenerla.

Sin embargo, esta división no es completamente tajante. Algunos casos donde se puede ver es en la tradición de epistemología bayesiana<sup>1</sup>, las explicaciones consecuencialistas de la

---

<sup>1</sup> En el tratamiento probabilístico de las creencias y de la racionalidad, normalmente llamado

normatividad epistémica (e.g. Goldman: 1988), las justificaciones evolutivas de la racionalidad (e.g. Dennett: 1986), o la propuesta de Grimm (2009) de considerar el valor de las creencias verdaderas como bienes de una comunidad. Sin embargo la taxonomía original sobre racionalidad epistémica y práctica es un útil como un punto de partida en la investigación. El objeto de estudio aquí es la racionalidad en relación con las creencias, el conocimiento, y las inferencias entre las mismas; empero, no hay razones para presuponer de antemano que hay una distinción tajante con respecto a la racionalidad práctica.

La primera parte de esta investigación busca responder a grandes rasgos la pregunta acerca de qué es la racionalidad epistémica, en adelante, referida simplemente como racionalidad. El marco de trabajo utilizado aquí consiste en entender la racionalidad como un tipo de razonamiento que establece relaciones inferenciales correctas entre un estado doxástico y otro; dichas relaciones son —al menos parcialmente— constitutivas de las creencias y de la agencia epistémica. Asimismo, será necesario mostrar de qué modo se debe entender una regla y cómo es que los razonamientos pueden capturarse por medio de reglas. En esta parte será también preciso explicar por qué los modelos de la racionalidad usados habitualmente, es decir, la lógica clásica de primer orden y la probabilidad, son inadecuados. Para esto será útil retomar algunos argumentos de Christopher Cherniak (1990).

En la segunda parte, será necesario establecer cuál debe ser el criterio de corrección para saber cuáles reglas de razonamiento son las correctas. El desiderátum principal es que el criterio provisto debe satisfacer las exigencias de la normatividad epistémica: dicho de otro modo, ¿qué es lo que le da fuerza a este mandato? El tipo de normatividad que se desarrolla aquí es consecuencialista, tal como la han propuesto diversos autores de manera reciente (e.g. Sosa: 2009) y por otra parte, el criterio de corrección está basado en las intuiciones epistémicas. Dicho esto último desde un punto de vista metodológico: las reglas deben ser obtenidas con un método basado en intuiciones (o derivadas a partir de un criterio intermedio obtenido por intuiciones). De acuerdo con una meta-epistemología consecuencialista, cuando

---

bayesiano por su recurso constante al teorema de Bayes y en general al principio de condicionalización. En esta literatura, los axiomas de la probabilidad son demostrablemente equivalentes a la coherencia en un sistema de apuestas, y donde las creencias mismas se definen en la disposición eventual a apostar que tendría un sujeto idealmente racional.

realizamos evaluaciones epistémicas de un agente, lo hacemos juzgando si actúa en concordancia con lograr la consecución de ciertos fines (e.g. típicamente, la obtención de creencias verdaderas). Las intuiciones epistémicas proveen normatividad porque reflejan prácticas y valores que se ajustan a estos mismos fines. Todo esto es importante, pues para elaborar una teoría de la racionalidad, es crucial tener una explicación de la normatividad epistémica en concordancia con un método que permita obtener las reglas de razonamiento.

La tercera y última parte de esta investigación se concentrará en el aspecto metodológico acerca cómo debe proceder una teoría de la racionalidad que utilice las intuiciones epistémicas como insumos para establecer una regla correcta. Para ello será propuesta una variante naturalista del equilibrio reflexivo, un método originalmente propuesto por Nelson Goodman en 1955 (Cfr. Goodman: 1983). En esta parte se mostrarán algunos problemas que ha presentado el equilibrio reflexivo interpretado de un modo racionalista tal como se ha desarrollado en buena parte de la literatura (e.g. Cohen: 1989) y de las ventajas que puede tener un equilibrio reflexivo naturalizado como el que proponen autores como Goldman (2007) y Bates (2005). A diferencia del equilibrio reflexivo clásico, el naturalizado es un método que justifica su validez a posteriori. En éste, las intuiciones epistémicas son tomadas como intuiciones semánticas ordinarias (o *folk*) a partir de los conceptos epistémicos que se poseen como estructuras mentales. Una característica fundamental del equilibrio reflexivo naturalizado, es que debido a la noción psicológica de concepto utilizada es necesario mostrar, por una parte, que las intuiciones epistémicas son estables, y por otra, que son compartidas; puesto que dichos conceptos deben expresar los valores y prácticas que provean la normatividad propia de la racionalidad. Para este punto resultará relevante revisar la evidencia provista por la epistemología experimental, es decir, el estudio empírico de las intuiciones epistémicas por métodos estadísticos (e.g. Stich, Nichols, Weinberg: 2001), así como algunas propuestas sobre cómo interpretarla.

Este trabajo no elabora una teoría concreta y detallada sobre la racionalidad. El objetivo es mostrar en términos generales qué es la racionalidad y cómo se debe proceder en su estudio para conocer las reglas concretas que la caracterizan; es decir, esta investigación defiende un programa de investigación sobre la racionalidad que cubra todos los aspectos importantes

(meta-epistemológicos y metodológicos) de manera explícita. De manera un tanto sorprendente, las conclusiones son en cierto modo conservadoras, pero también en parte revisionistas. Puesto que las intuiciones epistémicas se consideran el fundamento de lo que hace a algunas reglas correctas, esto se encuentra en armonía con una gran parte de la epistemología tradicional (o analítica) que se ha llevado a cabo en las últimas décadas. Sin embargo, puesto que aquí propongo que la justificación de esta metodología sea a posteriori, surge la necesidad de tener evidencia empírica relacionada con dichas intuiciones, por lo que será necesario adoptar una actitud cautelosa hacia esta tradición teórica. Así, el naturalismo a medio camino propuesto por esta investigación deja abierta la puerta para estudiar la racionalidad retomando resultados valiosos de la epistemología realizada con intuiciones, pero poniéndolos a prueba y reevaluándolos continuamente como cualquier dato empírico.

## **Agradecimientos**

Esta tesis fue realizada con la beca de maestría del Consejo Nacional de Ciencia y Tecnología y la beca de fomento a la titulación de de la Universidad Nacional Autónoma de México. También fue parte del trabajo de investigación llevado a cabo durante mi estancia como estudiante asociado del Instituto de Investigaciones Filosóficas de esta misma universidad. Agradezco a Ángeles Eraña por su apoyo incondicional como tutora, a Axel Barceló, Miguel Ángel Fernández, Claudia Lorena García y Eduardo García-Ramírez por sus correcciones detalladas; a José Alfredo Amor, Edgar González-Varela, Atocha Aliseda, Faviola Rivera, Alberto Ross, Luis Guerrero, Rafael Núñez, Roberto Rivadeneyra, Rafel García y Héctor Zagal por su disposición atenta y su ayuda durante los estudios de maestría. Agradezco el apoyo administrativo de Jorge Linares, Fabiola Villela, Noemí Vidal y Elizabeth Barajas. También a los colegas que leyeron y comentaron versiones de este texto: Teresa Rodríguez, Ingrid Constant, David Fajardo, Moisés Macías, Alejandro Cavallazzi, Azucena Palavicini, Antonio Pardo y Jorge Morales. Sólo me queda agradecer de todo corazón a mi familia y amigos por su apoyo incondicional durante este periodo, particularmente a mis padres Gloria y Alejandro, a Fátima y a Montserrat.



## 2. RACIONALIDAD Y REGLAS DE RAZONAMIENTO

### La racionalidad como objeto de estudio

En este apartado mostraré cuál es el objeto de estudio del presente trabajo y daré algunas aclaraciones sobre el modo en que me aproximaré al mismo. Para ello voy a esbozar una explicación de la racionalidad que se basa en los requisitos normativos existentes entre las creencias de un agente y las inferencias que éste debe de realizar. Dicha explicación será tomada de la propuesta de racionalidad mínima de Cherniak (1990).<sup>2</sup>

En otras palabras, se ilustrará qué significa que una creencia sea racional. Algunos aspectos relevantes para este análisis podrían ser los siguientes ejemplos, tomados de distintas teorías: la manera mediante el cual se adquiere una creencia (fiabilismo de método), las relaciones de apoyo en las creencias que forman parte de la estructura doxástica de un agente (coherentismo), y las relaciones materiales de incompatibilidad e implicación entre las creencias (inferencialismo). El objetivo final es explicar esta noción en términos que resulten más claros.

El problema de la racionalidad epistémica<sup>3</sup> consiste en dar una respuesta normativa y descriptiva acerca de la formación y el mantenimiento de las creencias en tanto que resultan adecuadas desde el punto de vista de la epistemología. Un estudio descriptivo busca esclarecer cómo es que de hecho ciertos agentes son racionales y qué hace que lo sean; mientras que un estudio normativo busca establecer cómo deberían ser las inferencias o las creencias de un agente para ser racional. ¿Cómo puede verse esta diferencia? Por ejemplo, es plausible que los seres humanos seamos en cierto grado racionales, pero que existan estándares de racionalidad aún más altos que no cumplimos. Es común considerar que el problema descriptivo pertenece a las ciencias empíricas y que el problema normativo

---

<sup>2</sup> Hasta donde tengo conocimiento, la única exposición completa de esta teoría se encuentra en el trabajo citado de Cherniak, aunque aparece explicada en trabajos de otros autores, como Stich (1993), Eraña (2002).

<sup>3</sup> En adelante siempre que use el término “racionalidad” estaré hablando de racionalidad epistémica, cuando me refiera a otros tipos de racionalidad lo indicaré explícitamente.

pertenece a la filosofía. Si bien es verdad que el estudio descriptivo de la racionalidad tiene una parte empírica relacionada con las ciencias cognitivas, la filosofía tiene un papel relevante al momento de ayudar a determinar una noción de racionalidad lo suficientemente clara. Por otra parte, no existe un consenso respecto a cuál ha de ser el papel que deban tener las ciencias cognitivas en el estudio normativo de la racionalidad. Se ha sugerido desde un papel auxiliar (Goldman: 1988), hasta un papel principal donde el problema normativo es completa o casi completamente empírico (postura por lo general atribuida a Quine, Cfr. 1969). Independientemente de estas cuestiones, siguiendo a Harman (1988), considero que en un estudio del concepto de racionalidad no es siempre posible (ni útil), distinguir tajantemente entre la parte normativa y la parte descriptiva del problema. Por tanto, esta distinción familiar no se enfatizará en este trabajo.

Tanto la racionalidad epistémica como la racionalidad práctica —una noción análoga respecto a la normatividad de las acciones y decisiones— resultan ser nociones poco claras en cuanto a su contenido, es decir, aún desde un punto de vista superficial. Tanto en la filosofía, como en distintas disciplinas (e.g. economía), e incluso en el habla cotidiana, la racionalidad se usa en sentidos muy diversos, sin que quede enteramente claro si se trata de uno o varios conceptos, así como la relación que pueda existir entre los mismos.

Acotando la discusión únicamente en su acepción epistémica, los contextos en los cuales se utiliza siguen siendo muy diversos, sin un consenso acerca de cuál es el significado preciso del término. Esto ha llevado a algunos epistemólogos (Goldman: 1988) a tener dudas legítimas respecto a que se trate de una noción filosóficamente interesante. El término “racional” en su sentido epistémico es un calificativo que puede aplicar tanto a agentes (naturales y artificiales), creencias particulares, conjuntos de creencias, inferencias, estrategias inferenciales, métodos de conocimiento y teorías. De manera inmediata, no parece existir algo en común a todos estos contextos, salvo que es un uso epistémico aprobatorio.<sup>4</sup>

---

<sup>4</sup> El problema puede ser más profundo que la confusión señalada por Goldman. Expresando esto del modo como se llevan a cabo las discusiones en meta-ética, se podría sospechar que hay una explicación no-cognitivista del término “racional”, o concretamente emotivista, donde un sujeto le

Si bien el adjetivo “racional” se utiliza tradicionalmente para realizar evaluaciones epistémicas (incluso de manera cotidiana), el eliminativista argüirá que ésta puede y debe realizarse con otras herramientas y criterios, ya sea de la epistemología o externos. Existen ejemplos de ambas cosas en la literatura. En el texto anteriormente citado (1988), Goldman considera otros términos de aprobación [appraisal] epistémica son suficientes (principalmente la justificación).

En un texto más reciente (2004), el mismo Goldman muestra cómo una caracterización formal y estipulada de la racionalidad colectiva hecha por Pettit (2001) resulta inadecuada como término de aprobación. Los argumentos de Goldman están dirigidos a mostrar que ser *racional* —en el sentido técnico definido por Pettit— para un grupo es irrelevante desde el punto de vista de la epistemología social. En este mismo texto, Goldman muestra que el *conocimiento colectivo* es una noción más adecuada para realizar evaluaciones. Con esto quiero ilustrar cómo la racionalidad, a pesar de ser uno de los términos más usados para evaluaciones epistémicas sobre el comportamiento de agentes, no es forzosamente un término indispensable: los eliminativistas de la racionalidad han ofrecido maneras alternativas de realizar dichas evaluaciones.

Un segundo ejemplo sobre cómo procede el eliminativismo respecto a la racionalidad es la teoría pragmática de evaluación cognitiva de Stich (1993). Según ésta, un sistema de reglas de razonamiento debe ser juzgado como adecuado de acuerdo a una finalidad concreta, haciendo preguntas del tipo ¿qué conjunto de inferencias es el más adecuado para lograr  $x$ ? La teoría sostiene que  $x$  puede ser cualquier objetivo, y es indiferente si se trata de una finalidad tradicionalmente considerada como epistémica (como creencias verdaderas). Una vez que se escoge *cualquier* objetivo (e.g. ganar dinero, vivir el mayor tiempo posible, etc.), una determinada manera de razonar es evaluada según qué tanto ayuda para obtener dicho objetivo. Tomemos un ejemplo. Existe evidencia de que los sujetos occidentales tienden a enfatizar el aspecto causal en sus razonamientos que los habitantes de otras culturas

---

llama a algo “racional” cuando le genera un sentimiento positivo desde el punto de vista epistémico. Por ejemplo, a un sujeto llama irracional a su vecino porque le molesta que se forme creencias sobre política internacional leyendo fuentes poco confiables, sin que el término “irracional” tenga un contenido semántico más allá de esta molestia.

(Nisbett: 2003). ¿Esto es un buen modo de razonar? De acuerdo con el eliminativismo de Stich, el modo en que razona un sujeto occidental típico no puede ser evaluado como bueno o malo sino en relación a un objetivo cualquiera. Los razonamientos que enfatizan más el aspecto causal podrían ser buenos para lograr un dominio del medio ambiente, pero malos (o peores que las maneras alternativas de razonar) para alcanzar la armonía social.

Tanto en el caso de Stich como en el de Goldman, no se niega que tenga sentido preguntarse cómo es posible evaluar de manera operativa o procedimental a los agentes cognitivos. Por ello, se acepta un explanandum —el hecho que algunos procedimientos inferenciales sean mejores que otros— a pesar de que se niega que la racionalidad sea una noción correcta para dar cuenta de dicho explanandum.

Este trabajo, en cambio, sigue la vía más tradicional de asumir que este explanandum puede englobarse bajo la noción de racionalidad, y que una teoría de la racionalidad responderá a las preguntas más importantes acerca de qué hace que ciertas inferencias sean adecuadas y otras inadecuadas. Empero, para que la racionalidad epistémica pueda ser una noción trabajable, será necesario elaborar algunas delimitaciones preeliminarias.

En este trabajo, la racionalidad será entendida como algo que se refiere a los individuos y a sus creencias ordinarias. Es decir, será pertinente concentrarse en casos del tipo: por qué podría ser racional para alguien pensar que por la mañana llovió afuera de su casa; y no del tipo por qué es más racional aceptar la mecánica relativista que la mecánica newtoniana (o por qué es racional que la comunidad de físicos acepte una y no la otra). Esto logra distinguir la presente investigación de una familia de problemas relacionados con la filosofía de la ciencia, como la elección de teorías o la racionalidad del conocimiento científico. Si una teoría sobre la racionalidad referida a creencias ordinarias se puede extender para resolver problemas sobre epistemología de la ciencia es una cuestión ajena a la presente investigación.

Aún acotando el problema a las situaciones epistémicas ordinarias, existe una divergencia importante entre los usos del término “racional.” Por una parte, se dice que los agentes pueden ser racionales, pero también sus creencias, y los razonamientos que realizan. En el

siguiente apartado explicaré de qué manera puede entenderse la racionalidad de un modo que estará ligado a las transiciones entre creencias; pero que al mismo tiempo, está relacionado con la constitución de un agente epistémico y con los requisitos epistémicos acerca de las creencias. El término “racional” podrá utilizarse para calificar tanto creencias, como inferencias y agentes, sin embargo mostraré que estas tres áreas de aplicación del término están fuertemente relacionadas.

## **La racionalidad y las relaciones inferenciales**

La intuición que subyace al intento de caracterizar la racionalidad —al menos en parte— como una relación entre creencias, surge de la idea de que la agencia epistémica no sería posible sin esto. La agencia epistémica puede entenderse, a grandes rasgos, como el hecho de que un sujeto sea capaz de realizar ciertas acciones —o de manera converso, que ciertas acciones sean atribuibles a un sujeto— relacionadas con el aspecto normativo de las creencias. El propósito aquí es desarrollar una explicación, al menos parcial, que dé cuenta simultáneamente de cómo se constituye un agente, una creencia y un razonamiento correcto a través de relaciones inferenciales.

La idea puede ilustrarse con un ejemplo. Pensemos en un sujeto que escucha un ladrido afuera de su casa y a partir de ello se forma la creencia de que hay un perro afuera. Asimismo, a partir de estas creencias se forma otras, como por ejemplo la creencia de que el perro no trae bozal, o de que hay algo que está haciendo ladrar al perro. De acuerdo con el tipo de teoría que se pretende esbozar aquí, al menos una parte de lo que hace que la creencia “Hay un perro afuera” sea una creencia, es que produce otras creencias relacionadas y que fue producida por una de ellas; lo que hace que el sujeto sea un agente epistémico es su capacidad de pasar de ciertas creencias a otras de un modo que al menos no resulta completamente arbitrario. Atribuiremos racionalidad si estas transiciones se llevan a cabo de una cierta manera, si son razonamientos correctos, los cuales a su vez llevan a constituir creencias y agentes racionales. Por supuesto, una explicación de este tipo no excluye la presencia de otro tipo de elementos: por ejemplo, las creencias podrían requerir

elementos representacionales o disposicionales; la agencia epistémica a su vez podría depender de una noción de acción, etc.

A continuación será conveniente ofrecer algunas alternativas existentes en la literatura acerca de cómo deben ser estas relaciones inferenciales. Cabe señalar que la primera alternativa aparece tan sólo como caso límite, donde no hay exigencias del tipo que acabo de describir.

a) **Requisitos vacíos:** En esta caracterización, la noción misma de creencia está desligada de las relaciones inferenciales entre creencias. El teórico de requisitos vacíos únicamente pide que el agente afirme la proposición que cree, o que tenga la disposición a afirmarla. Sería necesario, por tanto, dar una teoría alternativa acerca de qué constituye una creencia. De acuerdo con Cherniak (1990), Russell defendió que una creencia es aquello que un sujeto afirma explícitamente o afirmaría en caso de ser cuestionado.

b) **Racionalidad mínima:** Existen algunos requisitos inferenciales, de consistencia o de otro tipo (según la teoría) para que algo pueda contar como una creencia. Sin embargo basta con que el agente haga las inferencias (o revisiones) apropiadas algunas veces. La racionalidad es un término vago ligado tanto a la agencia como a decir que un sujeto tiene “creencias”.<sup>5</sup> La defensa más conocida de esta teoría es el libro *Minimal Rationality* de Christopher Cherniak (1990).<sup>6</sup>

c) **Racionalidad global:** El agente a partir de un conjunto particular de creencias debe realizar inferencias correctas la mayoría de las veces (Eraña: 2002), si bien no siempre.

d) **Racionalidad acotada:** El agente debe realizar las mejores acciones a partir de las limitaciones a las que está sujeto, tales como tiempo, capacidades cognitivas, información disponible, etc. (Gigerenzer: 2008). Al igual que la racionalidad óptima, esta caracterización también es maximizadora, sólo que toma en cuenta recursos y constreñimientos. Aquí la

---

<sup>5</sup> Conforme un sistema cognitivo tenga más relaciones inferenciales entre sus representaciones, será más racional, tendrá más agencia y éstas serán creencias en un sentido más fuerte.

<sup>6</sup> Desde el punto de vista exegético, también podría argumentarse que este es el tipo de teoría de la racionalidad que tenía en mente Davidson; pero también la racionalidad global.

consigna es inferir lo más que se pueda con los recursos disponibles.

e) **Racionalidad óptima:** El conjunto de creencias del agente debe estar deductivamente cerrado y ser consistente. Estas exigencias óptimas, en conjunto con la teoría de la decisión (basada en la probabilidad y la utilidad esperada) también es llamada visión estándar de la racionalidad.

Estas posturas son tesis acerca de la cantidad de razonamientos que deben ser correctos para que la racionalidad pueda atribuirse. El objetivo de este trabajo es realizar una indagación más general: se trata de encontrar los fundamentos de un razonamiento correcto. Empero, este trabajo no es compatible con todas las posturas acerca del primer problema.

Mostraré a continuación que la racionalidad mínima es la mejor manera de entender la relación entre racionalidad, creencias e inferencias correctas. Posteriormente señalaré algunas consecuencias de escoger la racionalidad mínima como base para trabajar acerca de las inferencias correctas.

Como señala Cherniak (1990), la versión de requisitos vacíos es implausible por motivos ontológicos (no captura lo que es de hecho agente cognitivo) y epistemológicos (no explica por qué una creencia es normalmente informativa acerca de otras creencias o de futuras acciones de un agente). Aún entre los teóricos que desean caracterizar a las creencias como un tipo de estado, existe la demanda de que haya una determinada estructura relacionando dichos estados (Armstrong: 1973).<sup>7</sup> La teoría de requisitos vacíos, al omitir requisitos inferenciales (o estructurales, en teorías como las de Armstrong) parece eliminar por completo la noción de lo que es una creencia, incluso como recurso heurístico para interpretar a un agente cognitivo.

Esto se puede observar con los criterios que regularmente usamos para atribuir creencias. Si alguien afirmara “creo que  $p$ ” pero también afirmara tener (demasiadas) creencias incompatibles con  $p$ , y (conociendo sus deseos<sup>8</sup>) viéramos que sus acciones son

---

<sup>7</sup> Estas relaciones estructurales que pide Armstrong son el análogo a las relaciones inferenciales que aparecen en una teorías más abstractas (como la de Cherniak), las cuales no descienden hasta detalles de tipo ontológico.

<sup>8</sup> Los deseos y las creencias están intrínsecamente relacionados. Stalnaker (1987), por ejemplo,

incompatibles o incluso contrarias a esta creencia, lo más adecuado sería no atribuir la creencia de  $p$  a esta persona. Sería mejor concluir que el sujeto dice que cree que  $p$ , pero no es verdad; o quizá si es demasiado incoherente, que no tiene capacidad de tener creencias.

La versión de requisitos vacíos también es inadecuada. Es posible que algún defensor hipotético de esta postura hiciera notar que existen ciertas relaciones a priori entre las aserciones y las creencias, e intentar derivar a partir de ello que la posibilidad de proferir una aserción de un sujeto es condición necesaria y suficiente para que ese sujeto tenga una creencia. Se puede mostrar esta relación haciendo un ejemplo análogo a la paradoja de Moore (1993). En esta paradoja, debemos considerar un sujeto que dice el siguiente enunciado:

**M1:** “Es verdad que ' $p$ ', pero no lo creo” (donde ' $p$ ' es cualquier proposición).<sup>9</sup>

Se trata de un enunciado paradójico que nos lleva a apreciar la relación intrínseca que existe entre creer una proposición y pensar que es verdadera. Estos son dos lados de lo que Engel (2001) denomina “el triángulo verdad-aserción-creencia.” Un análogo al enunciado mooreano, podría ser el siguiente:

**M2:** “' $p$ ', pero no lo creo”.

O de manera más precisa para evitar problemas:<sup>10</sup>

**M3:** “[afirmación de ' $p$ '], pero no lo creo”.

La inconsistencia patente de oraciones como M3 podrían dar una evidencia a favor de una

---

muestra cómo podrían definirse mutuamente.

<sup>9</sup> Alguien podría considerar que existen limitaciones psicológicas para lo que es posible creer, y que por tanto se pueden afirmar proposiciones o predicar la verdad de ciertas proposiciones sin que el sujeto que lo haga sea capaz de creerlas. Para quien desee utilizar esta manera psicológica de hablar de creencias, puede sustituir el término de “creencia” por “aceptación” de una proposición (este último término aparece en Cohen: 1989). En el contexto del presente trabajo, las creencias son consideradas desde un punto de vista epistémico, en el cual no difieren de la noción de aceptación.

<sup>10</sup> Para el enunciado M2, es posible pensar algunos contextos en donde no resulta paradójico. Por ejemplo, alguien podría ver un espectáculo de ilusionismo y afirmar “El mago cortó a la mujer en dos, pero no lo creo”. En este caso, el sujeto está citando un enunciado que ordinariamente serviría para afirmar la proposición ' $p$ ' (“El mago cortó a la mujer en dos” afirma que el mago, en efecto, cortó a la mujer en dos), pero en ese contexto el enunciado afirma otra proposición distinta, cuyo contenido sería algo como “Se observa que el mago cortó a la mujer,” de manera que no existe ninguna paradoja. Es fácil agregar una precisión para evitar este tipo de contraejemplos triviales.



caracterización de lo que es una creencia, y un agente racional, a partir de requisitos vacíos. Un agente racional es un agente capaz de realizar aserciones, y la posibilidad de llevarlas a cabo estipula las condiciones necesarias y suficientes para tener las creencias correspondientes. Concediendo que el teórico de requisitos vacíos lograra mostrar el modo preciso como esto se deriva de la naturaleza paradójica de enunciados como M3, ¿no sería esto una defensa adecuada de su postura?

Hasta aquí mi objetivo ha sido tratar de reconstruir una teoría en la que la aserción, o la disposición para una aserción, fuera suficiente para explicar qué constituye una creencia sin tener que recurrir a relaciones inferenciales con otras creencias. Me parece que esta estrategia no funcionaría por las siguientes razones.

En primer lugar, una explicación de este tipo presupone que hay una asimetría entre las aserciones y las creencias, según la cual es necesario entender las segundas con base en las primeras. ¿Por qué la capacidad de realizar aserciones tendría que ser un elemento constitutivo de la creencia? Al menos resulta perfectamente concebible que hubiera animales con creencias y que no fueran capaces de realizar aserciones.<sup>11</sup> Pero concedamos que la capacidad de realizar aserciones es condición necesaria para tener creencias. Independientemente de esto, es casi seguro que tener creencias también sea una condición necesaria para realizar aserciones (a menos que consideráramos como aserciones cosas que ordinariamente no lo son, como la emisión de signos de ciertos seres vivos que no poseen intencionalidad). Así, ambas cosas serían condiciones necesarias y suficientes (si dos cosas son mutuamente condiciones necesarias, son también mutuamente condiciones suficientes). ¿Por qué la asimetría en la explicación tendría que privilegiar a la aserción? Si bien no hay espacio para argumentar rigurosamente en favor de esto aquí, parece difícil negar que las creencias jueguen un papel causal en la aserción; y por ello que las aserciones dependan explicativamente de las creencias.

---

<sup>11</sup> A menos que se cayera en un absurdo de insistir que dichos animales, en condiciones ideales podrían, donde esas condiciones ideales incluyeran tener capacidades cognitivas para realizar actos de habla que de hecho no tienen.

La segunda razón por la cual la alternativa mooreana ofrecida aquí no es útil para proveer una teoría de requisitos vacíos, es que entendiendo a la aserción en su sentido más propio, surgen requisitos inferenciales, de modo que dejaría de ser una teoría de requisitos vacíos. La aserción que aparece en enunciados mooreanos no puede ser simplemente la pronunciación de un enunciado. Una aserción es un tipo especial de pronunciamiento, y no sólo una acción prosódica. Si aserción significara simplemente emitir los sonidos de un enunciado que expresa una proposición, las oraciones mooreanas: "El gato está en el tejado, pero no es verdad" o "El gato está en el tejado pero no lo creo" no resultan en absoluto paradójicas (por ejemplo, si un perico las pronunciara). Este sentido trivial de afirmación es el que tiene en mente Cherniak (1990) cuando rechaza a las teorías de requisitos vacíos como llanamente absurdas. Si aceptamos que una aserción, como acto de habla, es algo más complejo que va más allá de la parte prosódica y que incluye algunos elementos cognitivos, se evitan las objeciones de Cherniak, pues existe un límite impuesto por el aparato cognitivo. Los seres humanos, al menos, no podemos afirmar sinceramente cualquier cosa. Con frecuencia —quizá incluso de forma cotidiana— caemos en contradicciones sin darnos cuenta, mas no somos capaces de afirmar cualquier cosa. Pensemos enunciados con incompatibilidad material, como "el cielo está completamente azul y no está completamente azul en este momento"; algunas personas confundidas —como niños pequeños— podrían ser capaces de afirmar esto sinceramente, pero no cualquier adulto podría. Es decir, parte de los constreñimientos de nuestro aparato cognitivo se traducen en constreñimientos de tipo inferencial.<sup>12</sup>

La tercera razón es parecida a la segunda, pero únicamente acerca de los enunciados del tipo M1. Teníamos el problema de establecer cuáles deben ser las relaciones entre creencias; ahora tenemos un problema idéntico respecto a las aserciones. Las oraciones mooreanas del tipo M1 señalan que resulta paradójico no creer que es verdad una proposición que se afirma. Otra manera de ver la misma paradoja, es que resulta extraño para un sujeto afirmar una proposición que cree que no es verdadera. A su vez, es plausible

---

<sup>12</sup> Esto no implica de suyo que nuestro pensamiento intencional se lleve a cabo con manipulaciones simbólicas como pasa en las teorías computacionalistas de la mente.

que existan restricciones respecto a lo que un agente racional puede considerar como verdadero, y por tanto, sea capaz de creer.

En resumen, una teoría de la racionalidad no puede carecer de requisitos inferenciales. Aún la vía mooreana que parecía una respuesta plausible para asociar las creencias (y tal vez indirectamente a la agencia), hace reaparecer este tipo de requisitos. Esto no significa que la aserción no sea una manera válida de estudiar la racionalidad (por ejemplo la tradición griceana en filosofía del lenguaje la estudia de forma indirecta), pero por las razones de prioridad explicativa aducidas antes, parece mejor avocarse directamente a las relaciones inferencias y a las transiciones entre creencias.

Del resto de las alternativas, racionalidad óptima, acotada, global y mínima; es necesario señalar algo sobre las primeras dos. Voy a tomar algunos argumentos de Cherniak para mostrar que la racionalidad óptima y la racionalidad acotada no son alternativas adecuadas. Ambas tienen en común que operan con requisitos maximizadores, sólo que la segunda agrega constreñimientos a la maximización. Esto tendrá consecuencias relevantes, porque la racionalidad óptima y la acotada ofrecen criterios muy específicos acerca de qué constituye una inferencia correcta. Si éstas fueran alternativas viables, bastaría escoger bajo qué criterio (e.g. cantidad de creencias deducidas con lógica de primer orden) se está maximizando para resolver el problema de qué constituye una inferencia racional. Mostraré que no pueden ser una alternativa viable debido a que, por una parte, la maximización no captura del todo la noción de racionalidad. Aún quitando este problema, no queda claro cuáles deban ser los criterios de maximización, por lo cual incluso si las teorías maximizadoras de la racionalidad fueran adecuadas una investigación como la que llevo a cabo en este trabajo sería necesaria.

Los requisitos máximos están relacionados con lo que frecuentemente se denomina la visión tradicional o estándar de la racionalidad (Cfr. Stein: 1998). De lo que se trata es de escoger un paradigma de razonamiento que puede ser una teoría formal o un conjunto de teorías formales, y posteriormente señalar que un agente o inferencia es racional en la medida como se ajuste a estos paradigmas. Los más usuales son la lógica clásica, ya sea de primer orden o proposicional, y el cálculo de probabilidad, donde se asocia a la racionalidad con ciertas

propiedades metalógicas clásicas (como la correctitud), y con la clausura deductiva.<sup>13</sup>

Otra característica importante de este modelo es que no deben existir contradicciones dentro de las creencias del agente. Esto tiene consecuencias particularmente dramáticas si se utiliza la lógica clásica, debido a que es posible derivar cualquier cosa de una contradicción.

Los agentes naturales y los artificiales reales tienen recursos limitados de procesamiento de tiempo y espacio. La visión estándar de la racionalidad no incluye estos elementos. Esto no es una razón de suyo para argüir que se trata de un modelo inapropiado. Al menos podrían ofrecerse dos razones como respuesta a esta omisión. Una de ellas, es que un modelo debe incluir los aspectos adecuados para el nivel de generalidad que se busca; todas las ciencias hasta cierto punto realizan idealizaciones en el modo como representan los objetos que estudian. Los constreñimientos anteriores, de acuerdo con un defensor de la visión tradicional, no son relevantes en principio. En la práctica pueden ser importantes, para predecir lo que harán ciertos individuos en determinadas situaciones, o para elaborar agentes artificiales, así como para un ingeniero puede ser importante tomar en cuenta la fricción que no aparece en ciertas ecuaciones del físico.

Una segunda respuesta para este problema consiste en señalar que estos factores afectan la verosimilitud del modelo sólo en la parte descriptiva. Si los humanos, o la mayoría de los animales no-humanos, o incluso los agentes artificiales hasta ahora disponibles, se desvían de estos paradigmas formales de razonamiento, esto simplemente indica su falta de capacidad para razonar mejor. En la medida en que sus tipos de razonamiento se acerquen más a los teoremas de lógica clásica de primer orden o de cálculo de probabilidades, según esta visión serán más racionales.

Colyvan (2009) considera que las caracterizaciones tradicionales de la racionalidad son

---

<sup>13</sup> Desde un punto de vista histórico, la visión tradicional de la racionalidad se consolida con el nacimiento de la teoría de decisión y la interpretación clásica de la probabilidad en el siglo XVIII, sin embargo en la filosofía se pueden encontrar ejemplos de ella todavía en la segunda mitad del siglo XX. Esta visión es muy amplia, pues está basada en la probabilidad y la lógica deductiva, y de esos modelos se derivan otros: teoría de juegos, teoría de la utilidad, condicionalización bayesiana, la coherencia de apuestas, etc.

adecuadas porque una teoría de la racionalidad establece qué es un buen razonador, lo cual consiste en conseguir ciertos fines. Colyvan considera, como proposición empírica (y por tanto derrotable), que la lógica y la probabilidad son las mejores explicaciones acerca de cómo conseguir estos fines. Las razones ofrecidas por Cherniak para rechazar estos modelos tradicionales muestran que esta tesis de hecho es derrotada. Cabe señalar que aquí no se trata el tema de determinar si las reglas de la lógica clásica son correctas o incorrectas, sino de establecer si son adecuadas para realizar el trabajo de una teoría normativa del razonamiento.

Los ejemplos que pueden encontrarse en Cherniak (1990) muestran el grado de "parálisis" al que puede llegar un agente que intente seguir estos principios. Los problemas de complejidad computacional son tan comunes, y los resultados de adoptar un paradigma de razonamiento semejante son tan pobres, que resulta difícil mantener esta opción como viable. Un inconveniente peor es que algunos problemas son indecidibles; donde la solución seguiría buscándose al infinito para un problema que es insoluble. Encontrar una solución general de antemano para saber si un problema cualquiera es soluble, equivale a resolver el problema de la parada,<sup>14</sup> lo que no es soluble.

Algunos de estos problemas se pueden resolver cambiando de sistema lógico. Pero la elección de una nueva estructura formal para reemplazar a las teorías normativas tradicionales no sólo debe ser hecha pensando en resolver estos problemas, sino conociendo los principios epistémicos adecuados, lo cual es el tema de este trabajo.

Un modo inmediato de responder a los problemas de la visión tradicional, es señalar que se trata únicamente de un ideal normativo. Si bien un agente racional no puede cumplir con los requisitos infinitos de la omnisciencia lógica, es posible sostener que es racional en tanto se acerque a ellos. Cherniak (1990) tiene dos argumentos importantes para cuestionar la verosimilitud de que la omnisciencia lógica pueda ser un ideal normativo. El primero es que por motivos de complejidad computacional, las inferencias apropiadas para obtener una respuesta correcta podrían tardar demasiado tiempo (quizá más del tiempo de vida del

---

<sup>14</sup> El problema de la parada consiste en encontrar un procedimiento general para saber si cualquier programa realizado por una máquina de Turing llegará a detenerse eventualmente. Cfr. Kozen: 1997.

universo).

Otro motivo por el cual el modelo deductivo parece no ser correcto, es su énfasis en la validez lógica. Los argumentos de Cherniak (1990) para mostrar que ciertas reglas inválidas no afectarían la racionalidad de un agente son muy convincentes. Por ejemplo, pensemos en la siguiente regla: “Hacer conjunción de todas las proposiciones atómicas relevantes que se saben verdaderas y se están considerando en el momento actual. Excepto cuando son más de mil millones, en ese caso concluir lo que sea”. ¿Esta incorrección de una regla de inferencia debe hacer que juzguemos distinto la competencia de un agente racional? Aquí la respuesta depende de la perspectiva teórica que adoptemos. Desde una perspectiva como la de la filosofía de la lógica que se pregunta acerca de la inferencia racional y sobre cuál es el sistema que la captura mejor (lo que se denomina a veces estándar deductivo), entonces una invalidez como la descrita en el ejemplo de Cherniak sí podría ser relevante; pues el estándar de la deducción debe servir incluso en situaciones hipotéticas o contrafácticas. La situación contrafáctica relevante al ejemplo de Cherniak es aquella donde el agente realmente tuviera conocimiento de mil millones de proposiciones atómicas.

En cambio desde una perspectiva que tenga que ver con la racionalidad de un agente —es decir, no con buscar el mejor sistema formal— las propiedades metateóricas del sistema son triviales si no producen ninguna consecuencia práctica. El siguiente ejemplo basado en el anterior puede ser ilustrativo. Tomemos las reglas  $A$  y  $A'$ , para cualesquiera fórmulas  $\alpha$ ,  $\beta$ ,  $\gamma$ :

**Regla A:**

$\alpha$ ,  $\beta$

$\alpha \wedge \beta$

**Regla A':**

$\alpha$ ,  $\beta$

Si la longitud respectiva de  $\alpha$  y  $\beta$  es menor a  $1 \times 10^{99}$  símbolos entonces  $\alpha \wedge \beta$ , si la longitud respectiva de  $\alpha$  o  $\beta$  es mayor a  $1 \times 10^{99}$  símbolos entonces  $\gamma$ , para cualquier gama.

### Escenario:

El sujeto  $S$ , genera la creencia de los gatos son negros ( $\forall x (Gx \rightarrow Nx)$ ). Después, de manera independiente, genera la creencia de que el cielo es azul ( $Ac$ ).

### Resultado

a) Cuando  $S$  sigue  $A$ :  $\forall x (Gx \rightarrow Nx) \wedge Ac$

b) Cuando  $S$  sigue  $A'$ :  $\forall x (Gx \rightarrow Nx) \wedge Ac$

Queda claro a partir del ejemplo, que cuando  $S$  sigue la regla  $A$  obtiene los mismos resultados que cuando sigue  $A'$ . Desde un punto de vista meta-teórico,  $A'$  no es equiparable a  $A$ , puesto que  $A'$  es una regla incorrecta (prueba como teoremas fórmulas que no son consecuencias lógicas). Sin embargo, desde el punto de vista de las acciones de  $S$ , ambas reglas producen resultados indiscernibles.

Por las razones aducidas anteriormente, parece que la versión de requisitos máximos no es una caracterización adecuada de la racionalidad. La versión acotada es un intento de resolver algunos de los problemas planteados anteriormente, pero manteniendo una perspectiva maximizadora.

Mientras que la racionalidad de requisitos máximos exige omnisciencia lógica, o maximización de utilidad (en un contexto de teoría de la decisión) de manera absoluta, la racionalidad acotada sólo requiere que se cumplan estos máximos de acuerdo con los recursos del agente. En un contexto lógico, la racionalidad acotada indicaría evitar tanto la parálisis como pérdida de recursos, al exigir que se procese información sólo cuando los beneficios de hacerlo exceden los costos.<sup>15</sup> Éstas incluyen fundamentalmente información disponible, tiempo y recursos cognitivos. Este enfoque posee también exigencias excesivas,

---

<sup>15</sup> Cabe notar que la racionalidad acotada necesariamente incluye ya sea metas, deseos, fines o en general elementos relacionados con la racionalidad práctica.

como indica Gigerenzer (2008, p. 6): "El ideal puede socavar los intentos de hacer una teoría más realista al demandar nuevos mecanismos de optimización, tales como que sea necesario conocimiento sobre el costo y beneficio para llegar a un punto de detención de la búsqueda."

Una vez establecidos los problemas de la teoría que pide requisitos vacíos (a), y de las dos que tienen la característica de ser maximizadoras (d, e), hace falta examinar la racionalidad mínima y a la racionalidad global. La racionalidad global indica que "para poder atribuirle creencias a un sujeto, es necesario que éste lleve a cabo y seleccione a partir de su conjunto de creencias, un mayor número de inferencias correctas que incorrectas." (Eraña: 2002). En cambio la racionalidad mínima sólo exige que existan algunas inferencias apropiadas.<sup>16</sup> ¿Cuántas son suficientes? Cherniak (1990) propone que la racionalidad es un término vago. Entre la racionalidad y la carencia de racionalidad hay un continuo según la cantidad de inferencias correctas realizadas.

Si vinculamos a las creencias con las relaciones lógicas entre las mismas, entonces también habría que decir que las creencias se atribuyen vagamente (un agente demasiado inconsistente, se le atribuye creencia en menor grado, sólo un sujeto por completo inconsistente dejaría de tener creencias y por tanto de ser un agente cognitivo).

Parece plausible conceder que una vez que se adopta una perspectiva de la racionalidad basada en relaciones lógicas entre creencias, la racionalidad debe ser un término vago. Desde el punto de vista epistémico, no resulta sorprendente que ciertos términos sean vagos. Un sujeto podría estar justificado en creer que no se puede construir una máquina de movimiento perpetuo si su fuente es un artículo enciclopédico de una buena publicación, pero es razonable pensar que el físico que escribió dicho artículo sostenga la misma creencia con mayor grado de justificación. Plantinga (1993), por ejemplo, considera que existe cierta vaguedad en el conocimiento, la función propia, la garantía epistémica, y otros términos emparentados.

Sin embargo la idea de Cherniak es más radical. De acuerdo con ésta, la distinción

---

<sup>16</sup>Al igual que en el caso de la racionalidad acotada, aquí van implícitos elementos prácticos.



misma entre un agente epistémico y algo que no lo es, es vaga. Podemos pensar en un animal que tiene ciertos estados representacionales. Si la coherencia entre los mismos se fuera degradando y las conexiones lógicas entre ellos se fueran perdiendo, de acuerdo con la tesis de vaguedad de Cherniak, estos estados serían creencias y el animal un agente epistémico cada vez en menor grado, descendiendo gradualmente hasta llegar al punto en que no hay ninguna relación lógica entre los mismos.

La tesis de Cherniak tiene como resultado una idea excesivamente liberal sobre la agencia epistémica. Un defecto que inmediatamente salta a la vista, es que no tiene conexión alguna con la idea de agencia *simpliciter*, es decir, llevar a cabo una acción. Me parece que Cherniak se concentra en condiciones necesarias, pero de ningún modo provee condiciones suficientes, ni tampoco algo cercano a un esbozo de una teoría sobre la agencia epistémica. Quizá la teoría de Cherniak puede ser una especie de teoría eliminativista, que señalara que no hay una propiedad que tengan ciertos organismos tal que los haga ser agentes epistémicos, sino que lo que llamamos agencia epistémica es tan sólo un organismo cuyos estados intencionales tienen relaciones inferenciales. Sería raro pensar que ésta es la intención detrás de la teoría de racionalidad mínima, al menos no hay un plan eliminativista explícito en *Minimal Rationality*. Sin embargo supongamos que alguien quisiera defender la vaguedad acerca de la agencia epistémica en un espíritu similar a éste, en tal caso tendría que explicar cómo las relaciones inferenciales son una condición suficiente para tener creencias, conocer, revisar creencias, y otras actividades propias de un agente epistémico.

La agencia epistémica es el caso más problemático, pero existe un problema general con todos los términos que Cherniak quiere tomar como vagos. En líneas generales, el problema es que incluso si tomamos los requisitos lógicos únicamente como condición necesaria y que éstos se den de manera vaga, aún así resulta controvertido que la racionalidad o la atribución de creencias sean vagas. Por ejemplo, partiendo de una teoría que intente enmarcar este tipo de características en relación a las funciones biológicas es posible notar, como señala Burge (2009, p. 269), que "dondequiera que hay funciones, [se puede considerar] *a priori* que hay estándares para cumplirlas". Una teoría de la agencia (psicológica, epistémica, cognitiva, etc.) de este tipo determinará sin vaguedad los casos en que hay estos tipos de agencia, aún

si ciertos elementos como los aquí descritos son vagos.

En resumen, Cherniak no provee una teoría completa de la agencia epistémica, sino en todo caso de las condiciones necesarias. La idea de que los términos referentes a la racionalidad, agencia y creencias sean vagos, es algo que dependerá de una teoría más completa acerca de la agencia epistémica.

Hago un señalamiento más. Arriba mostré que las versiones más plausibles de una teoría inferencial del razonamiento eran la racionalidad mínima y la racionalidad global. Esta última, pone como requisito un alto número de relaciones inferenciales para que un agente pueda ser racional. Esta teoría no tiene la implicación de que la racionalidad deba ser un término vago. Lo anterior, sin embargo, no elimina la idea de que la racionalidad pueda tenerse por grados (una vez que se es racional).

Para fines de este trabajo no es necesario decidir cuál de estos dos enfoques es mejor, puesto que la investigación se centrará en la pregunta sobre cómo se puede considerar un razonamiento como apropiado, y no en la proporción o cantidad de razonamientos correctos que sea posible realizar.

En resumen, la caracterización de una creencia y de un agente capaz de tener creencias incluye como mínimo algunos tipos de relaciones apropiadas entre éstas, de tipo inferencial en un sentido amplio que aún hace falta aclarar. Una teoría de requisitos vacíos es implausible, y las dos teorías maximizadoras (racionalidad acotada, racionalidad optimizadora) son tan ajenas a la práctica epistémica real (en los casos como los que plantea Cherniak) que parece dudoso que puedan ser un modelo de la racionalidad epistémica, incluso normativo. Los modelos restantes, son tanto la racionalidad mínima como la racionalidad global. Ambos incluyen una noción de inferencia apropiada o adecuada, pero que corresponde enteramente a las estructuras formales de los modelos optimizadores.

## **El razonamiento en un marco de reglas**

En el apartado anterior expliqué a grandes rasgos en qué consiste la visión inferencial de la

racionalidad y las creencias. La teoría fue planteada de dos maneras equivalentes sin hacer distinción entre ellas: como las relaciones existentes dentro del conjunto de creencias de un agente, y como las transiciones que se dan de unas creencias hacia otras nuevas. Este segundo enfoque es el que adoptaré en el resto del trabajo.

En esta investigación acerca de la racionalidad asumiré un marco basado en reglas. La motivación y las razones para llevar esto a cabo son parecidas a las que ofrece Goldman en *Epistemology and cognition* (1988) para estudiar así a la justificación. Explicaré brevemente por qué resulta una manera adecuada de investigar la racionalidad a pesar de que existen inconvenientes en apariencia.

Goldman hace notar que los términos referentes a la justificación contienen un componente deóntico que queda capturado en un lenguaje de reglas. En segundo lugar, centrar la discusión en reglas, es un modo de organizar y estructurar el debate acerca de la justificación de manera neutral, antes de defender una teoría propiamente acerca de la justificación (Goldman: 1986, p. 59). En este caso, en vez de justificación sería acerca de la racionalidad.

Resulta verosímil que si algo tiene un componente normativo, la obligación que representa pueda expresarse a través de una o varias reglas. Por ejemplo, si  $S$  debe realizar un razonamiento  $r$  en un momento  $t$ , es posible formular una regla o un sistema de reglas de la cual se derive este imperativo. Un razonamiento es un proceso, que en el caso de agentes, puede ser capturado como producto de una competencia análoga a las competencias lingüísticas (Stein: 1998) y esto a su vez se puede expresar bajo la forma de una regla. Ésta puede ser entendida como una formulación abstracta de una clase de procesos.

Un caso en el que no sería fructífero adoptar un marco de reglas es si éstas tuvieran una forma demasiado general e imprecisa, tales como “Un razonamiento debe intentar llevar a conclusiones relevantes,” “Un razonamiento debe estar dirigido hacia cumplir los deberes epistémicos del agente,” o “Un razonamiento debe buscar la coherencia”. Suponiendo que tales soluciones fueran aceptables, este tipo de posturas sería menos compatible con este proyecto.

Cabe reconocer entonces que el marco de reglas quizá no es capaz de acomodar ni

sistematizar, al menos de modo sencillo, cualquier concepción acerca de la racionalidad y el razonamiento, o del conocimiento y la justificación. Dicho de otro modo, hay buenas razones para asumir un marco de reglas al comenzar una investigación, lo cual genera preguntas como ¿qué es una regla correcta?, ¿cuáles reglas son correctas? Empero, quizá sea necesario abandonar dicho marco dependiendo de los resultados finales de la investigación.

Una vez aclarado que este es el marco de la investigación, hace falta una definición de de razonamiento. Al hablar de razonamiento (y secundariamente, de revisión de creencias), lo hago desde una perspectiva epistémica. Debe existir un nivel de explicación desde la psicología (o quizá otras ciencias cognitivas) mediante el cual se realicen estos procesos. Con ello no quiero decir que el razonamiento (o un tipo razonamiento determinado, como el *modus ponens*) sea en sí un proceso psicológico.<sup>17</sup> Es posible que lo que llamamos razonamiento no corresponda a un sólo proceso cognitivo, sino que sea realizado por un gran número de competencias distintas, sin descartar que exista un componente social o cultural en el mismo. Desde el punto de vista epistémico estos detalles resultan ser más bien secundarios.

El siguiente ejemplo ilustra el nivel de generalidad que busca la epistemología. Por ejemplo, confundir la ausencia de evidencia con la evidencia de ausencia probablemente es una falla desde un punto de vista epistémico (cabe mencionar llega a ser causa de negligencia médica). Éste sería entonces un modo inadecuado de razonar para cualquier agente, sin importar cómo funcione su aparato cognitivo ni qué mecanismos psicológicos concretos haya utilizado para llevar a cabo esta inferencia.

Una vez aclarado que el tipo de explicación que se busca aquí no es psicológica, hace falta dar una definición provisional de razonamiento:

*Proceso mediante el cual un agente llega de un conjunto<sup>18</sup> de estados doxáticos a otros sin*

---

<sup>17</sup> Hablo sobre psicología porque me parece el nivel de descripción más adecuado para indicar lo que son las creencias. Sin embargo, cabe notar que podría haber desacuerdo, ya sea con respecto a que exista una explicación psicológica autónoma (e.g. sin necesidad de apelar a aspectos neurológicos), o de que la psicología sea el nivel adecuado de explicación ya sea para los fenómenos intencionales, o concretamente para las creencias.

<sup>18</sup> Presumiblemente el conjunto inicial es no-vacío.

*necesidad de recibir nuevos insumos.*

Esto debe ser tomado como una caracterización inicial, de modo que no tiene un carácter definitivo y mucho menos pretende ser un análisis del término “razonar.” Cabe notar también que ésta resulta neutral respecto de los detalles que deberá llenar una teoría del razonamiento, tales como si debe ser consciente/reflexivo (realizado a nivel personal), o si el medio ambiente forma parte del proceso de razonamiento —constituyendo un caso de cognición situada— o se trata de algo que tiene lugar exclusivamente dentro del aparato cognitivo del organismo que razona. En todos estos casos hay un cambio doxástico como el descrito, de modo que los detalles concretos acerca de cómo se realiza el razonamiento no deben afectar una investigación con un nivel de abstracción propio de la epistemología.

Intentos más precisos de dar una definición de razonamiento han sido llevados a cabo por Walton (1990) y Harman (1988, 1999), y más recientemente Sperber (2012), entre otros. Debido a que esta investigación no es propiamente sobre el razonamiento sino sobre la relación entre la racionalidad y las reglas de razonamiento, considero conveniente tener una definición más general y menos comprometida (si bien, necesariamente menos completa).

A continuación caracterizaré las reglas de razonamiento. Siguiendo la definición provisional de arriba, se trata de las reglas que siguen los sujetos para llegar de unas creencias a otras. Para hablar de razonamiento es común apelar a las reglas de deducción natural que usan algunos sistemas formales (e.g. Cherniak: 1990), y que precisamente reciben este nombre debido a que —supuestamente— son intuitivas y fácilmente aplicables para sujetos que no poseen conocimientos técnicos de lógica. La regla de conjunción, aplicada al razonamiento, indica que si creemos  $P$  y también creemos  $Q$ , estamos autorizados a creer “ $P$  y  $Q$ .”

Es casi seguro que una caracterización con reglas puramente formales tenga limitantes, que haya casos de razonamiento genuino que no puedan ser capturados de este modo. Por ejemplo, de la creencia “El círculo 1 es completamente rojo,” es fácil para un humano obtener la creencia “El círculo 1 no es completamente azul.” ¿Bajo qué regla se puede describir este

razonamiento? Una manera podría ser bajo una regla usando lógica de segundo orden:<sup>19</sup>

*Para todo  $x$ , para todo  $P$  y para todo  $Q$ , si  $P$  y  $Q$  son predicados que hablan de ser completamente de un color, y  $P$  no es igual a  $Q$ , y  $x$  es  $P$   
 $x$  no es  $Q$*

Esta regla requiere cuantificar sobre predicados y no sólo sobre individuos. Si bien no es enormemente complicada, es una inferencia que tiene un cierto grado de sofisticación innecesaria. Es poco probable que cuando concluimos que una figura no es azul lo hagamos apelando a un axioma formulado en lógica de segundo orden sobre los predicados de color. Pero notemos que desde un punto de vista semántico ser completamente rojo excluye ser completamente azul. Otro modo de ver la inferencia de que el círculo sea rojo a que no sea azul, es pensar que esta incompatibilidad semántica es lo que se usa para realizar una inferencia, sin necesidad reglas.

Otro caso de razonamiento que no se ajusta fácilmente a un modelo inferencial, como hace notar Harman (1988) es el razonamiento siguiente: “observo que hay una silla en la habitación, por tanto hay una silla en la habitación”. Considerar que esto es un razonamiento presupone que al menos algún tipo de estado perceptual puede ser considerado un estado doxástico (pues la transición se hizo a partir de una percepción a una creencia. Quisiera dejar de lado esa cuestión para concentrarme en el punto que quiere señalar Harman. Lo que se busca ilustrar aquí es que no es necesario comprometerse a que todas las reglas de razonamiento tengan que ser inferencias formales. Nuevamente, es posible pensar en una regla formal para expresar esto. En lógica de primer orden podría decirse: “Para todo  $x$ , si  $x$  es una imagen perceptual y  $x$  es observado en  $z$  / existe un  $w$  tal que  $w$  es el objeto que causa  $x$ , y  $w$  se encuentra en  $z$ ”.

¿Cómo deben expresarse las reglas de razonamiento?

Lo que haré a continuación es buscar un modo de caracterizar las reglas de razonamiento. Primero mostraré dos maneras distintas de estudiarlas, que son el sentido descriptivo y el

---

<sup>19</sup> Además de esta regla, para poder aplicar esta inferencia se requieren reglas sobre la cuantificación, concretamente la instanciación universal.

sentido normativo. Posteriormente voy a evaluar dos maneras que son candidatas para capturarlas: como argumentos o como principios epistémicos. Mostraré que los principios epistémicos son un modo más general, y que es capaz de incluir a los argumentos como caso especial.

La distinción entre reglas en sentido normativo o descriptivo es de carácter metodológico. Por supuesto, existe un sentido trivial en el que cualquier regla es normativa. Toda regla indica lo que se debe hacer bajo ciertas circunstancias. La distinción anterior no se refiere a esto, sino a que existe un conjunto de reglas que puede utilizar un ser racional (por ejemplo, los humanos), y el conjunto de inferencias que son adecuadas en el sentido descrito por los apartados anteriores.

Al primer estudio se le puede llamar epistemología descriptiva (Stein: 1998). Al estudio de esto se adjudica a la psicología cognitiva, pero no es necesario descartar inicialmente otros enfoques.<sup>20</sup> El sentido normativo, está más relacionado con la epistemología tradicional, y consiste simplemente mostrar qué tipos de razonamiento son adecuados.

Estas dos cuestiones no son del todo independientes. En palabras de Harman (1988, p. 7): “Por una parte parece difícil llegar a principios normativos sin considerar cómo razonan las personas realmente, lo cual pertenece a una teoría descriptiva. Por otra parte, parece que cualquier teoría descriptiva conlleva una cierta cantidad de idealización, por lo que es normativa hasta cierto punto.” Esta investigación estará centrada en las reglas de razonamiento en sentido normativo.

¿Cómo se puede caracterizar una regla de razonamiento? Propongo examinar dos alternativas:

- a) Como argumento formal.
- b) Como principio epistémico.

---

<sup>20</sup> Las investigaciones comunes, tales como lo que se estudia en economía conductual, parten de dos presupuestos que han resultado fructíferos. El primero, puesto en duda por los estudios de cognición situada, es que la unidad de análisis relevante son los agentes. El segundo, cuestionado por la neuroeconomía, es la autonomía del nivel de explicación de la psicología. Considero que aún no hay razones de peso para abandonarlos.

La primera alternativa implica caracterizar a los razonamientos como inferencias, por ejemplo, la eliminación de la conjunción: “si un sujeto  $S$  tiene la creencia  $A$  y  $B$ , entonces debe creer  $A$ .”

Existen diversas razones que podrían inclinarnos a escoger la primera alternativa. En primer lugar, uno de los intereses que existen al buscar un modo abstracto de caracterizar que es un razonamiento es que la caracterización no prejuzgue la investigación. En este sentido, la noción de argumento es bastante flexible, los argumentos pueden ser derrotables, probabilísticos, o heurísticas sencillas. Otra ventaja de considerar al razonamiento de una manera argumentativa es que no se presuponen otras nociones que podrían complicar el estudio del razonamiento, particularmente nociones de tipo epistémico, tales como la justificación. Una tercera ventaja es que los principios de inferencia parecen ser muy fructíferos para estudiar empíricamente el razonamiento, por la relativa facilidad de verificar si los sujetos los cumplen.

Sin embargo, existen algunos problemas con esta caracterización. El primero es que no todos los razonamientos pueden representarse de manera formal, ya que existen razonamientos heurísticos que se basan en tipos semánticos (García: 1999). Otro problema es que hacer demasiado énfasis en el razonamiento como procesamiento podría dificultar distinguirlo de otro tipo de regularidades que normalmente no identificaríamos con el razonamiento. Brandom (2000) desarrolla este tema con respecto a la justificación utilizando el ejemplo de un termómetro, que muestra con regularidad la temperatura, a pesar de lo cual no pensamos que merezca ser llamado racional. Un tercer problema, en particular usando como modelo lógicas monótonas, es que una visión es que el modelo tendría como resultado una acumulación ciega de consecuencias, lo cual que presentaría los problemas mostrados por Cherniak. Un ejemplo sencillo es el siguiente. Supongamos por un momento que la regla de adición, también llamada introducción de la disyunción, es una regla correcta para razonar. Al usar esta regla, de la creencia  $P$ , podrían seguirse las creencias  $P \vee Q$ ,  $P \vee R$ ,  $P \vee S$ , y así para cualquier proposición atómica (también para proposiciones no-atómicas, pero esto es suficiente para el ejemplo). Dado el número tan grande de proposiciones que un sujeto puede representarse, a partir de tener la creencia  $P$  podría permanecer de manera



indefinida agregando disyunciones. Un agente paralizado, realizando tareas cognitivas inservibles, es una imagen que dista mucho de lo que normalmente se entiende por racionalidad.

Otro ejemplo es el siguiente (Harman: 1988): De la creencia en la implicación de  $P \rightarrow Q$ , y de  $P$ , un agente regularmente puede escoger entre aceptar  $Q$  o negar alguna de las dos premisas; pero un modelo puramente argumentativo parecería que lo obliga a aceptar  $Q$ . En muchas ocasiones, cuando nos damos cuenta de que nuestras creencias tienen implicaciones no deseadas, revisamos el cuerpo de creencias en vez de aceptar estas implicaciones. Por ello un argumento deductivo difícilmente puede capturar qué es la racionalidad. Este segundo problema, podría resolverse potencialmente adoptando una lógica más sofisticada.

Como mostraré a continuación, adoptar una descripción de las reglas de razonamiento como principios epistémicos puede evitar los compromisos que parece traer la forma argumentativa, de modo que la solución no dependa de que los epistemólogos formales sean capaces de realizar modelos adecuados para capturar las reglas de razonamiento bajo la forma de argumentos. Por otra parte, ya que los principios epistémicos pueden tomar una forma inferencial, será posible mantener las ventajas de la descripción argumentativa sin tener que sostenerla.

Como mencioné anteriormente, los principios epistémicos pueden ser principios de generación o principio de transmisión (o ambos). Los principios de generación especifican circunstancias en las que una creencia se justifica independientemente de su relación con otras creencias, mientras que los principios de transmisión buscan establecer cuál es la justificación de una creencia dada otras (Boghossian: 2003). Los principios epistémicos que podrían servir aquí son los de transmisión, y en concreto solamente algunos que hicieran referencia a esta forma argumentativa.

García (1999) propone capturar las reglas de razonamiento *en sentido normativo* por medio de enunciados condicionales del siguiente tipo: “Si  $S$  cumple tales y tales condiciones, entonces está epistémicamente permitido que  $S$  infiera tal y tal creencia o estado

intencional.” Esto también puede adaptarse para hablar del razonamiento en sentido descriptivo: “Si  $S$  cumple tales y tales condiciones, entonces le parece correcto a  $S$  inferir tal y tal creencia o estado intencional”. Al elaborar la definición de razonamiento no con base a obligatoriedad, sino a permisos, se evitan problemas que surgen a partir de concebirlo como argumento, pero esto no excluye a las reglas de tipo formal.

En el mismo texto, García (1999) muestra cómo su definición puede cubrir una regla formal: “Si  $S$  tiene ciertas creencias cuyos contenidos tienen determinadas formas lógicas  $F_1... F_n$  [...] entonces está epistémicamente permitido a  $S$  inferir otra creencia con forma lógica  $F_n$ ”.<sup>21</sup> En adelante, este trabajo se concentrará en determinar si es posible establecer una normatividad para las reglas de razonamiento y decir de qué manera podría hacerse.

La racionalidad así concebida puede entenderse como parte de la epistemología operativa, tal como la define Williamson (2008):<sup>22</sup> como dar un conjunto de reglas en el que se puede saber en principio si se cumple con ellas entonces “la racionalidad puede concebirse como el requisito de conformidad con la epistemología operativa” (p. 2). Para dar entonces una respuesta completa acerca de si un agente es racional, sería necesario determinar qué cantidad o proporción de inferencias correctas debe realizar, lo cual no es parte de este trabajo.

---

<sup>21</sup> La definición completa incluye: “y si estas creencias son epistémicamente aceptables”, sin embargo esto me parece erróneo. Si se refiere a que son epistémicamente aceptables desde el punto de vista del agente, esto ya estaba capturado al decir que son sus creencias. Si se refiere a que son epistémicamente aceptables en sentido objetivo, esto parecería introducir un problema, ya que la fuente de las creencias a partir de las cuales se razona resulta irrelevante para el proceso de razonamiento, pero no así para la aceptabilidad epistémica.

<sup>22</sup> El artículo de Williamson es en contra de la posibilidad de la epistemología operativa. La epistemología no-operativa, por el contrario, toma en cuenta factores que no son transparentes al agente. Por ejemplo, la posibilidad de que sus creencias sean falsas o erróneas.

### 3. UNA ESTRATEGIA PARA VALIDAR REGLAS DE RAZONAMIENTO

#### Criterio de Corrección

Según lo establecido en el capítulo anterior, una teoría de la racionalidad epistémica indica el modo adecuado de revisar creencias o adquirir nuevas. Uno de los resultados una investigación de este tipo sería un conjunto de reglas adecuadas. La manera como Goldman (1988) sugiere plantear esto es a través de un criterio de corrección. La estrategia puede resumirse en dos pasos:

- 1- En primer lugar, es necesario resolver cuál es el criterio (o criterios) que debe cumplir una regla para ser adecuada.
- 2- Posteriormente, hace falta ver qué reglas o clases de reglas cumplen con este criterio.<sup>23</sup>

Este modo de proceder presupone que existe un criterio o una serie de criterios en común a todas las reglas de razonamiento que sean correctas. También se puede recurrir a un paso intermedio: establecer un meta-criterio, es decir, un criterio para escoger criterios. Considero que esta distinción no es interesante, pues la estrategia descrita no restringe cuáles deban ser los niveles de criterios. Lo importante es que en última instancia exista un principio que permita decidir si una regla es adecuada o inadecuada.

Los siguientes son ejemplos de algunas reglas de razonamiento. La motivación para mostrar la siguiente lista únicamente es proveer algunos ejemplos qué tipo de reglas serían candidatos para ser aceptables o inaceptables, es decir, para saber cuál es el nivel de generalidad de una regla como la que se tiene en mente aquí:

- a) R1: Si el sujeto cree en una proposición con forma de condicional y cree en el antecedente, entonces debe creer en el consecuente (*Modus ponens*).

---

<sup>23</sup> Un tercer paso podría ser comprobar si los humanos, o determinadas especies de animales razonan de este modo. Para una exposición del debate de racionalidad (humana) ver Stein (1998, cap. 1), y Stich (1993, cap. 1).

- b) R2: Realizar *modus ponens* si existen los elementos que señala R1, pero además sólo si resulta relevante dada la información actual.
- c) R3: No asignar probabilidades que violen los axiomas de Kolmogorov (incoherentes).
- d) R4: Asignar probabilidades tomando en cuenta la evidencia según la probabilidad condicional.
- e) R5: Confiar en la propia memoria *prima facie*, es decir, si no existen buenas razones para no hacerlo.
- f) R6: Confiar en el testimonio de otra persona *prima facie*, es decir, si no existen motivos para desconfiar.

La cantidad de reglas correctas puede ser muy grande o quizá infinita. Por ello un criterio de corrección puede tener importancia más allá de un paso intermedio para escoger reglas correctas, pues debido a que su cantidad es tan grande, unos pocos criterios podrían ser más informativos cuando lo que interesa es saber en qué consiste una inferencia racional.

Algunos ejemplos de candidatos para ser criterios correctos son los siguientes. Como en la lista de reglas que mostré anteriormente, la intención es únicamente ejemplificar cómo puede ser un criterio. Una vez más, la intención no es poner estos criterios a consideración, sino mostrar cuál es el nivel de generalidad que se busca:

- a) Que la regla sea formulable como un teorema de la lógica clásica o bien del cálculo de probabilidad (Stein: 1998).
- b) Que la regla produzca un número suficiente de creencias verdaderas (Goldman: 1979).
- c) Para que una regla sea racional debe estar en concordancia con la formulación de las normas comunitarias locales.
- d) Que la regla aparezca espontáneamente como válida a un sistema cognitivo producto de la selección natural. Una justificación de este criterio se puede ver en los argumentos evolucionista en favor de la racionalidad humana (e.g. Dennett: 1989, Feldman: 1988).

e) Que la regla cumpla determinados requisitos de segundo orden, como que aparezca como válida después de algún tipo de reflexión.

f) La regla debe tener “validez ecológica.” Esto significa que debe ser rápida, frugal y robusta<sup>24</sup> con poder de predicción (Gigerenzer: 2008).

Como se puede ver, es relativamente fácil aplicar criterios como éstos y por tanto, resolver el problema de si una regla de razonamiento es correcta o incorrecta. De este modo, se resolvería el problema más importante respecto a qué tipos de inferencias constituyen a un agente racional.

Cabe señalar que el proyecto de buscar un criterio no es necesariamente un análisis, como es común encontrar en la literatura; es decir una búsqueda de condiciones necesarias y suficientes. La prueba de Turing puede ilustrar un caso de un criterio que no está elaborado de manera analítica. Alan Turing (1950) sugirió que una computadora que pudiera ser confundida con un humano al conversar (es decir, que el 50% de las veces los interlocutores pensarán que es humano), debería ser considerada como inteligente. Presumiblemente, pasar la prueba de Turing no es condición suficiente para poseer inteligencia,<sup>25</sup> y tampoco es un análisis de los aspectos principales sobre lo que significa la inteligencia; ni siquiera nos dice algo interesante sobre el tema. Sin embargo, aún así puede defenderse como un criterio bueno para reconocerla. Dennett (1998) ilustra el sentido original de dicha prueba usando como ejemplo una prueba bajo su propio nombre.

La prueba Dennett para saber si uno vive en una gran ciudad es si en un día cualquiera uno puede realizar lo siguiente: escuchar una orquesta sinfónica, ver un cuadro de Rembrandt y un evento atlético profesional, almorzar *quenelle de brochet*. Es evidente que no hay ninguna necesidad conceptual que impida que un pequeño pueblo tenga una orquesta sinfónica, un

---

<sup>24</sup> La frugalidad se refiere a que utilice pocos recursos computacionales, o específicamente cognitivos. La robusteza se refiere a la capacidad de hacer predicciones para sucesos novedosos.

<sup>25</sup> Es fácil mostrar esto con un ejemplo. Pensemos en un programa que tuviera todas las oraciones posibles en español con respuestas para todas ellas. Este programa simplemente buscaría las frases que se le digan en una lista y daría la respuesta pre-programada para dicha frase. Empero, esto no bastaría para mantener una conversación coherente, para ello el programa tendría que tener todas las conversaciones posibles. Con esto sería suficiente para pasar la prueba de Turing. Se trata de algo imposible por limitaciones de almacenamiento y tiempo, pero posible en principio. Sin embargo, este programa no sería inteligente en ningún sentido relevante.

cuadro de Rembrandt y el platillo mencionado. Quizá incluso exista en el mundo un pequeño poblado que cumpla con estas características, sin embargo se trata de un buen criterio para saber si un poblado es una gran ciudad.

Esto, aplicado al problema de los criterios de corrección epistémica, lleva a enfatizar que éstos no tienen que ser un análisis de lo que significa una regla racional. Si bien un análisis puede proveer un criterio, no todos los criterios corresponden a un análisis. Lo último es más fuerte. Tal como se plantea esta investigación, cada criterio de manera individual no corresponde a un análisis de la racionalidad. Esto contrasta con aquellos teóricos que sostienen que, por ejemplo, la ley de Bayes, la teoría de la decisión o las máquinas de Turing, capturan la racionalidad al modo de una idealización o abstracción. En cambio aquí, lo que constituye la racionalidad son el conjunto de reglas de razonamiento, que a su vez determinan el conjunto de relaciones inferenciales que debe haber entre creencias. Los criterios sirven para generar las reglas de razonamiento que necesitamos, y como mencioné arriba, para tener una visión global acerca de cuáles son estas reglas.

### **Métodos basados en intuiciones**

Escoger un método para validar reglas de razonamiento plantea dos problemas fundamentales:

- a) Hace falta una explicación que diga por qué el método provee reglas adecuadas; dicho de otro modo: ¿qué significa que una regla sea buena desde un punto de vista epistemológico? Esto lleva a problemas de normatividad epistémica (y secundariamente, de meta-epistemología).
- b) Debe ser posible aplicar el método, al menos de manera imperfecta. Estos problemas de aplicación son propios de cada teoría.

Como se puede ver en los ejemplos de reglas y criterio que utilicé en el apartado anterior, las alternativas son muy diversas, ejemplos tan distintos como el pragmatismo social y el naturalismo epistemológico son una muestra de esta diversidad. Algunos de estos modos podrán ser a priori y otros a posteriori; y sin importar esta distinción, los criterios pueden o no

ser sujetos a formalizaciones en matemáticas (como el programa de heurística de Todd y Gigerenzer).

En este trabajo propongo que los métodos que utilizan intuiciones epistémicas son capaces de proveer un modo para validar reglas de razonamiento. Si bien se trata de algo ampliamente practicado en la epistemología contemporánea, es necesario mostrar cómo se resuelven los problemas tanto de normatividad epistémica como de aplicación.

Lo que haré en este apartado será caracterizar los métodos basados en intuiciones. Stich, Nichols y Weinberg (2001) en un trabajo crítico, señalan tres características generales que tendría que cumplir cualquier estrategia de validación epistémica que sea de este tipo (MI):<sup>26</sup>

- a) La estrategia de tomar intuiciones como insumo (no exclusivamente).<sup>27</sup>
- b) Debe tener como producto, principios o aserciones normativas sobre cómo debemos proceder al razonar.
- c) El producto debe depender, al menos en parte, de las intuiciones insumo.

A pesar de que el uso del término “intuiciones” en la acepción que le damos es reciente, los filósofos desde hace siglos apelan a las intuiciones, no siempre de manera explícita (Goldman: 2007). En esta investigación recurriré a las intuiciones en el sentido de aceptación acerca de la aplicabilidad de un término. Este sentido es el que ha recibido primordialmente en buena parte de la filosofía contemporánea.

¿Qué son las intuiciones epistémicas? ¿Qué es lo que estamos diciendo cuando señalamos que nuestra intuición dice que no está justificado formar una creencia lanzando un dado o una moneda? La interpretación más plausible acerca de las intuiciones, y una que aparece frecuentemente en la literatura (e.g. Goldman: 2007; Nagel: 2007; Stich, Nichols, Weinberg: 2001), es una explicación lingüística.

---

<sup>26</sup> El nombre sugerido es “Intuition Driven Romanticism” (romanticismo basado en intuiciones), debido a las creencias de los poetas decimonónicos de que la verdad se encuentra en el interior.

<sup>27</sup> Por ejemplo, imaginemos un método *M* muy sencillo donde las intuiciones no son el único insumo. Este método consiste en comparar intuiciones preteoréticas con resultados científicos. El método consistiría en tomar un par de proposiciones contradictorias, una preteorética y otra validada por el consenso científico, y escoger una de las dos. Este método no tiene mucha utilidad, pero quedaría dentro de lo que Stich *et al.* consideran método basado en intuiciones.

Sabemos que existen términos epistémicos: justificación, racionalidad, conocimiento. Decir que un sujeto tiene la intuición de que una creencia que fue formada lanzando una moneda no está justificada, simplemente es decir que el sujeto no considera que se aplique el término “justificación” a ese caso; tal como otro sujeto podría decir que no un mueble para sentarse si respaldo no debe ser llamado “silla.”

La intuición de lo que cuenta como creencia justificada no tiene nada distinto con respecto a la intuición de lo que cuenta como una silla o cualquier otra cosa. Las intuiciones epistémicas son simplemente intuiciones semánticas acerca de términos epistémicos. Esto ofrece una explicación de lo que sucede cuando un filósofo utiliza intuiciones como evidencia. En el caso particular de la epistemología, describe la enorme literatura que existe sobre el análisis del conocimiento y la justificación: el llamado proyecto analítico. En esta interpretación metodológica, lo que hace Edmund Gettier (1963) con su famoso contraejemplo al análisis tradicional del conocimiento, es mostrar que éste no se ajusta a ciertas intuiciones semánticas (propias, pero con la pretensión de ser compartidas) acerca de la palabra “conocimiento”.

En la epistemología analítica, las intuiciones son una manera de comprobar si un análisis es correcto. El análisis, entendido en sentido técnico, consiste en proveer condiciones necesarias y suficientes para un término. Un buen análisis debe capturar los usos correctos del término en contextos normales (p.e. “normalmente *decimos* que un niño sabe que  $p$  aún cuando no pueda dar una razón para  $p$ ”) y por el desarrollo de experimentos mentales para probar el límite de nuestras intuiciones sobre el significado de estos términos (p.e. “si modificaran el cerebro de alguien para tener automáticamente la creencia de que la temperatura es de 30 grados cuando la temperatura es de 30 grados, ¿sabría que la temperatura es de 30 grados?”). Los casos y preguntas planteados por la epistemología analítica son respondidos afirmativa o negativamente (“sí está justificado”, “no está justificado”, etc.). En adelante, y a falta de una teoría alternativa sobre cómo darle sentido a la práctica filosófica de apelar a las intuiciones, partiré del hecho que los MI funcionan en el modo que acabo de describir.

Los MI además de utilizar intuiciones como insumos, tienen que resolver un problema en



relación al modo de acomodarlas. Para ello basta notar que no todas las intuiciones pueden ser aceptadas. En primer lugar, porque no existe una razón *prima facie* para rechazar los resultados contra-intuitivos. En las ciencias empíricas y en las matemáticas los resultados contra-intuitivos son considerados valiosos, ¿por qué habría que rechazarlos en la filosofía? (Dennett: 2005). Las intuiciones son los insumos de este tipo de métodos, pero debe haber un proceso relativamente sofisticado antes de llegar a un producto, que puede ser un análisis, regla o criterio; de manera que no es necesario ni deseable que todos los resultados sean intuitivos. Esto último puede parecer incompatible con la explicación semántica de los MI. ¿Cómo podría un método de este tipo tener como resultado que un término, digamos “justificación”, resulte aplicable en un caso contra-intuitivo, es decir en un caso que no nos parece aplicable? (Es igualmente problemático lo converso). Si nuestras intuiciones semánticas es lo que estamos tomando como evidencia, y obtenemos algunos resultados contrarios a estas intuiciones, ¿no resulta esto un poco extraño? Este problema lo trataré más adelante.

En segundo lugar, otra razón por la cual no pueden aceptarse todas intuiciones epistémicas es porque no siempre son consistentes. Esto se puede ver claramente en la literatura producida por los epistemólogos analíticos a partir de los años setenta. Por ejemplo, el análisis causal del conocimiento cubre muchos casos donde es intuitivo (al menos para los filósofos) atribuir conocimiento a los sujetos de los experimentos mentales, y este mismo análisis excluye casos donde no resulta intuitivo atribuir conocimiento (como los famosos casos Gettier). Sin embargo existen contra-ejemplos al análisis causal elaborados independientemente por Goldman y Nozick (Goldman: 1988). Esto ejemplifica algunos de los conflictos que puede haber con respecto a las intuiciones. Por una parte, el análisis causal tiene evidencia intuitiva favorable, y por otra parte tiene contra-ejemplos importantes. Nozick y Goldman, respectivamente, proponen casos sofisticados que incluyen cajones llenos de termómetros y graneros con falsas fachadas. ¿Lo correcto es aceptar el análisis causal y admitir como un resultado contra-intuitivo que sí hay conocimiento en estos dos escenarios, o considerar que estos contra-ejemplos refutan el análisis causal? Los filósofos tienen muchos modos implícitos de tomar estas decisiones. Pueden señalar que un resultado es demasiado

implausible, o que las condiciones y cláusulas de un análisis llegarían a ser excesivamente complicadas.

En el caso particular de un análisis o similar, éste es buen candidato cuando no ha sido posible encontrar un contraejemplo, es decir, una consecuencia  $A'$  del análisis  $A$ , donde  $A'$  resulta contraintuitiva. Pero quizá con todo, el análisis  $A$  siga pareciendo tan bueno que debamos aceptar esta consecuencia, o abandonarlo sea costoso por otras razones, de modo que finalmente  $A$  podría ser considerado correcto aún con su consecuencia contraintuitiva. Un criterio que lograra sistematizar casi perfectamente nuestras intuiciones epistémicas, no debería ser rechazado por producir algunas consecuencias incómodas.

Por tanto un MI razonable requiere de una manera de acomodar los resultados contraintuitivos, de señalar cuándo deben ser aceptados y cuándo rechazados. Más adelante expondré un subconjunto de estrategias denominadas de equilibrio reflexivo que han figurado prominentemente en la literatura acerca de racionalidad. Se trata de una solución muy plausible acerca de cómo manejar el balance de consecuencias intuitivas y contraintuitivas. También es muy general, de modo que obtendremos teorías muy distintas según el modo como se ponga en funcionamiento. Posteriormente hablaré del equilibrio reflexivo como un MI que puede resolver estos problemas metodológicos.

### **Las fuentes de la normatividad epistémica y el problema normativo**

Un importante problema al que se enfrentan los MI es lo que algunos autores han llamado el problema normativo (Stich, Nichols, Weinberg: 2001). El problema señala: “¿Qué razón hay para que un [MI] produzca resultados con fuerza normativa? (pp. 8-9). Es decir, ¿qué tienen de especial las intuiciones acerca de lo racional para que sean capaces de proveer un criterio que efectivamente produzca reglas correctas?

La respuesta más común a este problema se basa en el significado mismo de nuestros términos epistémicos, así la expone Kornblith (1993) criticando a Goldman:

“Supongamos que alguien preguntara, [...] ¿qué hay de bueno en la creencia justificada [...]? Goldman parece comprometerse con la siguiente respuesta. Que esté justificada es simplemente algo bueno; es parte del significado del término. [...] Un argumento similar podría construirse para otros términos epistémicos de validación [approval]. La fuerza normativa parece derivarse únicamente de consideraciones semánticas.”

Esto lleva al problema de buscar por qué la validación es parte del significado mismo del término. Para responder a esto, es necesario apelar a bienes epistémicos –típicamente la verdad– que son parte del significado de los términos epistémicos. En teoría, las intuiciones epistémicas capturarían este tipo de valores implícitos. Por ejemplo, quizá la semántica del término “racional” (aplicado a las inferencias), incluye bienes valorados como los siguientes:

- a) La inferencia llega a creencias con capacidad explicativa.
- b) La inferencia se realiza en un tiempo razonable.
- c) La inferencia tiene un grado aceptable de fiabilidad, en relación a la verdad de las creencias obtenidas.
- d) La inferencia aumenta la coherencia del conjunto doxástico.
- e) Las creencias obtenidas por medio de la inferencia ayudan a que el agente realice los fines que se propone.

Esto es sólo un ejemplo del tipo de bienes que podría indicar implícitamente un término epistémico. Empero, existen cuatro problemas que debe resolver una respuesta de este tipo al problema normativo. El primero es expuesto por Ahlstrom (2008): aunque el origen de la normatividad por consideraciones semánticas pueda explicarse con estas metas (creencias verdaderas, etc.), el análisis conceptual no garantiza que nuestras nociones epistémicas las obtengan con ellas del mejor modo. Dicho de otra manera, nosotros creemos que lo que llamamos “racionalidad” ayuda a alcanzar ciertos fines (coherencia, éxito pragmático); pero

sería posible que no fuera así. Por poner un ejemplo, el concepto de “rey” utilizado en el absolutismo, podía incluir ciertos valores y metas; por ejemplo, que el rey fuera el encargado de mantener la concordia de sus súbditos. Sin embargo, un análisis de los efectos reales de la monarquía podría mostrarnos que en realidad el rey no cumple con esta función, o que incluso genera discordia. ¿Qué evita que suceda algo similar con los términos epistémicos?

El segundo es que la conexión conceptual anteriormente descrita puede garantizar que nuestros conceptos epistémicos generen reglas con la suficiente relación con nuestras metas epistémicas. Sin embargo este hecho es compatible con la posibilidad de que exista un conjunto de normas alternativo que resulte mejor para llegar a estas mismas metas. Aquí también ayudará un ejemplo. Supongamos que el médico le indica a un paciente que tiene que correr todas las mañanas para que resuelva un problema de circulación que tiene en las piernas. El objetivo que se busca, curar el problema circulatorio, es lo que le da fuerza normativa a la instrucción de correr. Empero, si existen otras maneras de mejorar la circulación, correr deja de ser imperativo de un modo absoluto (o dicho en términos kantianos, categórico).<sup>28</sup>

El tercer problema señalado por Stich (1993), es que una vez que un grupo de personas pertenecientes a una cultura logran establecer una normatividad epistémica por los medios aquí descritos, no existe ninguna razón por la cual otras culturas deberían adoptarla, pues un modo consecuencialista de realizar evaluaciones cognitivas es sensible al contexto. “La probabilidad de que un sistema de procesos cognitivos haga un mejor trabajo que otro asegurando cualquier meta plausible, dependerá de factores como la existencia de un sistema de escritura, la existencia de comunidades de expertos y la relación de estas comunidades con los arreglos sociopolíticos de la sociedad como un todo.” (pp. 139, 140).

El cuarto problema podría ser que este tipo de respuesta requiere demostrar otras tesis, como que la verdad está relacionada con el éxito pragmático, o que la verdad es correspondencia.

---

<sup>28</sup> Para una aplicación de las nociones kantianas de imperativo categórico e hipotético aplicadas a la epistemología, ver por ejemplo Kornblith (1993).

Para el primer problema, el mismo Ahlstrom<sup>29</sup> (2008) ofrece una solución; sin embargo algo similar había sido descrito por Goldman (1988) en un contexto distinto (el escepticismo) llamándolo “epistemic bootstrapping.” En esencia, este tipo de respuesta señala que basta con tener las herramientas epistémicas mínimas para construir mejores herramientas, y repetir este procedimiento las veces que sea necesario. El análisis de los conceptos es sólo un punto de partida, pero después de éste, es necesario establecer de manera empírica si nuestras prácticas cumplen con estas metas que resultaron en un análisis.

El segundo problema está relacionado con las diferencias. ¿Por qué unas culturas tendrían que dejar sus propios estándares a partir de un análisis de conceptos de otras culturas? Un primer modo de responder a esto es intentando mantener pretensiones universalistas; es decir, que los resultados del análisis puedan aplicarse a cualquier grupo social. Aquí el universalista puede echar mano de los argumentos que relacionan a la agencia epistémica humana con el hecho de que su fin es buscar verdades (Dennett: 1989), lo cual tendría por consecuencia que los estándares epistémicos de distintas culturas tendieran a ser convergentes; y en caso de ser divergentes un grupo aceptará que las prácticas epistémicas del otro son mejores, puesto que estos dos grupos buscan los mismos fines y bastaría con ver quién los cumple mejor. Parece algo evidente que la verdad debiera ser un valor universal, pero también se ha argumentado que no tiene por qué ser así (ver por ejemplo, Stich: 1993).

Otra respuesta de tipo universalista, alguna vez propuesta por Kornblith (1993), consiste en conectar las creencias verdaderas (aunque podrían ser otros fines epistémicos) con la estructura general de nuestros deseos, de una manera que la estructura de la agencia humana (y de cualquier ser racional que tenga deseos) incluye el hecho de valorar la verdad. De manera independiente al éxito que puedan llegar a tener de este tipo de argumentos, considero que un criterio no tiene que ser universalmente aplicable para ser normativo. La racionalidad mínima requiere normas de evaluación epistémica, pero estas normas no tienen

---

<sup>29</sup> El artículo de Ahlstrom defiende un uso naturalista de las intuiciones, y señala que el análisis como método principal lleva a este tipo de limitaciones en relación al mejoramiento epistémico. Sin embargo provee una solución al problema de la normatividad resultante de dichos análisis.

que aplicar en cualquier contexto (Eraña: 2002), lo cual hace posible sostener un pluralismo (pero no necesariamente un relativismo epistémico en un sentido más fuerte, pues no es relativo a cada sujeto).

Esto constituye una respuesta satisfactoria al problema normativo, siempre que los individuos de un mismo contexto sociocultural compartan las suficientes intuiciones. Del rechazo de un estudio de la racionalidad puramente a priori, se sigue que una parte de la investigación debe ser empírica (Eraña: 2002) y que es necesario incluir diversos grupos sociales. Si los métodos basados en intuiciones (MI) estudian conceptos que no son universales, entonces sus resultados (sin ayudarse de otro tipo de investigación) sólo son limitados en tanto que no proveen una investigación exhaustiva acerca de qué es la racionalidad. En resumen, considero que existen argumentos buenos tanto desde un punto de vista para pensar que los términos epistémicos son universalmente compartidos como para pensar que no lo son, y que en ambos casos existe la posibilidad de apelar a las intuiciones epistémicas.

El último problema señalado, es que una respuesta en estas líneas al problema normativo implicaría un compromiso —aparentemente inaceptable— con una teoría de la verdad no-doxástica. De lo contrario, si la verdad depende de factores epistémicos, entonces ya no podemos utilizarla para explicar nociones epistémicas, bajo pena de circularidad.

La preocupación por evitar la circularidad es comprensible si lo que se busca es explicar lo epistémico por medio de lo no-epistémico. Sin embargo, esto no es el objetivo de esta investigación. Ni siquiera un análisis —el caso más fuerte de criterio— es inválido o trivial tan sólo por ser circular en este sentido. Para un proyecto naturalista que pretende dar cuenta de la normatividad sin un lenguaje normativo como la meta-filosofía denominada “canberrista” (Braddon-Mitchell, Nola: 2009) resulta esencial que no aparezcan nociones normativas en las explicaciones. Pero esto es en cierta manera independiente de un estudio de la racionalidad epistémica, donde se puede explicar una noción normativa (la racionalidad) con otras diferentes. Una meta-filosofía como la anteriormente descrita puede hacer imperativo analizar a su vez estas nociones normativas, lo cual no le resta validez a la primera reducción teórica.

El objetivo de este apartado ha sido mostrar de manera general algunos modos como se

puede resolver el problema normativo con una respuesta consecuencialista. Sin embargo desarrollar el tema a fondo rebasaría los alcances de este trabajo, cuyo objetivo es establecer cómo se pueden conocer las reglas de razonamiento adecuadas para avanzar en una teoría de la racionalidad.

#### **4. EL EQUILIBRIO REFLEXIVO COMO MÉTODO**

##### **Equilibrio Reflexivo Histórico y Clásico**

En el capítulo anterior mostré que algunos métodos basados en intuiciones (MI) pueden proveer una alternativa viable para decidir cuáles son las reglas de razonamiento adecuadas. Para ello, deben de satisfacer dos tipos de criterio. El primero de ellos es relativo al problema de la normatividad epistémica. Es decir, establecer por qué es obligatorio o permisible seguir las reglas obtenidas por un MI. La respuesta que ofrecí en el capítulo anterior, a grandes rasgos, consiste en relacionar las intuiciones con conceptos, y éstos con ciertos fines o bienes epistémicos.

El segundo aspecto que debe satisfacer un MI para ser una solución viable al problema de encontrar un criterio tiene que ver con el aspecto metodológico. Es necesario que las intuiciones tengan una relación apropiada con los conceptos epistémicos, de manera que sea posible investigar los conceptos a través de las mismas. En el capítulo anterior mostré cómo pueden responderse estas preguntas en cuanto al primer aspecto, el de la normatividad. Queda por responder la otra parte, referente a la aplicabilidad.

Un problema distinto es establecer para un método concreto, si puede satisfacer estos criterios. Esto resulta más complicado que establecer la verosimilitud inicial de los MI (el trabajo del capítulo anterior). Afortunadamente, considero que al menos existe un MI que puede responder de manera satisfactoria a ambos criterios (si bien con algunos problemas que quedarán pendientes).

El método que considero que responde mejor a los problemas planteados es el equilibrio reflexivo (ER). Éste consiste, a grandes rasgos, en considerar las consecuencias intuitivas y

contraintuitivas de aceptar una regla para decidir si es aceptable. El ER es una propuesta de Goodman (1983), que sido debatida subsecuentemente en el contexto de la presente investigación por Cohen (1989), Stich (1993), Cummins (1998), Goldman (2007), y otros; así como en ética (Daniels: 1996).

El ER propuesto por Goodman no coincide en todos los puntos con el que se ha discutido más recientemente. Por ello considero necesario hacer una distinción, al menos entre dos tipos de ER. El equilibrio reflexivo histórico (ERH) es el termino que usaré para referirme a lo que se propone en el texto de Goodman. Para referirme al equilibrio reflexivo del que hablan autores contemporáneos (Stich: 1993 utiliza el término de “proyecto neo-goodmaniano”), le llamaré equilibrio reflexivo clásico (ERC). La diferencia principal entre ellos, es que el equilibrio reflexivo clásico es una teoría más específica, que llena algunos vacíos que Goodman no aclara explícitamente. El ERC tal como lo describen Goldman (1988) y Stich (1986, 1993), —y también en trabajos de ética— se considera explícitamente un método acerca de intuiciones (MI). Otros presupuestos del ERC que también son importantes (pero negociables), es que se trata de un método individual, practicado en primera persona. El texto de Goodman no es explícito sobre ninguno de estos puntos. Conociendo los escritos posteriores que discuten el equilibrio reflexivo, resulta natural interpretar a Goodman de manera que su propuesta (ERH) sea parecida o igual a la que aparece en éstos (ERC). Sin embargo la interpretación del ERH ofrece otras posibilidades, un ejemplo claro de esto es el trabajo de Catherine Elgin, *Considered Judgements* (1999). Elgin elabora una teoría de la justificación, con espíritu goodmaniano, muy distinta al ERC, pues considera más bien factores sociales.

¿Goodman tenía en mente una teoría pragmatista social al estilo de Elgin, o una (aquí llamada “clásica”) como la expuesta por Stich? Para responder esta pregunta sería necesario realizar una labor de tipo exegético, que no es parte de los objetivos del presente trabajo. Por ello, estipular una distinción entre ERH y ERC permite dejar de lado el problema. Sin importar si ERH es una formulación elíptica de las teorías que se han hecho de ERC o no, esto no afectará la presente exposición.

Mi objetivo a continuación será reseñar brevemente lo que dice Goodman sobre el método



para contextualizar el equilibrio reflexivo. Nelson Goodman (1983) lo sugirió como un modo de justificar la inducción –sin darle un nombre al método– en el tratado clásico de filosofía de la ciencia *Fact, fiction and forecast* publicado por primera vez en 1955. Rawls llama a un método similar (mas no idéntico) “equilibrio reflexivo” [reflective equilibrium] en *A theory of justice* (2005, primera edición 1971) con una referencia explícita a Goodman. A partir de este último trabajo el criterio empezó a ser usado y discutido en teoría moral, ética aplicada, bioética y filosofía política. Pero también es retomado como tema en la epistemología y en el debate de racionalidad de la filosofía de las ciencias cognitivas (Cfr. Stein: 1998).

Goodman plantea su método en el contexto de la justificación de una regla inductiva, lo cual es modo de plantear el problema clásico de la inducción distinto al usual (normalmente no se habla de reglas inductivas, sino de inferencias inductivas). La estrategia de Goodman (1983) consiste en disolver el problema de Hume, del modo en que —a partir de Kripke (1982)— comúnmente se denomina “solución escéptica”. Es decir, una solución que acepta la conclusión de un argumento escéptico (en este caso, que la inducción es injustificable de manera no-circular), pero que muestra que la misma no tiene las consecuencias fatales que el argumento original pretende tener. Hume mismo ya había ofrecido una solución escéptica a su propio problema al señalar que los orígenes de la inducción son psicológicos y no epistémicos.

Para Goodman existe un “problema inductivo”<sup>30</sup> tanto como existe un “problema deductivo”, pues todas las reglas las justificamos del mismo modo (incluso las morales, estéticas, etc.). Tanto las inferencias inductivas como las deductivas se justifican apelando a reglas generales. En el caso de las reglas deductivas son reglas sancionadas “por una práctica deductiva aceptada”:

“La validez de una deducción no depende de su conformidad con reglas arbitrariamente válidas que podamos ingeniar, sino en conformidad con reglas válidas. [...] ¿Pero cómo se determina esta validez? [...] Considero que la respuesta se encuentra mucho más cerca de la superficie. Los principios de inferencia deductiva se justifican por su conformidad con la práctica deductiva aceptada. La

---

<sup>30</sup> Para una exposición concisa del problema de la inducción ver George Dicker, *Hume's Epistemology and Metaphysics*, 1998, cap. 3.

validez depende de las inferencias deductivas que de hecho hacemos y aprobamos.” (pp. 63-64).

Queda pendiente el problema de cómo justificar una regla particular. Goodman consideró en su momento que eran infructuosos los intentos de explicar de manera psicologista o como autoevidentes a los principios deductivos. Ofrece la siguiente estrategia:

Si una regla lleva a inferencias inaceptables la abandonamos como inválida. La justificación de las reglas generales, depende así de los juicios que aceptan o rechazan inferencias particulares. Esto parece fragantemente circular. Lo que he dicho es que las inferencias se justifican en cuanto a su conformidad con reglas generales y que las reglas generales se justifican en conformidad con inferencias válidas. Pero este círculo es virtuoso. El punto es que las reglas y las inferencias particulares se justifican al ponerse en concordancia [agreement] unas con otras. Una regla se corrige si lleva a inferencias que no estamos dispuestos a aceptar; y una inferencia es rechazada si viola reglas que no estamos dispuestos a corregir. El proceso de justificación es delicado, consistente en hacer ajustes mutuos entre las inferencias aceptadas y las reglas aceptadas, y en la concordancia que se logra yace toda la justificación que se necesita para ambas.”. (p. 65, énfasis en el original).

Por ejemplo, tomemos una regla que resulta inicialmente aceptable para muchas personas, la regla de afirmación del consecuente:

*Si p entonces q*

*q*

*Por lo tanto p*

Esta regla tiene muchas consecuencias como la siguiente aplicación. Supongamos que el día de el equinoccio de otoño, un individuo realiza el siguiente razonamiento, cuyas premisas son verdaderas:

*Si es el primer día de primavera entonces hay equinoccio.*

*Hay equinoccio.*

*Por tanto es el primer día de primavera.*

Las consecuencias de aplicar esta regla son muy poco intuitivas. O bien es primavera y otoño al mismo tiempo, o bien el sujeto debe aceptar que no es otoño (a pesar de que tiene evidencia muy fuerte de que sí lo es), etc. Del mismo modo, pueden construirse muchos argumentos que tengan como resultado conclusiones falsas, y éstas a su vez consecuencias lógicas falsas. Las infinitas consecuencias que provienen de aceptar la regla de la afirmación del consecuente superan el atractivo que originalmente posee.

El equilibrio reflexivo clásico es un método basado en intuiciones, donde se decide la aceptación o negación de una regla con base a lo intuitivo o contraintuitivo que resulte dicha regla, y a las consecuencias intuitivas o contraintuitivas que ésta tenga. Las intuiciones son tomadas como disposiciones a aceptar o no aceptar las reglas o consecuencias de las reglas. Como señala Hintikka (1999), buena parte del uso de intuiciones como evidencia por parte de los filósofos contemporáneos no se compromete con una teoría acerca de qué son estas intuiciones (más allá del sentido mínimo del término) y de cómo funcionan. Es notorio, desde un punto de vista histórico, que los filósofos que apelaron a intuiciones como Platón y Kant (pero no todos, e.g. Locke), tenían una teoría que explicaba por qué podían ser evidencia válida (Hintikka: 1999). El ERC bien puede caer dentro de esta crítica, a lo cual se ha respondido elaborando respuestas más completas sobre las intuiciones (y en particular las intuiciones epistémicas). Esto lo explicaré posteriormente.

¿Sobre qué deben ser las intuiciones del ERC? En el ejemplo anteriormente mencionado utilicé la falacia de la afirmación del consecuente, es decir, una regla. Esto es lo más cercano a la descripción del ERH, y es el modo como lo han tratado autores como Cohen (1989), Stein (1998), Stich y Nisbett (1980). Pero el ERC también puede ser aplicado a un criterio de corrección con el cual se juzgarían las reglas, Goldman (1988).

Una motivación para adoptar el primer tipo de estrategia puede ser considerar como plausible

la teoría psicológica de que existen competencias de razonamiento. Cohen (1989), se basa en intuiciones preteóricas acerca de la corrección de inferencias particulares, las cuales se supone que son universales. Estas intuiciones se pueden considerar análogas a las intuiciones lingüísticas de los hablantes competentes de un lenguaje en el contexto de la lingüística computacional (Stein: 1998). Los hablantes competentes del español tienen la intuición de que “el gato está en el granero” es una oración gramaticalmente correcta, debido a las competencias gramaticales innatas que poseen; aunque pueden tener errores de *ejecución* (por distracción, etc.). De acuerdo con Cohen, evaluar si una regla de razonamiento es correcta implica contrastarla con las intuiciones particulares sobre las inferencias particulares que se producen a partir de ellas.

Sin embargo, considero que los problemas notados tempranamente por Stich y Nisbett (1980) resultan devastadores para un ERC aplicado sobre reglas. Por ello expondré a continuación estos problemas.

Un problema importante que tiene tomar esta interpretación es que reglas patentemente inválidas resultan estar justificadas. El siguiente ejemplo es ilustrativo (Stich y Nisbett: 1980). Es común que los aficionados a los juegos de azar tomen por cierta una regla de inferencia que por ellos se denomina “falacia del apostador”. Ésta señala que para un juego de azar justo, la probabilidad de obtener un resultado es mayor cuando no ha salido en  $n+1$  tiros que en  $n$  tiros. Esta falacia puede refutarse muy fácilmente por medios matemáticos o de manera empírica, con un análisis estadístico que comprobará que la ausencia de un suceso en experimentos<sup>31</sup> anteriores no favorece su aparición si el método que usamos es aleatorio. El problema es que para quien sostiene la falacia (“el apostador”), esta regla se encuentra en equilibrio reflexivo. El apostador al considerar sus intuiciones acerca del azar no encuentra nada indebido, e incluso cree que obtiene soporte empírico de su regla cada vez que participa en juegos de azar. Cuando después de muchas tiradas de la ruleta sin que salgan números rojos, sale por fin un número rojo, el apostador lo atribuye a ello.

Una posibilidad es aceptar que esto no resulta un contraejemplo al ERC como método directo

---

<sup>31</sup> Este es el término técnico en probabilidad para referirse a una tirada de dado, un lanzamiento de moneda, etc.

para mostrar que una regla es correcta (a), y que por tanto la falacia del apostador es una inferencia correcta mientras el apostador la sostenga en equilibrio reflexivo. De hecho cualquier cosa estaría justificada mientras un individuo la sostenga en equilibrio reflexivo. Llamémosle a esto ERC irrestricto. Además del ejemplo de Stich y Nisbett, existe evidencia anecdótica de sujetos que aceptan una gran cantidad de inferencias tradicionalmente falaces, aún después de reflexionar sobre ellas (Cfr. Stein: 1998). ¿Es completamente implausible mantener el ERC irrestricto (aplicado a reglas)? Tal parece ser la postura de Stich (1988, 1994), al señalar que no se trata de una alternativa seria. Al parecer, la razón de fondo es la idea de que esta postura podría sancionar cualquier inferencia. Sin duda una teoría de la racionalidad epistémica tan permisiva sería inútil.

¿Pero en realidad tiene estas consecuencias la respuesta que Stich se niega siquiera a considerar? Seguramente existen límites psicológicos para lo que un individuo pueda aceptar (o creer) en equilibrio reflexivo, y por tanto el ERC irrestricto no podría sancionar cualquier inferencia, porque no toda transición sería posible desde un punto de vista psicológico. Sin embargo, en este caso se requiere un modo de caracterizar lo que es una creencia independientemente de las relaciones inferenciales, a diferencia de lo que sucede en la teoría de la racionalidad mínima que se adoptó en el primer capítulo de esta investigación. Es decir, una noción de creencia que no sea puramente epistémica, pues de otro modo existiría una circularidad viciosa. Lo que se requeriría entonces para evitar dicha circularidad es una noción de creencia ya sea definida a partir de los deseos, o caracterizándola como un tipo de representación mental. Debido a que en este trabajo estoy utilizando una noción de creencia puramente epistémica, el ERC irrestricto no es una opción disponible. Pero es importante señalar que no es una vía absurda.

Ya que el ERC irrestricto aplicado a reglas no es una opción disponible, es necesario buscar una solución a los contraejemplos del tipo que propone Stich. Existen dos respuestas al respecto. Una es el equilibrio reflexivo amplio que defienden autores como Rawls y Daniels (Stein: 1998), y otra es el recurso a expertos.

El recurso a expertos es una manera de incluir factores sociales en la justificación de una regla (una especie de división del trabajo epistémico). Según la teoría del ER, desde un punto

de vista individual, un sujeto no tiene un criterio superior al ER para juzgar sus prácticas epistémicas. Sin embargo “la deferencia a la autoridad no sólo es la práctica habitual de la gente educada, sino que es lo correcto desde un punto de vista normativo” (Stich y Nisbett: 1980). En el caso de la falacia anteriormente descrita, a pesar de que el apostador se encuentra en ER, la comunidad de expertos en cálculo de probabilidades no podría sostener la regla en ER.

Un primer problema con el recurso a un experto, es que desde un punto de vista puramente epistémico, no existe un modo de identificar a los expertos de manera no circular. Si un experto se define como aquél que realiza juicios correctos, y el criterio más alto para determinar juicios correctos es el ER entonces existe un problema de circularidad (Stich: 1993).

De modo que es necesario encontrar una manera de identificar a los expertos que no sea epistémica. Esto resulta más difícil de lo que parece. Salvo ciertos casos problemáticos, por lo general es relativamente fácil identificar a un experto debido a que provee información adecuada. Sin embargo, para Stich y Nisbett este tipo de evaluaciones no están disponibles, pues lo adecuado se define como aquello sancionado por un experto. Para contrastar puede verse el siguiente ejemplo: Goldman (2001) propone la fiabilidad como uno de los criterios mediante los cuales el novato puede escoger a qué experto seguir. Pero esto es posible porque Goldman tiene una teoría epistemológica con la cual puede defender de manera independiente la importancia de la fiabilidad (con respecto a creencias verdaderas). Aquí en cambio, en el contexto de validar el ERC como fundamento de una teoría de la racionalidad, el recurso a la fiabilidad sería gratuito.

¿De qué otra forma es posible identificar a los expertos sin presuponer algunas reglas como correctas? El éxito pragmático es una opción. Los expertos serían aquellos cuyos consejos hacen que los novicios que los siguen obtengan mejores resultados prácticos. Pero en este caso parece innecesario hablar de expertos. Si el éxito práctico es un última instancia la base de una teoría sobre la racionalidad, no es necesario basarlo en expertos. Por otro lado, el éxito pragmático no es una noción lo suficientemente robusta, ya que está sujeto a problemas relacionados con la suerte epistémica.

La moraleja parece ser la siguiente, si los expertos van a tener un papel de fundamento en una teoría de la racionalidad, no se les puede definir a estos de manera puramente epistémica. La solución que identifican Stich y Nisbett (1980) es adoptar una noción de autoridad como socialmente designada, lo cual también presenta problemas. En primer lugar, presenta una noción de inferencia racional que se apoya demasiado en el aspecto social. Es concebible *prima facie* una sociedad en que los expertos se designaran al azar y cuyas opiniones fueran respetadas, a menos que haya factores de tipo sociológicos que restringieran este tipo de casos. Esto claramente va en contra de la finalidad que se busca al proponer una teoría de la racionalidad epistémica. Alguna teoría social sobre los expertos podría señalar que un experto es designado con base en bienestar social que produce, si las cosas son así, tomar a las autoridades epistémicas como las designadas no caería en los problemas como el que acabo de describir. Esto último es una cuestión empírica perteneciente al ámbito de la teoría social, por lo cual no considero que existan razones *a priori* a favor o contra ella. Sin embargo, considero que esto no debería ser parte de los fundamentos de una teoría sobre la racionalidad, pues en ese caso tendríamos una teoría epistemológica que en el fondo sería una sociología del conocimiento. En este caso sería más propicio adoptar un enfoque abiertamente sociológico.

Una segunda razón en contra de identificar las autoridades epistémicas con las autoridades epistémicas de facto (aquéllas designadas como tales), es un problema señalado por Stich y Nisbett (1980). Analizan el caso hipotético de un “rebelde cognitivo”. Éste no es más que un individuo que no acepta seguir la opinión de los expertos, y además el rebelde no cuestiona que los individuos en cuestión sean expertos. El problema con esto, es que si el ER experto fuera un análisis correcto de la justificación, el rebelde cognitivo estaría contradiciéndose al señalar que aquellos a quienes consideran expertos se encuentran equivocados en una ocasión particular; estaría diciendo algo así como “esto está justificado y no está justificado”. Los autores sugieren reparar el análisis del siguiente modo: “La regla *r* está justificada si se conforman con las prácticas inferenciales de la persona o grupo de personas que yo (el hablante) considero apropiada o apropiados” (Stich y Nisbett: 1980, p. 201).

Este análisis del ERC (aplicado a reglas) en versión de experto permite que el rebelde

cognitivo pueda rechazar a las autoridades epistémicas socialmente designadas sin contradicción, lo cual parece ser lo correcto.<sup>32</sup> Sin embargo sólo es un análisis útil para la objetividad entre los “conservadores cognitivos”, es decir entre quienes aceptan a las autoridades epistémicas. La conclusión de Nisbett y Stich es que el ER experto no puede ser un análisis de la justificación y que únicamente deja espacio para “para argumentos racionales de justificación *entre* conservadores cognitivos”, es decir, entre quienes aceptan las mismas autoridades epistémicas, pero a su vez éstas no pueden ser evaluadas de manera epistémica. Esta solución, como admiten los mismos proponentes, no es una análisis satisfactorio de lo que es una regla correcta, pero ya que el objetivo de este trabajo es buscar un criterio (y no un análisis en sentido riguroso), ¿podría servir entonces el recurso a los expertos? Lo que no queda claro es de dónde procede el carácter normativo de las intuiciones de los expertos. ¿Qué tienen de especial sus intuiciones para que tengamos que obedecerlas?

Una opción puede ser que la autoridad provenga de sus respectivas disciplinas, lo cual hace que la postura colapse en un naturalismo metodológico. Es decir, las intuiciones de un científico son autoritativas porque están informadas por alguna ciencia. Esto es incompatible con un proyecto de basar lo que es una inferencia correcta en intuiciones epistémicas, es decir, un MI. Las intuiciones de los científicos (incluyendo las ciencias formales) tienen un papel dentro de la práctica ordinaria, pero esto no significa que las ciencias empíricas sean métodos basados en intuiciones (Kornblith: 2005). Si desea saber si una roca con una determinada descripción es o no sedimentaria, el geólogo tras escuchar la descripción puede considerar durante unos momentos las consecuencias de ambas alternativas y finalmente dar una respuesta; pero esto no es un método justificado a priori, son intuiciones informadas por las teorías que la geología sostiene, y en cualquier momento pueden ser cuestionadas por más datos empíricos. Parece, sin embargo, que no todas las preguntas epistémicas están relacionadas con alguna ciencia. No siempre nos preguntamos cuál es el mejor modo de apostar, o el mejor modo de identificar alguna roca. Existe la posibilidad de preguntar en

---

<sup>32</sup> Otra estrategia sería admitir que el rebelde epistémico se está contradiciendo, pero resulta muy poco intuitiva.



abstracto si una determinada inferencia es una buena regla de razonamiento. ¿Quién es el experto que debería responder a esa pregunta? Un lógico no podría, pues estudia diferentes lenguajes que varían en las reglas que los constituyen.

Aún salvando el problema de Stich y Nisbett acerca de que el recurso a expertos no es un análisis, parece que no es posible mantener un MI que cumpla con las intenciones que motivaron el ER clásico. En tanto que las maneras que las intuiciones de las personas que la mayoría (conservadora) reconoce como expertos son adecuadas, esta propuesta colapsa en una teoría de tipo sociológico, y no provee recursos independientes para identificar expertos.

Parece necesario entonces abandonar el recurso a expertos. Una segunda familia de alternativas es hacer más amplios los insumos a partir de los cuales se realiza la reflexión, lo que se denomina “equilibrio reflexivo amplio” [wide reflective equilibrium] (Stein: 1998). En la discusión anterior, se asumía que las intuiciones relevantes eran epistémicas. La propuesta de sustituirlo por un ER amplio, implicaría obligar a quien evalúa la regla a incluir intuiciones de otras áreas del conocimiento (metafísicas, éticas, etc.). Por ejemplo, en el caso del apostador, la creencia de que si no ha salido 6 en un dado justo durante muchas tiradas es más probable que salga, puede derrumbarse al considerar que implicaría la siguiente disyunción:

*O las tiradas anteriores tendrían una relación causal con las posteriores, o aún sin tener ninguna relación causal, pueden influir en el resultado.*

Como se puede ver, ambos cuernos del dilema resultan contraintuitivos desde un punto de vista metafísico. Por tanto, parecería que al ampliar los insumos, el apostador no aceptaría las consecuencias de sostener la falacia. Así, el ER amplio resuelve el contraejemplo del apostador planteado por Stich y Nisbett, sin embargo trae nuevas complicaciones. ¿Qué tan lejos tiene que llevar la reflexión un individuo? Esto no queda claro, pues siempre podría haber nuevas áreas del conocimiento capaces de cambiar las intuiciones en una dirección relevante, pero no queda claro en qué casos. ¿Cuándo son relevantes las intuiciones éticas, metafísicas, matemáticas, físicas, etc.? Del mismo modo, parece difícil sostener que las intuiciones preteóricas sean confiables en todos estos ámbitos. Supuestamente, para una

defensa de los MI como la propuesta aquí, los conceptos epistémicos llevan un importe normativo por los valores que son parte de su contenido. Las intuiciones morales podrían encontrarse en el mismo caso, ¿pero qué hay de las intuiciones lógicas o metafísicas? No existe un argumento análogo que indique que tengan un valor normativo en términos consecuencialistas, ni hay *prima facie* motivos para pensar en uno. De hecho existe una gran cantidad de literatura que puede interpretarse de modo que diga que los seres humanos somos muy propensos a aplicar reglas inválidas de lógica y probabilidad (Stein: 1998), las cuales aparentemente juzgamos como buenas a pesar de ver algunas de las inferencias que resultan de ellas.

El recurso a expertos y el ERC amplio son las mejores alternativas para mantener el ERC acerca de reglas particulares. Ninguna de las dos alternativas resuelve el problema patente sugerido por Stich y Nisbett (1980). Y si bien en vez de resolverlo, se puede evitar el problema del todo aceptando un ERC irrestricto, como indiqué arriba no es una alternativa que esté disponible, ya que es incompatible con la teoría de la racionalidad inferencial desarrollada aquí (mas no una razón independiente para rechazarla). La conclusión debe ser que, desde la perspectiva adoptada en este trabajo, el ERC aplicado a reglas es inviable.

Sin embargo, si tomamos la segunda interpretación donde el ERC se refiere a la aceptación de un criterio y no una regla, estos problemas no aparecen, al menos inmediatamente. Se sabe que existen reglas falaces sostenidas por individuos en ERC, como la falacia del apostador, pero no hay contraejemplos inmediatos de personas que sostengan criterios epistémicos inaceptables en ERC. Quizá nuestros juicios sobre reglas son excesivamente vulnerables a sesgos y errores, pero un criterio de corrección no.

En este caso el ER sería un meta-criterio que permitiría a su vez encontrar criterios para escoger reglas. Por ejemplo el ER podría arrojar como resultado que el criterio C-1 es el correcto (p.e. escoger la regla que produzca más verdades, o que disminuya las contradicciones al mínimo, etc.). Una vez que se aplica el criterio, las reglas mismas y las inferencias que producen podrían no estar en equilibrio en el sentido directo descrito en los párrafos anteriores (a). El criterio C-1 podría generar una regla R-1, y la regla R-1 generar las inferencias  $I_1$ - $I_n$ , todas ellas contraintuitivas. Pero en ese caso no habría nada que corregir. Si

la aplicación del ER fuera para cada regla, R-1 se tendría que eliminar. Pero en el caso de que el ER sea un meta-criterio, mientras  $I_1-I_n$  no tengan un peso suficiente para que rechacemos C-1, debemos sostener R-1.

El ERC aplicado a criterios necesita recurrir a los conceptos epistémicos, tales como justificación, o racionalidad. Explorando diversos casos hipotéticos, es posible saber cuáles son los factores relevantes para que estos conceptos epistémicos puedan aplicarse. Por ejemplo, si un epistemólogo sospecha que la causalidad puede jugar un papel relevante, puede diseñar experimentos mentales en los que los nexos causales se rompen de maneras extrañas. Si se considera en estos casos que el concepto en cuestión no se aplica, entonces se tiene una intuición a favor de rechazar dicha caracterización (nuevamente aclaro que no tiene que ser un análisis en sentido riguroso). Así, las distintas caracterizaciones compiten con base en las consecuencias intuitivas y contraintuitivas que ofrece cada una.

Un ejemplo conocido de este tipo de estrategia es el que realiza Goldman en *Epistemology and cognition* (1988), al escoger como criterio de corrección la fiabilidad, pero afirma explícitamente que el modo de escoger su criterio es a través del ERC. En trabajos posteriores (2007), Goldman mantiene la misma estructura de trabajo, pero se compromete con una explicación naturalizada del ER, lo cual será el tema del siguiente apartado.

La necesidad de dar una explicación psicológica de cómo procede el análisis conceptual se debe a que el ERC es demasiado abstracto. Me parece que en cierta medida está basado en una confianza ciega en que el análisis conceptual tradicional es un método viable. Este optimismo excesivo es por supuesto inadecuado. Las alternativas más notorias son el neo-racionalismo de autores como Peacocke (2004), y una explicación naturalista de cómo es posible la exploración conceptual por medio de intuiciones. La vía que exploro y defiendo como válida es la segunda.

## **El equilibrio reflexivo naturalizado**

A muy grandes rasgos, la justificación de las reglas epistémicas puede realizarse a priori o a posteriori. Al momento de evaluar un método bajo estos términos conviene distinguir dos

niveles de preguntas. En primer lugar, podemos ver si metodológicamente la epistemología se llevará a cabo de manera tradicional (o de sillón), reflexionando sobre las propias intuiciones en primera persona; o utilizando datos más allá de las intuiciones propias. En segundo lugar, es posible preguntarse de qué manera se garantiza que este método resulte válido, o dicho de otra manera, cómo se justifica el método; qué es lo que garantiza que el modo como estamos llevando a cabo la investigación dé resultados correctos. Tenemos entonces estas posibilidades para hacer epistemología a partir de los conceptos epistémicos:

- a) Método tradicional justificado a priori.
- b) Método tradicional justificado a posteriori.
- c) Método empírico justificado a posteriori.

El ERC es un método tradicional, debido a que no requiere de datos empíricos para llevarse a cabo. ¿Pero cómo es su justificación? Una alternativa sería realizarla dentro del proyecto neo-racionalista, propuesto por teóricos como Bonjour, Bealer, y otros. Por ejemplo, Peacocke (2004), señala lo siguiente acerca de un concepto epistémico muy relevante, el conocimiento: “El racionalismo moderado busca explicar los casos de conocimiento a priori apelando a la naturaleza de los conceptos que tienen los contenidos que son conocidos a priori” (p. 199). Este tipo de posturas neo-racionalistas cuestionan la opinión común en la filosofía (Cfr. Bates: 2005) de que una vez borrada la distinción analítico-sintético, una investigación conceptual de este tipo puede llevarse a cabo sin una explicación psicológica ulterior.

La segunda alternativa es explicar a posteriori los mecanismos del ER, aún si éste es un método tradicional. Es decir, mantener la estrategia tradicional de hacer epistemología de sillón, pero dar un recuento plausible de por qué es posible (desde un punto de vista psicológico) hacerlo. Puesto que la investigación es acerca de un criterio (y no de reglas) a través de conceptos epistémicos, esto puede ser considerado una postura naturalista sobre el análisis (Barceló: en prensa), sin que se trate necesariamente de defender la metodología naturalista en epistemología.

La tercera alternativa sugiere abandonar la epistemología de sillón por completo. ¿Cómo se

sustituye ésta? Hay dos posibilidades a grandes rasgos. La primera es adoptar un naturalismo, que puede ser eliminativista (Quine: 1969) o bien no-eliminativista (Kornblith: 2005) con respecto de las nociones epistémicas (justificación, conocimiento, creencia, racionalidad). Empero, la alternativa que será evaluada aquí, es una postura alejada del naturalismo metodológico, pero aún empírica. Mientras que el naturalismo abandona la relevancia de las intuiciones epistémicas al buscar, o bien en el caso eliminativista estudiar objetos como activaciones neuronales, etc., o en el caso no-eliminativista a los objetos tradicionales de la epistemología pero con métodos de las ciencias naturales (e.g. Kornblith sugiere estudiar al conocimiento humano tomando elementos de la etología cognitiva), es posible hacer epistemología usando una metodología empírica apelando a las intuiciones epistémicas. En este caso, la herramienta que se puede usar es el análisis estadístico de cuestionarios psicológicos sería una herramienta muy valiosa para conocer las intuiciones epistémicas de una población, lo cual ha sido llamado experimentalismo moderado (Liao: 2008) en otro contexto.<sup>33</sup>

Llamaré a la segunda y tercera alternativa equilibrio reflexivo naturalizado (ERN). Sobre esto me concentraré a continuación, dejando pendiente para una futura investigación la viabilidad de las metodologías tradicionales justificadas a priori. Cabe señalar, que el ERC tal como se utiliza en los años 80 y 90, no se decanta por ninguna de estas dos interpretaciones, sino que se asume que la investigación conceptual es posible sin especificar cómo.

El resto de la sección tratará acerca de este ERN, primero con una breve descripción, y después con un cuestionamiento de la viabilidad del proyecto. Si bien no defenderé el ERN frente a una justificación a priori (la primera alternativa), cabe mencionar que dos motivaciones buenas para ello pueden ser las teorías externistas del significado y el cuestionamiento de Quine a la distinción analítico-sintético (Bates: 2005).

¿Cómo es posible el ERN a priori? Diversos teóricos han abandonado la metodología a priori “misteriosa” (Bates: 2005) del proyecto analítico en epistemología para sustituirla por un método hipotético deductivo que realiza el mismo trabajo. En esta versión más reciente,

---

<sup>33</sup> Como se verá a continuación, a la práctica de explorar empíricamente las intuiciones de una población se le llama filosofía experimental.

Goldman parte de una noción de conceptos que llaman “mentalismo” donde el concepto es un estado psicológico real que permite a alguien utilizar un término (Nagel: 2007). Los conceptos son “estructuras psicológicas que juegan un rol causal en el comportamiento lingüístico” (Bates: 2005). Las intuiciones son una fuente básica de evidencia empírica derrotable acerca de la posesión de estos conceptos. No es necesario sostener que las intuiciones o la disposición a aplicar los conceptos tenga un papel constitutivo, basta con que resulten una evidencia lo suficientemente buena.

En este caso, los conceptos que se investigan son epistémicos, concretamente aquello que hace que una regla de razonamiento sea correcta. Un término usual para investigar este tipo de criterios es el de “justificación”, pero también directamente el término “racional”.

Tal como propongo en este trabajo, los conceptos epistémicos son interesantes desde un punto de vista epistemológico por los bienes epistémicos que llevan implícitos, los cuales son la fuente de la normatividad epistémica. Por ello, suponiendo que una investigación realizada dentro de este marco fuera sobre el término “justificación,” esto no implicaría que el objetivo fuera realizar un análisis de dicho término. Considero que esta manera de proceder sigue en espíritu a Goldman, pues si bien propone al fiabilismo como un análisis del término “justificación” (Cfr. 1979), cuando contextualiza esto en un trabajo más amplio (Cfr. 1984) la importancia radica en encontrar un criterio de corrección.

El ERN consiste en juntar casos de intuiciones que ilustran el uso ordinario de un concepto para posteriormente generar una hipótesis que los cubra. La hipótesis puede ser descartada debido a que no cubre las intuiciones; pero al igual que en el método hipotético-deductivo, algunas intuiciones pueden ser descartadas por utilidad teórica, como sucede en investigaciones de tipo empírico con algunos resultados experimentales que son considerados anomalías. Es por ello que el equilibrio reflexivo resulta un MI ventajoso para dar una explicación naturalista de un método basado en intuiciones. O dicho de otra manera, el equilibrio reflexivo es un modo adecuado de describir esta metodología.

Las intuiciones se convierten en datos empíricos precisos cuando son definidas como disposiciones a aceptar la aplicación de un concepto como correcta. A partir de una serie de

situaciones planteada, un sujeto tiene que decidir si a éstas podrían aplicarse los conceptos epistémicos que busquen investigar si una regla o un criterio para justificar una regla es adecuado (posiblemente bajo el concepto de justificación).

Este método hipotético-deductivo se parece a la filosofía tradicional cuando se practica desde la primera persona. Es decir, cuando el epistemólogo tradicional realiza una investigación de sillón acerca de las situaciones en que aplicaría los conceptos epistémicos cotidianos para conocer sus propios conceptos epistémicos. Los contraejemplos sirven como evidencia en contra de las hipótesis.

El hecho de que dos sujetos reaccionen de manera distinta ante una aplicación de un término, es evidencia a favor de que en realidad poseen distintos conceptos (Stich, Nichols y Weinberg: 2003). Por ejemplo, si el sujeto 1 dice que el oficinista Smith del caso Gettier (Cfr. Gettier: 1963) no sabe que el hombre que tendrá el trabajo tiene 10 monedas en su bolso, y el sujeto 2 dice que Smith sí lo sabe, esto (descontando errores de ejecución, distracciones, etc.) indica que el sujeto 1 y el sujeto 2 poseen distintos conceptos de conocimiento. Esto es un compromiso teórico con respecto a cómo se da la individuación de conceptos psicológicos. Stich, Nichols y Weinberg (2003) hacen notar a manera de ejemplo que esto no sucedería con la teoría de individuación de Fodor.

Las intuiciones no necesariamente son constitutivas de los conceptos, pero sí tienen una relación evidencial. De acuerdo con Goldman (2007), poseer un concepto hace que “uno tenga la disposición a tener intuiciones a favor en casos de aplicación correcta e intuiciones en contra en casos de aplicación incorrecta” (p. 15). Con ello se puede mantener la idea de que las intuiciones son buena evidencia para saber qué conceptos tiene un sujeto. Posteriormente distingue una investigación en tercera persona de una investigación en primera persona. El análisis en tercera persona sería una actividad de tipo empírico (ERN con un metodología a posteriori), mientras que el análisis en primera persona sería similar al análisis tradicional pero con una justificación naturalista (ERN con metodología tradicional).

¿De quién deben ser las intuiciones que se exploran? En este punto también existe desacuerdo. Una vez que adoptamos la visión psicologista (o mentalista) expuesta aquí, la

consecuencia es que las intuiciones deben ser de cualquier sujeto, lo que comúnmente se denomina como "folk." Como indica Goldman (2007, p. 18): "El análisis filosófico se ocupa principalmente de conceptos comunes [...] En este terreno, no hay ninguna división experto/novato. Así, en caso de que existan diferencias asociadas con una sola palabra, no pueden resolverse apelando a expertos." Jackson (1998) en cambio no considera que la investigación debe ser acerca de conceptos, sino de significados. Términos como "justificación" o "conocimiento", tienen un significado público correcto, que muchos individuos pueden no haber comprendido en su totalidad. Supongamos que Jackson tiene razón y que existe un individuo con intuiciones epistémicas erradas. Supongamos por un momento que las intuiciones "correctas" con respecto a los casos Gettier son las que no atribuyen conocimiento. Esta persona considera que Smith, en el caso Gettier, sabe que la persona que conseguirá el trabajo tiene diez monedas en su bolso. También esta persona muestra intuiciones similares en otros casos Gettier, y en general en casos que involucran suerte epistémica. El individuo, desde un punto de vista como el de Jackson, tiene un concepto (personal, psicológico) de conocimiento distinto al de los sujetos que no comparten dichas intuiciones. En esto no hay una diferencia con Goldman. Sin embargo de acuerdo con la metodología que propone Jackson, puesto que el objetivo de la investigación es el significado del término "conocimiento", lo que cabe decir es que el concepto psicológico que posee el individuo con intuiciones erradas no es el adecuado (y por ello no podrá hacer análisis conceptual en primera persona).

De acuerdo con esta postura que comparten Goldman y también Weinberg y Alexander, no existe ninguna razón para que las intuiciones deban proceder de filósofos profesionales. De hecho "los filósofos están especialmente equipados para atender los aspectos importantes de los experimentos mentales y de los conceptos involucrados en las aserciones filosóficas. Así, las intuiciones de los filósofos, imbuidas con un cierto tipo de pericia, tienen un valor evidencial mayor que el de los no-filósofos" (Weinberg y Alexander: 2007, p. 59). Si se considera que las intuiciones de los filósofos deben ser consultadas, al menos para dirimir diferencias entre las intuiciones de la población general, entonces es necesario dar buenas razones para ello.



Jackson no da razones suficientes para mostrar que las intuiciones sobre términos filosóficos deban variar (Goldman: 2007). En el caso del agua, el significado público correcto de agua requiere ciertos conocimientos de química o incluso física de partículas, que el experto debe conocer. El concepto psicológico de agua de este experto coincidirá con el significado correcto del término “agua.” Esta es la descripción de la división del trabajo lingüístico que propuso Putnam (1975) con gran aceptación. Pero en el caso de la epistemología normativa, no existen conocimientos análogos a los de la física o química. ¿Qué otra buena razón podría haber para dividir el trabajo lingüístico? La otra razón aducida tendrá que ver con la cantidad de reflexión que requieren las intuiciones para ser un insumo adecuado. Mencionaré esto más adelante. Lo que resulta claro es que la división del trabajo lingüístico obedece a ciertas exigencias, como en el caso del agua (no todos pueden estar especializados en química), y no debe invocarse gratuitamente.

La comparación entre un método hipotético deductivo y el ERN no se debe confundir con la idea de que el ER refleja la práctica científica, como hace notar Cummins (1998) y en cierto modo Kornblith (2005). Cummins señala en un artículo referido en tono crítico: “Este tipo de ajuste mutuo entre teoría y datos es una característica familiar de la práctica científica. Cualquier autoridad que el ER parezca tener proviene, creo, de un reconocimiento tácito o implícito de que tiene la misma forma que este tipo de inferencia científica” (1998, p. 114). El procedimiento no sólo es similar a un modo prominente en el que la ciencia experimental se lleva a cabo, como hace notar Cummins, sino que son de hecho hipótesis empíricas acerca de las estructuras (en este caso conceptos) que posee un sujeto. Así, como indica Goldman, la justificación a priori “proviene de, sobreviene en, los procesos psicológicos responsables para la creencia”<sup>34</sup> (2007, p. 19) Se trata de una auténtica justificación psicológica de la metodología basada en intuiciones, y no sólo de una analogía superficial con la metodología científica.

Por ejemplo, un trabajo de Nichols, Stich y Weinberg muestra evidencia utilizando epistemología experimental (la realización de encuestas sobre la aplicación de conceptos

---

<sup>34</sup> Por ejemplo, nuestras capacidades aritméticas por el modo como funcionan proveen justificación a priori.

epistémicos) realizando un estudio de a diversas poblaciones, divididas por etnicidad y nivel socioeconómico. A los sujetos se les muestra una situación epistémica diseñada por los experimentadores y deben contestar si atribuyen conocimiento. Las situaciones están diseñadas para explotar las diferencias culturales de ambos grupos. Las hipótesis a probar son las siguientes:

- 1- Las intuiciones epistémicas varían en cada cultura.
- 2- Las intuiciones epistémicas varían de un contexto socioeconómico a otro.
- 3- Las intuiciones epistémicas varían según los cursos de filosofía que ha tomado una persona.

La conclusión del estudio es que existen diferencias importantes en cuanto a la atribución de conocimiento de los distintos grupos. Por ejemplo, un experimento es presentar un contraejemplo Gettier. El éxito de los contraejemplos Gettier ha sido algo poco cuestionado en epistemología. En este tipo de casos, se da un ejemplo en que un sujeto tiene una creencia verdadera y justificada, pero donde la verdad de la creencia no tiene que ver con la justificación. Uno de los resultados experimentales más sorprendentes es que cuando se enfrentan con un relato de este tipo, los sujetos provenientes de Asia oriental atribuyen conocimiento en un porcentaje significativamente mayor al personaje del caso Gettier que los sujetos occidentales.

En la parte interpretativa del estudio, las diferencias encontradas entre occidentales y orientales se atribuyen al hecho de que los occidentales le dan mayor importancia a la causalidad, mientras que los asiáticos orientales tienden a valorar más la autoridad comunitaria. De acuerdo con la perspectiva adoptada en esta investigación, el resultado indica que ambos grupos poseen distintos conceptos. Por tanto, una posible alternativa para continuar con una investigación conceptual basada en intuiciones, sería adoptar un pluralismo epistemológico. Los estándares normativos para distintos grupos poblacionales

son variables. Esto no es incompatible con posturas acerca de la racionalidad sostenidas antes de que existieran estos resultados (e.g. Eraña: 2002). Sin embargo los siguientes resultados que reseñaré, y lo que considero que son interpretaciones correctas de sus autores, muestran de manera definitiva que la investigación conceptual de sillón justificada a priori no es una vía disponible.

## **Equilibrio reflexivo naturalizado informado y experimentalismo moderado**

Comenzaré este apartado realizando una distinción entre ERN ingenuo e informado. Posteriormente mostraré que el ERN informado y la postura denominada experimentalismo moderado pueden resistir las críticas realizadas desde la filosofía experimental, y al mismo tiempo mantener los beneficios explicativos de la teoría de la racionalidad defendida en este trabajo en cuanto a una explicación de la normatividad.

Propongo llamar metodología tradicional ingenua, a aquella que supone *prima facie* que es posible investigar todos los conceptos a través de intuiciones. En contraste, una metodología tradicional informada procede a partir de los conceptos e intuiciones que la filosofía experimental (u otro medio de calibrar intuiciones) ha mostrado ser lo suficientemente compartidos, estables y confiables. El optimismo con respecto a la epistemología tradicional, supone que las anomalías no son lo suficientemente importantes como para que no sea posible continuar la investigación conceptual con experimentos mentales en primera persona. Las razones que aduciré a continuación muestran que hay suficientes razones para pensar que el ERN no puede ser una metodología tradicional ingenua. El hecho de que las divergencias en intuiciones hayan sido encontradas en casos en los que había consenso entre los filósofos profesionales, es una prueba de la dificultad de saber de antemano cuáles son los conceptos compartidos. El caso Truetemp, el experimento mental del sujeto que sabe automáticamente la temperatura en ciertas condiciones, resulta ilustrativo (Weinberg y Alexander: 2007). El caso fue propuesto en contra de los análisis

externistas de la justificación.

Como mencioné anteriormente, las siguientes razones impiden que se pueda proseguir una MI tradicional mezclada con un pluralismo epistémico. El trasfondo socioeconómico resulta relevante. Al parecer los individuos que pertenecen a estratos más altos, tienen estándares epistémicos más estrictos, lo cual se refleja estadísticamente sus intuiciones. Se trata de un relativismo poco sustentable: “La versión del relativismo según la cual las [estrategias basadas en intuiciones] llevarían a implicar que las normas epistémicas apropiadas para los ricos son distintas de las normas epistémicas apropiadas para los pobres” (Stich, Nichols y Weinberg: 2001). Mientras que las diferencias en intuiciones epistémicas por motivos étnicos podrían ser el principio de una teoría epistemológica compatible con el relativismo cultural, un relativismo socioeconómico resulta algo incompatible con la epistemología normativa.

La posición que le atribuye un papel más problemático a estos resultados para las MI puede llamarse “experimentalismo radical.” La siguiente formulación de la postura es clara y concisa: “Los experimentalistas radicales creen que debemos abandonar la práctica filosófica estándar de utilizar intuiciones como evidencia porque sus estudios empíricos han mostrado que las intuiciones varían con factores como la educación y el trasfondo cultural” (Liao: 2008, p. 23).

Los epistemólogos experimentales han adaptado sus pruebas e interpretaciones para responder a una diversa gama de objeciones. Por ejemplo, ante la objeción de que las diferencias podrían ser sólo en grado de fuerza, Swain, Alexander y Weinberg (2008) han incluido gradaciones en las pruebas que muestran que en casos importantes las diferencias van más allá del grado (Liao: 2008).

Además de estas objeciones, existen otras que son incompatibles desde la perspectiva llevada a cabo en este trabajo. Por ejemplo, Williamson (2004) arguye que las intuiciones son más que evidencia psicológica. Las intuiciones son evidencia sobre ciertos hechos modales, así como la percepción provee evidencia sobre ciertos hechos que trasciende la evidencia psicológica. La perspectiva que adopto en este trabajo utiliza una interpretación psicológica de la evidencia intuitiva, misma que rechaza Williamson.

Otra objeción a la interpretación de los experimentalistas radicales es aquella que señala que las pruebas experimentales no examinan el tipo de intuiciones adecuadas. Según esta postura las intuiciones deben ser robustas, es decir, producto de un largo proceso de discusión. Una vez más, esta línea de respuesta no está disponible para lo que he propuesto en esta investigación, debido a los presupuestos acerca de la individuación de conceptos que acepté con anterioridad, lo cual se encuentra en el corazón del ERN. De acuerdo con esta postura (Cfr. Goldman: 2007), la diferencia en la disposición aplicarlo necesariamente implica que se posee un concepto distinto. Si el filósofo responde de manera diferente a los casos Gettier, significa que posee un concepto distinto de conocimiento. Las posturas son las siguientes:

- a) Epistemología tradicional.
- b) Experimentalismo moderado.
- c) Experimentalismo radical.

Quienes tienen una inclinación por dar una explicación parecida a lo que llamo ERN, incluso cuando consideran que el proceso de análisis podría ser relativamente autónomo piensan que verificaciones experimentales ocasionales pueden ser útiles o benéficas. La distinción entre ERN ingenuo e informado no tiene que ser tajante, pues puede variar según el grado de información que se considere necesario o conveniente. Algunos defensores del ERN con metodología tradicional como Goldman (2007) y Jackson (1998), no han tenido interés por una investigación a posteriori como la del experimentalismo moderado, pero consideran que es importante calibrar las intuiciones o corregirlas con base a este tipo de evidencia.

La clasificación entre experimentalismo radical y experimentalismo moderado es de Liao (2008). Originalmente, se trata de una clasificación acerca de las posturas de quienes

realizan estudios de filosofía experimental y su postura con respecto a los resultados. Sin embargo, me parece útil tomarlos como términos para describir una postura ante el uso o valor de la experimentación con respecto a las intuiciones. Estas posturas son implícitas, y muchas veces quienes han realizado estudios de filosofía experimental no mantienen una sola consistente (Gasparatou: 2010).

El experimentalismo radical considera que resultados como los anteriormente mencionados tienen como consecuencias que ya no es posible sostener una meta-filosofía donde las intuiciones jueguen un papel primordial. En cambio los experimentalistas moderados, consideran que las intuiciones tienen un interés filosófico, y que la filosofía experimental es necesaria (o al menos útil) para examinarlas.

Algunos teóricos que se presentan optimistas respecto a los prospectos de mantener la continuidad con la filosofía tradicional, como Liao (2008) y Goldman (2007), consideran que la filosofía experimental puede tener un valor para revisar la estabilidad de los conceptos, para resolver conflictos de intuiciones, y para detectar errores como “la parcialidad, emociones obnubiladas y heurísticas incorrectas” (Liao: 2008, p. 19). Debido a los fenómenos anteriormente expuestos, y a que los experimentos han examinado casos en el que existía consenso éxito de estos esfuerzos ha sido muy limitado.

Pero existe otra manera de utilizar a las intuiciones evaluadas por medio de pruebas. Esto es compatible con lo que describí con utilizar el ERN como método posteriori. Existen defensores del experimentalismo moderado que a pesar de desconfiar de los métodos de sillón de los filósofos tradicionales “parecen creer que las intuiciones de sentido común son filosóficamente relevantes y que pueden ser usadas para promover una tesis” (Gasparatou: 2010, p. 32). De acuerdo con esta concepción del experimentalismo, un argumento en favor de un criterio de corrección determinado podría ser el siguiente:

1- Existe evidencia experimental de que los sujetos, o cierto grupo de sujetos, tienen la intuición de que el criterio C-R es correcto (se pueden usar distintos conceptos epistémicos para investigar la corrección del criterio).

2- Por tanto, hay razones *prima facie* para pensar que el criterio C-R es correcto.

El problema con este tipo de metodología, según Gasparatou (2010, p. 47), es que los experimentalistas moderados se encuentran en un dilema, pues sostienen que las intuiciones son poco confiables y al mismo tiempo que la investigación filosófica depende sustancialmente de ellas.

Uno de mis objetivos ha sido mostrar cómo tiene sentido realizar una investigación a posteriori con respecto a las intuiciones epistémicas para buscar un criterio de corrección. Una consecuencia adicional es que si se encuentra una estabilidad como la que pide Goldman (2007) en ciertas condiciones o ciertos conceptos, también será posible llevar a cabo una metodología de equilibrio reflexivo.

Me parece que hay tres cosas que vale la pena señalar. En primer lugar, el papel general que deben jugar las intuiciones en la teoría. Graham y Horgan (1994) ofrecen una visión interesante acerca de cómo la teoría debe dar cuenta de las intuiciones, aún en el caso de que sean erradas. Por ejemplo, en el caso de la ilusión de Mueller-Lyer, existe una intuición de que una de las dos flechas es más grande. Esta intuición no es cognitivamente penetrable, es decir, no desaparece aún sabiendo que son del mismo tamaño (o incluso después de medirlas). Sin embargo, la intuición no se desecha simplemente, sino que es necesario explicar cómo se produce y por qué.

La teoría final que debe acomodar las intuiciones incluye todos los datos relevantes y no sólo los datos epistémicos, de modo que se trata de algo parecido a un ER amplio, si es que cabe hacer esta distinción en este contexto. Como admiten los autores, lo que proponen es una propuesta metafilosófica en sentido amplio. El ER aquí no es un método concreto de la epistemología, sino que es la forma general como se debe realizar cualquier teoría filosófica. Como los autores señalan, este modo de acomodar los datos en una teoría es distinto que la típica metodología a posteriori (p. 219). Por tanto, aún si el experimentalismo moderado es la única manera de estudiar los conceptos epistémicos, no desaparece la necesidad de equilibrar las intuiciones. El ejemplo de Graham y Horgan es ilustrativo en este punto, la ilusión de Mueller-Lyer se equilibra al momento de realizar mediciones exactas de las líneas

y de dar una explicación psicológica acerca de por qué se produce la ilusión. Esto va más allá del método hipotético deductivo para proponer y defender una teoría científica. Después de todo, no es necesario explicar el origen cognitivo de la ilusión Mueller-Lyer para saber que las líneas son del mismo tamaño, sino que basta con medirlas. El requisito de acomodar todas las intuiciones es más fuerte, más que hacer uso de las intuiciones, exige encontrar un lugar para las mismas.

En segundo lugar, las presuposiciones adoptadas en el segundo capítulo de este trabajo tienen como consecuencia que los conceptos epistémicos son relevantes, aún en el caso de que no sean universalmente compartidos. Esto no exime de realizar un estudio más complejo de conjuntos de conceptos que quizá tendrán un parentesco de familia. Para ello puede ser útil estudiar intuiciones más modestas, o incluso sensaciones más básicas epistémicamente relevantes, como la sensación de que se sabe algo, como propone Jennifer Nagel (2007), quien revisa la literatura sobre la sensación de que se sabe [feeling of knowing], que resulta ser más estable que una epistemología realizada a un nivel más elevado. Por otra parte, nada de lo propuesto aquí exige que los conceptos epistémicos sean universales, ni que se pueda derivar una normatividad universalmente aplicable a partir de los mismos. La racionalidad mínima es neutral con respecto a este punto, al igual que la teoría mentalista sobre individuación de conceptos e intuiciones como evidencia de los mismos. Si la filosofía experimental provee evidencia de que los conceptos epistémicos no son compartidos es posible adoptar el pluralismo epistémico sin abandonar las tesis principales defendidas aquí. En lugar de leer estos resultados como una refutación de que los conceptos epistémicos puedan resolver el problema de la racionalidad, es posible interpretar esta evidencia como apoyo de que la normatividad epistémica es algo que surge en el contexto de una comunidad, como se ve en algunas propuestas de distintos tipos (Eraña: 2002; Kusch: 2002). Si la postura consecuencialista defendida en este trabajo es correcta, el debate sobre cuáles son los conceptos (o sensaciones básicas) adecuados estará relacionado con los bienes epistémicos (e.g. número de verdades obtenidas, proporción de verdades sobre errores, etc.). La circularidad aquí no es viciosa, pues es posible comenzar con algunos conceptos epistémicos que plausiblemente puedan ser importantes, y a partir de analizar los valores



que contienen, utilizar esta información para saber qué conceptos epistémicos puedan ser más relevantes. Esto no resulta nuevo en la práctica de la epistemología tradicional, donde resultaba común empezar investigando un concepto usual (como la justificación), y encontrar que otra noción resulta más adecuada o importante (la fiabilidad, la agencia virtuosa).

Los conceptos epistémicos son probablemente la única manera de acercarnos a las nociones epistémicas, ya que las posturas meta-epistemológicas pragmatistas (Stich: 1994) o naturalistas (Kornblith: 2005) no resultan adecuadas. Las posturas teleológicas, basadas en los bienes o fines epistémicos, logran dar una respuesta adecuada al problema de la normatividad epistémica. ¿Cómo estudiar a los conceptos a través de intuiciones una vez que existe inestabilidad de las intuiciones, y al mismo tiempo evidencia de que los conceptos no son compartidos del modo en que pensaban los filósofos tradicionales? El experimentalista moderado debe encontrar estas divergencias, dividir en grupos poblacionales cuando sea relevante, y encontrar cuáles son concretamente los valores epistémicos implícitos en los conceptos epistémicos. En caso de que exista confirmación empírica de que los conceptos epistémicos relevantes son socialmente compartidos, es posible adoptar lo que llamé una metodología tradicional informada, que puede ser permanentemente mejorada y corregida con más evidencia experimental. Así, el experimentalismo moderado es compatible y se complementa con una metodología de sillón, siempre que se reconozca que ésta depende de una justificación a posteriori acerca de los conceptos, y de una teoría de individuación de conceptos basada como la adoptada por Goldman (2009).

## 5. CONCLUSIÓN

A pesar de que la racionalidad es una noción compleja y difícil de definir, se trazó aquí una estrategia que se puede seguir para establecer tanto de manera general en qué consiste. La propuesta es considerar a la racionalidad como una relación inferencial entre distintas creencias por medio de razonamientos correctos. Para responder esto, se sugiere una postura meta-epistemológica consecuencialista, donde los conceptos epistémicos son los encargados de proveer las metas epistémicas. Finalmente, se ofrece una metodología concreta para llevar a cabo el análisis de estos conceptos: el equilibrio reflexivo naturalizado, cuya aplicación requiere de apoyarse en la epistemología experimental.

De este modo, el presente trabajo cumple el objetivo de darle tratamiento a una problemática clara en relación a un programa de investigación acerca de la racionalidad, y de mostrar cómo se pueden resolver las cuestiones más importantes. Sólo hasta la tercera parte del trabajo se llega a una propuesta metodológica concreta. Esto se debe a que fue necesario establecer soluciones a los distintos problemas epistemológicos y meta-epistemológicos que conlleva el estudio de la racionalidad.

Sin embargo, este paso previo es de enorme importancia para comprender la racionalidad epistémica. Gracias al modo como se llevó a cabo este estudio es posible garantizar que la metodología propuesta puede ser satisfactoria primero, desde el punto de vista de la agencia, y en segundo lugar, desde el punto de vista de la normatividad epistémica. De este modo, se puede ver claramente que el equilibrio reflexivo naturalizado resulta un modo adecuado de investigar acerca de la racionalidad.

En cierto sentido, una parte muy importante del programa de trabajo consiste en retomar los métodos tradicionales de la epistemología, pero al mismo tiempo, buscando dar una explicación de fondo acerca de cómo funcionan y de dónde procede su normatividad. Más que regresar a una epistemología a priori ingenua, la conclusión implica concebir desde una nueva perspectiva el análisis conceptual.

Teniendo en cuenta todos los problemas que resultan por adoptar métodos basados en

intuiciones, resulta sorprendente que apenas se haya cuestionado esta metodología durante los primeros años de la epistemología analítica, cuando surgieron las discusiones clásicas acerca del análisis del conocimiento (como respuesta a los contraejemplos Gettier) y de la fuente de la justificación.

Sin embargo, el estudio de los conceptos epistémicos aún puede ser un método viable para resolver la pregunta urgente acerca de qué es lo que constituye un razonamiento adecuado. Los programas de investigación de corte naturalista con frecuencia han evaluado los métodos epistemológicos con base a virtudes como su fiabilidad o su frugalidad computacional, pero sin dar cuenta de los criterios utilizados. Mientras estos criterios no se hagan explícitos y no se le dé la atención adecuada a sus fundamentos teóricos y metodológicos, el naturalismo epistemológico no puede apelar a ellos. Lo anterior no significa que el análisis conceptual sea la única manera de llevar a cabo esta labor, por ejemplo, Kornblith (2002) ha propuesto justificar los criterios naturalistas de fiabilidad a partir de un marco evolutivo. Sin embargo, el análisis conceptual es la mejor alternativa para satisfacer los criterios de normatividad epistémica.

Así, las intuiciones epistémicas son una manera de responder a estas necesidades, dando cuenta de manera crítica tanto a los distintos cuestionamientos acerca de su pertinencia como la fuente para resolver el problema de un criterio de corrección. Dentro de los métodos basados en intuiciones, el equilibrio reflexivo resulta interesante por dos motivos principalmente. El primero de ellos, ya notado por Goodman, es que parece conformarse con nuestra práctica filosófica. El segundo, aquí resaltado, es que permite sostener algunos resultados contraintuitivos de una manera razonable.

El equilibrio reflexivo naturalizado combina las virtudes del equilibrio reflexivo clásico, con una justificación concreta acerca de por qué funciona. Asimismo, debido a que parte de una noción psicológica de concepto tiene una ventaja respecto a un proyecto racionalista. Mientras que la relación entre las intuiciones y los conceptos en sentido fregeano es poco sustentable, tal parece ser que las intuiciones epistémicas sí tienen una relación estrecha con los conceptos psicológicos. Sin embargo, debido a la explicación mentalista de los conceptos adoptada en este trabajo, el equilibrio reflexivo naturalizado requiere ser validado por

evidencia empírica de que los conceptos son socialmente compartidos.

El experimentalismo moderado, la idea de que los métodos estadísticos sobre intuiciones epistémicas pueden ayudar a estudiar los conceptos epistémicos, tiene una doble función. En primer lugar, puede ayudar a validar, corregir y suplementar al equilibrio reflexivo naturalizado. Pero en caso de que no existan las condiciones para realizar epistemología a priori en determinados temas, o incluso de manera global, el experimentalismo moderado provee un sustituto adecuado. La epistemología experimental puede explorar los conceptos epistémicos de una manera adecuada, y en una forma que satisface los presupuestos básicos del equilibrio reflexivo, que consiste en balancear y acomodar las intuiciones.

Queda también como posibilidad abierta un análogo a posteriori a la epistemología de sillón por medio de intuiciones informadas; es decir, cuando la persona que analiza en primera persona tiene evidencia empírica para pensar que sus intuiciones *prima facie* son estables y van a coincidir lo suficiente con las intuiciones comunes. En este escenario, un caso límite de experimentalismo moderado donde las encuestas sólo se usan para calibrar intuiciones, el cambio fundamental no está en la metodología sino en la justificación de la misma, que deja de ser a priori.

Ambos usos de la epistemología experimental dejan un programa de investigación abierto, que no sólo es capaz de responder a los resultados disponibles, sino que se puede hacer valer de ella de manera importante para encontrar criterios epistémicos adecuados. Estos criterios a su vez, son capaces de proveer criterios acerca de cuáles son las reglas normativamente adecuadas para los agentes racionales. Así, este trabajo cumple con los objetivos propuestos y sugiere el camino que deben seguir futuras investigaciones sobre el tema.

## 6. BIBLIOGRAFÍA

- K. Ahlstrom (2008). 'Epistemology and Empirical Investigation'. *Grazer Philosophische Studien* 76(1): 109–134.
- J. Alexander & J. Weinberg (2007). 'Analytic epistemology and experimental philosophy'. *Philosophy Compass* 2(1).
- J. Alexander (2010). 'Is Experimental Philosophy Philosophically Significant?'. *Philosophical Psychology*, 23 (3).
- W. Alston (1986). 'Epistemic Circularity'. *Philosophy and Phenomenological Research* 47(1): 1–30.
- D. Armstrong (1973). *Belief, Truth and Knowledge*. Cambridge University Press.
- A. Barceló (en prensa). *Análisis y estructura*.
- J. Bates (2004). 'Reflective equilibrium and underdetermination in epistemology'. *Acta Analytica* 19(32): 45–64.
- J. Bates (2005). 'The Old Problem of Induction and the New Reflective Equilibrium'. *dialectica* 59(3): 347–356.
- G. Bealer (2002). 'Modal Epistemology and the Rationalist Renaissance', en *Conceivability and Possibility*, pp. 71–125. Oxford University Press.
- J. L. Bermúdez & A. Millar (eds). (2003). *Reason and Nature: Essays in the Theory of Rationality*. Oxford University Press.
- P. Boghossian (2001). 'How are objective epistemic reasons possible?'. *Philosophical Studies* 106(1-2).
- P. Boghossian (2003). *Objective epistemic reasons*. Oxford University Press.
- L. BonJour (1998). *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge University Press.

- R. Brandom (2000). *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press.
- T. Burge (2003). 'Perceptual Entitlement'. *Philosophy and Phenomenological Research* 67(3): 503–548.
- C. Cherniak (1990). *Minimal Rationality*. MIT Press.
- (2004). *The Oxford Handbook of Rationality*. A.A.V.V. Oxford University Press.
- L. J. Cohen (1981). 'Can human irrationality be experimentally demonstrated?'. *Behavioral and Brain Sciences* 4: 317–370.
- L. J. Cohen (1989). *The Dialogue of Reason: An Analysis of Analytical Philosophy*. Oxford University Press.
- M. Colyvan (2009). 'Naturalising Normativity', en D. Braddon-Mitchell & R. Nola (eds)., *Conceptual Analysis and Philosophical Naturalism*. MIT Press.
- R. Cummins (1998). 'Reflections on Reflective Equilibrium', en W. Ramsey & M. R. DePaul (eds)., *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Rowman and Littlefield Publishers.
- N. Daniels (1996). *Justice and Justification: Reflective Equilibrium in Theory and Practice*. Cambridge University Press.
- D. Dennett (1981). *Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press.
- D. Dennett (1989). 'True Believers', en *The Intentional Stance*. Bradford.
- D. Dennett (1998). 'Can machines think?', en *Brainchildren: Essays on Designing Minds*. Bradford.
- C. Z. Elgin (1999). *Considered Judgment*. Princeton University Press.
- P. Engel (2001). 'The norms of thought: Are they social?' *Mind & Society*. 2(1): 129, 148.

- A. Eraña (2003). 'Normatividad epistémica y estructura heurística del razonamiento'. *Crítica, Revista Hispanoamericana de Filosofía*. 35(10): 69–108.
- R. Feldman (1988). 'Rationality, Reliability, and Natural Selection'. *Philosophy of Science* 55(2): 218.
- C. L. García (1999 ). 'Razonamiento heurístico y racionalidad'. Manuscrito inédito.
- C. L. García (2004). 'Especificidad de dominio y sesgo en la psicología del razonamiento humano.'. *Signos Filosóficos* 6: 63–92.
- R. Gasparatou (2010). 'Experimental appeals to intuition'. *Crítica, Revista Hispanoamericana de Filosofía* 42(124): 31–50.
- E. Gettier (1963). 'Is Justified True Belief Knowledge?'. *Analysis* 23(6):121.
- G. Gigerenzer (2008). *Rationality for Mortals: How People Cope with Uncertainty*. Oxford University Press.
- A. Goldman (1976). 'Discrimination and Perceptual Knowledge'. *The Journal of Philosophy* 73(20): 771–791.
- A. Goldman (1979). 'What Is Justified Belief?', en Pappas (ed)., *Justification and Knowledge*. Springer.
- A. Goldman (1986). *Epistemology and Cognition*. Harvard University Press.
- A. Goldman (1993). 'Epistemic Folkways and Scientific Epistemology'. *Philosophical Issues* 3: 271–285.
- A. Goldman (2007). 'Philosophical Intuitions: Their Target, Their Source, and Their Epistemic Status'. *Grazer Philosophische Studien* 74(1): 1–26.
- N. Goodman & H. Putnam (1983). *Fact, Fiction, and Forecast, Fourth Edition*. Harvard University Press.
- G. Graham & T. Horgan (1994). 'Southern Fundamentalism and the End of Philosophy'. *Philosophical Issues* 5: 219–247.
- G. Harman (1974). *Thought*. Princeton University Press.

- G. Harman (1984). "Practical aspects of theoretical reasoning." en *The Oxford Handbook of Rationality*. pp. 45-56.
- G. Harman (1988). *Change in View: Principles of Reasoning*. MIT Press.
- G. Harman (1999). *Reasoning, Meaning, and Mind*. Oxford University Press, USA.
- J. Hintikka (1999). 'The Emperor's New Intuitions'. *The Journal of Philosophy* 96(3).
- C. Hooker (1994). 'Idealisation, Naturalism, and Rationality: Some Lessons From Minimal Rationality'. *Synthese* 99(2).
- F. Jackson (1998). *From Metaphysics to Ethics*. Oxford University Press.
- H. Kornblith (1993). 'Epistemic Normativity'. *Synthese* 94(3): 357–376.
- H. Kornblith (2005). *Knowledge and Its Place in Nature*. Oxford University Press.
- S. Kripke (1982). *Wittgenstein on Rules and Private Language*. Harvard University Press.
- S. M. Liao (2008). 'A Defense of Intuitions'. *Philosophical Studies* 140(2): 247–262.
- G. E. Moore (1993). *G.E. Moore: Selected Writings* Routledge.
- R. Morado (2000). *Nuevos paradigmas de inferencia racional.*, pp. 88–99. Plaza y Valdés.
- J. Nagel (2007). 'Epistemic Intuitions'. *Philosophy Compass* 2(6).
- J. Nagel (2012). 'Intuitions and Experiments: A Defense of the Case Method in Epistemology'. *Philosophy and Phenomenological Research*.
- S. Nichols, *et al.* (2003). 'Metaskepticism: Meditations in Ethnoepistemology', en *The Sceptics*. Ashgate.
- C. Peacocke (1998). 'Implicit Conceptions, Understanding and Rationality'. *Philosophical Issues* 9: 43–88.
- C. Peacocke (2004). *The Realm of Reason*. Oxford University Press.
- P. Pettit (2001). 'Deliberative Democracy and the Discursive Dilemma'. *Noûs* 35(s1): 268–299.
- A. Plantinga (1993). *Warrant: The Current Debate*. Oxford University Press.



- J. Pollock (2008). 'Irrationality and cognition', en *Epistemology: New Essays*. Oxford University Press.
- H. Putnam (1975). 'The Meaning of 'Meaning'.' *Minnesota Studies in the Philosophy of Science* 7: 131–193.
- W. V. O. Quine (1969), "Epistemology Naturalized" en *Ontological Relativity and Other Essays*, Columbia University Press.
- J. Rawls (1951). 'Outline of a Decision Procedure for Ethics'. *Philosophical Review* 60(2): 177–197.
- J. Rawls (2005). *A Theory of Justice*. Belknap Press of Harvard University Press.
- G. Seay & S. Nuccetelli (2009). 'Reasoning, Normativity and Experimental Philosophy'. *American Philosophical Quarterly*. 49 (2).
- A. Sen (1977). 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory'. *Philosophy and Public Affairs* 6(4): 317–344.
- E. Sosa (2009). *A Virtue Epistemology*. Oxford University Press.
- D. Sperber (2012) 'Reasoning as social competence' en *Collective Wisdom*, Cambridge University Press, 2012.
- R. C. Stalnaker (1987). *Inquiry*. Bradford.
- E. Stein (1998). *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*. Oxford University Press.
- S. Stich & R. Nisbett (1980). 'Justification and the Psychology of Human Reasoning'. *Philosophy of Science* 47(2): 188–202.
- S. Stich (1988). 'Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity'. *Synthese* 74(3): 391–413.
- S. Stich (1993). *The Fragmentation of Reason*. Bradford.
- S. Stich, et al. (2001). 'Normativity and Epistemic Intuitions'. *Philosophical Topics*, 29 (1-2) 429-260.

- S. Swain; J. Alexander; J. Weinberg. (2008). 'The instability of philosophical intuitions.' *Philosophy and Phenomenological Research* 76 (1): 138-155.
- A. Turing. (1950), "Computing Machinery and Intelligence," *Mind*, 59 (236): 433–60.
- R. Wedgwood (2002). 'The Aim of Belief'. *Philosophical Perspectives* 16(s16): 267–97.
- J. M. Weinberg (2007). 'How to Challenge Intuitions Empirically Without Risking Skepticism'. *Midwest Studies in Philosophy* 31(1).
- T. Williamson (2008). 'Why epistemology cannot be operationalized', en *Epistemology: New Essays*. Oxford University Press.