

# CREENCIA, RACIONALIDAD Y AGENCIA COGNOSCITIVA† (Belief, Rationality, and Cognitive Agency)

Angeles ERAÑA\*

Manuscrito recibido: 2000.11.24.

Versión final: 2001.11.3.

\* Instituto de Investigaciones Filosóficas, UNAM, Circuito Mario de la Cueva, Ciudad de la Investigación en Humanidades, Coyoacán 04510 México, D.F., México.  
E-mail: mael@filosoficas.unam.mx

BIBLID [0495-4548 (2002) 17: 43; p. 137-159]

RESUMEN: Uno de los supuestos prevaletentes en la epistemología contemporánea es que existe una conexión conceptual entre las nociones de creencia y racionalidad. En este artículo examino una noción débil de racionalidad de acuerdo con la cual *no* es necesario suponer la existencia de *un único conjunto de normas del razonamiento correcto* para poder atribuirle creencias a un sujeto. Desde este punto de vista, lo único que se requiere para la atribución de creencias es *entender* las afirmaciones del sujeto en cuestión y, para ello, sólo es necesario suponer que éste es similar a nosotros en algunos aspectos epistémicamente relevantes.

Descriptores: racionalidad mínima, epistemología, atribución de creencias, principio de humanidad.

ABSTRACT: *A prevailing assumption in contemporary epistemology is that there exists a conceptual connection between the notions of belief and rationality. In this paper I examine a weak notion of rationality, according to which it is not necessary to assume the existence of a unique set of norms of rationality in order to be able to assign beliefs to a subject. From this point of view, belief attribution only requires that we understand a subject's utterances and, to do this, we only need assume that she is similar to us in some epistemic relevant issues.*

Keywords: minimal rationality, epistemology, belief attribution, principle of humanity.

## SUMARIO

Introducción

I. La tesis de racionalidad *global*

II. La tesis de racionalidad mínima

III. Racionalidad, agencia cognoscitiva e interpretación intencional

IV. El principio de humanidad *modificado*

Conclusión

Bibliografía

*Introducción*

Uno de los supuestos prevalecientes en la epistemología contemporánea es que existe una conexión conceptual entre las nociones de creencia y racionalidad: para poder atribuirle creencias a un sujeto éste debe ser suficientemente racional. Sin embargo, no hay un acuerdo respecto al alcance de dicha conexión, de manera que la pregunta de cuán racionales tienen que ser los sujetos para poder atribuirles creencias y, en general, para considerarlos como agentes cognoscitivos, está en el centro de la discusión. En este trabajo defenderemos una noción débil de racionalidad que nos permitirá hacer ver que no es necesario suponer, como lo han hecho un número importante de teorías de la racionalidad contemporáneas, que existe un único *conjunto de normas del razonamiento correcto del cual todo sujeto en todo tiempo y lugar, deba usar alguna de ellas para ser considerado como agente cognoscitivo*.

Entre otras, podemos encontrar las dos siguientes respuestas a la pregunta de cuán racionales tienen que ser los sujetos para considerarlos como agentes cognoscitivos: 1) *no* es necesario que los sujetos sean racionales en *ningún* sentido y; 2) se requiere que los sujetos actúen, por lo menos en ciertas situaciones y en *algún* sentido, racionalmente. Esta última respuesta admite dos versiones distintas: una que exige que el sujeto sea *idealmente* racional, esto es que haga *todas y sólo aquellas acciones que, según sus creencias, lo llevan a satisfacer sus deseos* y otra que *no* requiere que el sujeto sea idealmente racional y que afirmaríamos que es suficiente que el sujeto haga *algunas* acciones aparentemente apropiadas (aunque no todas). En lo que sigue examinaremos dos tesis sustentadas en un requisito de racionalidad *no* ideal: la que llamaremos tesis de "racionalidad global" y la de racionalidad mínima que defiende Cherniak en su libro *Minimal Rationality*.<sup>1</sup>

En la primera sección de este trabajo examinaremos la primera de estas tesis y haremos ver algunas de sus limitaciones y problemas.<sup>2</sup> Después, argumentaremos a favor de la noción cherniakiana de agencia cognoscitiva y, con ella, defenderemos la idea de que las maneras como lo sujetos *de hecho* razonan deben jugar un papel importante en el establecimiento de las normas que dictan las maneras como *deben* hacerlo. Sin embargo, a diferencia de Cherniak -y de la tesis de racionalidad global- en la tercera sección defenderemos la idea de que lo único que se requiere para considerar a un sujeto como agente cognoscitivo es suponer que éste es semejante a nosotros en ciertos aspectos epistémicamente relevantes que más adelante especificaremos. Finalmente propondremos una versión *modificada* del principio en el que se sustenta esta idea, a saber, el llamado "principio de huma-

nidad". Haremos ver que dichas modificaciones nos permiten aceptar un requisito débil para la atribución de creencias, al mismo tiempo que nos permiten llevar a cabo evaluaciones *generalizables* respecto a lo que es un "buen" proceso cognoscitivo.

### I. La tesis de racionalidad global

Un requisito de racionalidad global exigiría algo como lo siguiente: si un sujeto *A* tiene un conjunto particular de creencias y deseos, *A* haría la *mayoría* de las acciones aparentemente apropiadas, donde una acción aparentemente apropiada es aquella que, de acuerdo con las creencias de *A*, tenderá a satisfacer los deseos de *A*. La tesis que se desprende de este requisito sostiene 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; en otras palabras, el sujeto debe poseer una habilidad heurística que *la mayor parte del tiempo* lo lleve a escoger, de entre *todas* las inferencias que serían correctas, a partir de todas y cada una de sus creencias, aquellas que le serán útiles para realizar sus fines. Si *A* no cumpliera con estos requisitos -en otras palabras, si *A* no fuese racional *la mayor parte del tiempo*, donde ser racional es actuar y razonar de acuerdo con algunos principios aceptables de aplicación universal- entonces no podría ser considerado como agente cognoscitivo.

Desde este punto de vista, un sujeto que *sistemáticamente* 1) sostuviese ciertas creencias aun cuando la evidencia que tuviese en contra de ellas fuese mayor a la que tuviese en su favor o; 2) mostrase gran divergencia entre sus creencias, deseos y acciones o; 3) sostuviese inconsistencias internas en su sistema total de creencias o; 4) razonase en desacuerdo con los principios básicos de razonamiento correcto, no sería interpretable por nosotros ya que no podríamos entender sus emisiones. La idea central es que si no fuésemos capaces de dar cuenta de las emisiones y de los actos de un sujeto como afirmaciones y conducta *racional*, no podríamos atribuirle creencias ni, en general, considerarlo como agente cognoscitivo ya que cualquier explicación teórica de sus acciones mostraría inconsistencias sistemáticas que harían imposible la asignación de condiciones de verdad a sus emisiones. Así, si bien para negarle la agencia a un sujeto no es suficiente que éste cometa *algunos* errores al llevar a cabo ciertas inferencias o que, en ocasiones, actúe en desacuerdo con los principios del razonamiento correcto, *no* puede haber agentes cognoscitivos que sean sistemáticamente irracionales; en otras palabras, no podríamos identificar las actitudes proposicionales de un su-

jeto irracional y, en consecuencia, no tendríamos razón alguna para pensar que es un agente cognoscitivo. Veamos.

La tesis de racionalidad global parte del supuesto según el cual las creencias de un sujeto sólo son inteligibles a través de su papel en la interpretación del lenguaje; la idea es que la noción de creencia adquiere sentido en el marco de un patrón de razonamiento coherente, el cual se manifiesta en el habla del sujeto. Así, la noción de agencia cognoscitiva que se desprende de la tesis de racionalidad global requiere de una teoría de la interpretación que a la vez asigne significado a las afirmaciones de un sujeto y le atribuya creencias y otras actitudes proposicionales. Esta teoría de la interpretación presupone un principio de caridad *débil*, según el cual, *a menos que haya evidencia suficiente para lo contrario*, debemos interpretar a los sujetos de manera tal que la mayoría de sus creencias resulten verdaderas, la mayoría de las inferencias que llevan a cabo sean correctas y no sostengan inconsistencias sistemáticas en el interior de su sistema total de creencias (Davidson 1973, 1974, 1975).

De acuerdo con esta postura, para poder interpretar a un sujeto, además de asumir un principio de caridad débil, debemos suponer que a) hay un mundo común que proporciona la misma información (o evidencia) al hablante y al intérprete; b) la mayoría de las oraciones que enuncia el hablante reflejan sus creencias; c) la mayoría de las conexiones racionales que existen entre los contenidos de las creencias del hablante y otros contenidos intencionales coinciden con las conexiones racionales que el intérprete establece entre los contenidos de sus creencias y los contenidos de sus otros estados mentales. La idea central es que toda posible interpretación intencional debe partir del supuesto de que los hablantes, la mayor parte del tiempo, llevan a cabo inferencias correctas "según nuestro propio punto de vista de lo que es correcto" (Davidson 1973, p. 148) ya que de otra manera no podrían llevar a cabo, la mayor parte del tiempo, acciones aparentemente apropiadas, ni podríamos juzgar su conducta como comportamiento racional.

El requisito de racionalidad global lleva implícita una idealización de acuerdo con la cual existe *un único conjunto de principios de razonamiento correcto* que se manifiesta en las emisiones del sujeto, *que determina las maneras como todos los sujetos deben razonar y que es compartido por todo agente cognoscitivo*. Desde este punto de vista, dicho conjunto, además de ser constitutivo de nuestra noción intuitiva de racionalidad, caracteriza apropiadamente nuestra competencia racional: los seres humanos tenemos un conocimiento subyacente de las normas de racionalidad que se mani-

fiesta en una habilidad compartida para utilizarlas de manera correcta la mayor parte del tiempo, por lo cual todos los sujetos debemos (en condiciones normales) ser capaces de llevar a cabo ciertas inferencias correctas y de elegir ciertas acciones aparentemente apropiadas.<sup>3</sup> Si hiciéramos caso de esta idealización, entonces los errores que cometemos en nuestro razonamiento cotidiano no podrían explicarse como *sesgos* de nuestra competencia racional, esto es, como desviaciones *sistemáticas* de nuestro razonamiento con respecto a alguna norma del razonamiento correcto, ya que si ese fuese el caso sería imposible atribuirle creencias a los sujetos y, por tanto, no habría agentes cognoscitivos. Así, la tesis de racionalidad global afirma que las desviaciones de nuestra conducta racional con respecto a las normas de razonamiento correcto se deben a errores de ejecución, esto es, a fallos en la aplicación de los principios de razonamiento correcto debidos a factores de interferencia en nuestro razonamiento, como pueden ser el olvido, la distracción u otros motivos de este tipo. La idea central es que el comportamiento racional de los sujetos se explica "apelando a la operación de nuestra competencia racional en combinación con errores de ejecución causados por factores situacionales o psicológicos" (Stein 1996, p. 61); en otras palabras, el comportamiento cognoscitivo de los sujetos se explica a través del supuesto de que la mayor parte del tiempo razonan de acuerdo con las normas del razonamiento, aunque a veces los utilizan incorrectamente.<sup>4</sup>

Nótese que el requisito de racionalidad global da cuenta de algunas de las limitaciones humanas al reconocer que los sujetos podemos cometer errores en nuestro razonamiento y, ocasionalmente, actuar irracionalmente.<sup>5</sup> En este sentido, dicho requisito puede constituirse en un ideal regulativo: nos propone un fin al que debemos aspirar no sólo porque alcanzarlo sería valioso, sino también porque acercarnos a él y movernos en esa dirección tiene consecuencias benéficas para nosotros (C.A. Hooker 1994). Sin embargo, la explicación del comportamiento racional de los seres humanos que ofrece el requisito de racionalidad global implica que los errores de razonamiento que comúnmente cometemos *no* son un indicativo de la necesidad de revisar o corregir las normas que constituyen nuestra noción intuitiva de racionalidad. En este sentido, la idealización implícita en dicho requisito es "degenerativa":<sup>6</sup> no puede incorporar las desviaciones de nuestro razonamiento respecto al conjunto de normas de razonamiento correcto -del que todo sujeto debe utilizar al menos una para ser considerado como agente cognoscitivo- como posibles correcciones a alguna norma particular. Si este fuese el caso, entonces nuestra capacidad predictiva del comporta-

miento futuro de un agente cognoscitivo se vería mermada, además de que se dificultaría diseñar estrategias que permitan corregir los errores. Veamos.

La distinción competencia-ejecución que subyace en la tesis de racionalidad global implica que los principios normativos, a partir de los cuales se establecen las reglas del razonamiento correcto, son *independientes* de cualquier descripción de las maneras como los sujetos particulares *de hecho* razonan o como *pueden* hacerlo y, por tanto, que la atribución de agencia cognoscitiva depende, exclusivamente, de que la mayor parte del tiempo, elijamos los rumbos de acción prescritos por las normas del razonamiento correcto.<sup>7</sup>

Desde este punto de vista, la conexión conceptual que existe entre racionalidad y creencia *no* se establece a través de un estudio empírico de las maneras como los sujetos razonan, ya que "la racionalidad no es una propiedad optativa de las creencias, sino la precondition de la atribución e individuación de una creencia" (Kim 1993, p. 336). La idea de fondo es que lo que significa "ser racional" es adecuadamente capturado por un conjunto de normas de razonamiento correcto y, por tanto, las maneras como procedemos, las desviaciones de la conducta idealizada y nuestro equipo psicológico resultan irrelevantes para el desarrollo de una teoría de la racionalidad. En otras palabras, las maneras como los sujetos *de hecho* razonan no juegan ningún papel en el establecimiento de las normas del razonamiento correcto (ni son una condición necesaria o suficiente para su modificación o revisión) y, por tanto, la posibilidad de atribuirle agencia cognoscitiva a un sujeto no depende de ellas. Si esto es correcto, entonces la idealización implícita en el requisito de racionalidad global no constituye una descripción simplificadora del comportamiento cognoscitivo real de los seres humanos, sino como un postulado meramente conceptual acerca de las maneras como los sujetos *deben* razonar para ser considerados como agentes cognoscitivos.

Ahora bien, si aceptásemos, de acuerdo con la idealización antes mencionada, que la única información relevante para dar una explicación plausible del comportamiento de un agente cognoscitivo es aquella que apela a la constitución de su competencia racional, entonces cualquier predicción de su conducta futura apuntaría a un comportamiento cognoscitivo acorde con las normas de razonamiento correcto. Sin embargo, los seres humanos tendemos a razonar de acuerdo con algunas reglas heurísticas que a menudo son diferentes de (o inconsistentes con) dichas normas. Una teoría cognoscitiva afín a la idealización implícita en el requisito de racionalidad glo-

bal no podría dar cuenta de este hecho y, por tanto, su capacidad predictiva se vería disminuida y sería exitosa sólo en aquellos casos en los que los sujetos razonasen de acuerdo con las supuestas normas de razonamiento correcto. Además, si aceptásemos que los errores se deben a factores de interferencia no previsibles ni siempre identificables por los intérpretes, y que el conjunto de normas del razonamiento correcto constitutivo de nuestra competencia racional no es susceptible de corrección, entonces la teoría en cuestión no podría darnos pautas para ser "mejores" agentes cognoscitivos.

Como hemos visto, la tesis de racionalidad global considera que, para poder atribuirle agencia cognoscitiva a un sujeto, debemos *suponer* que éste es capaz de razonar, *la mayor parte del tiempo*, de acuerdo con los principios normativos básicos de razonamiento, los cuales prescriben las maneras en que deben relacionarse los contenidos de cada una de nuestras creencias entre sí y con otras actitudes proposicionales. Desde este punto de vista, el hecho de que haya sujetos que sostienen algunas creencias (la minoría) diferentes de las nuestras o que utilizan algunos (los menos) mecanismos de producción de creencias diferentes a los que nosotros usamos, no pone en riesgo la atribución de agencia cognoscitiva. Es suficiente que esos sujetos compartan con nosotros los principios normativos de razonamiento y actúen la mayor parte del tiempo de acuerdo con ellos. Así, si aceptamos que una condición necesaria y suficiente para llevar a cabo una interpretación intencional es utilizar el principio de caridad débil, entonces podríamos aceptar diferencias no sustantivas entre sistemas de creencias de unos y otros sujetos, pero "no podríamos estar en condiciones de juzgar que otras personas tienen conceptos o creencias radicalmente diferentes de los nuestros" (Davidson 1974, p. 202).

Como veremos en la siguiente sección, una noción de agencia cognoscitiva más débil, como la que sugiere el requisito de racionalidad mínima, nos permite desarrollar una teoría cognoscitiva que, además de servir como ideal regulativo, parta de una idealización útil que nos dé las herramientas para explicar y predecir el comportamiento cognoscitivo real de los agentes cognoscitivos.

## II. La tesis de racionalidad mínima

La tesis de racionalidad mínima, que sostiene Cherniak en su libro *Minimal Rationality*, afirma que un sujeto A puede ser considerado como agente cognoscitivo sólo si A *tiene un conjunto particular de creencias y hace algunas*, aunque no necesariamente todas, *las acciones que tienden a satisfacer sus*

*deseos* (Cherniak 1986, p. 9). A diferencia de la tesis de racionalidad global, este autor considera que un sujeto no necesariamente debe hacer la *mayoría* de las inferencias correctas para ser considerado como agente cognoscitivo; es suficiente que el sujeto haga *algunas* inferencias que *puede* hacer dadas sus limitaciones y capacidades cognoscitivas. El agente, dice Cherniak, no sólo debe intentar algunas de las acciones que son apropiadas, dado un conjunto particular de creencias, sino que además *no* debe intentar suficientes acciones que son *inapropiadas* dado ese mismo conjunto de creencias. Es necesario añadir este requisito negativo pues, sin él, una persona que actúa de manera azarosa -su comportamiento está determinado por "volados", por ejemplo- podría ser considerada como agente racional, ya que eventualmente tomaría una (o más) decisión(es) que lo llevará(n) a actuar de manera apropiada.

Así, el requisito de racionalidad mínima dice que, para ser agente cognoscitivo, uno debe (1) seleccionar algunas inferencias correctas que son *útiles* bajo circunstancias específicas para satisfacer sus deseos; (2) hacer exitosamente algunas inferencias correctas y; (3) no hacer un gran número de acciones que son *inapropiadas* dado el conjunto de creencias que uno posee. Después Cherniak añade un cuarto elemento a este requisito: las inferencias que hace el sujeto deben ser *factibles* para él, donde una inferencia es factible para *A* si *A* la elige del conjunto de inferencias (correctas) que le son accesibles en ese momento y si *A* puede llevarla a cabo de manera exitosa. La posibilidad de que un sujeto lleve a cabo una inferencia de manera exitosa no depende exclusivamente de la dificultad de la inferencia misma (de sus características sintácticas, por ejemplo), también depende de las condiciones bajo las cuales se lleva a cabo: para cada sujeto, dependiendo de sus capacidades, hay inferencias que son más difíciles aunque parezcan más fáciles para otros. Supongamos que *q* se sigue de *p* y que *S* cree que *p*. Supongamos también que la inferencia de *p* a *q* es útil para *S* en este momento. De esto no se sigue que *S* hará la inferencia de *p* a *q* ya que es posible que, en la situación en que se encuentra *S*, la inferencia de *p* a *q* no sea factible para *S* (*S* puede no tener las capacidades cognoscitivas requeridas para *esa* inferencia); que *S* crea que *q* no es el caso, lo cual lo lleva a rechazar *p* o, simplemente, que la creencia de *S* de que *p* esté en su memoria de largo alcance y, por tanto, que *S* no pueda acceder a ella en *ese* momento.

La tesis de racionalidad mínima tal como Cherniak la plantea es una tesis *descriptiva* a la cual puede asociarse una tesis *normativa*. La tesis descriptiva simplemente *describe* las condiciones bajo las cuales un sujeto



puede ser considerado como agente cognoscitivo; la tesis normativa asociada propone un ideal regulativo al que cualquier agente cognoscitivo debe aspirar. En tanto que la primera de ellas dice que para ser considerado como agente cognoscitivo se requiere que un sujeto haga *algunas* inferencias correctas que sean factibles para él, es posible afirmar que el conjunto de inferencias que se requieren para que alguien pueda ser considerado como agente cognoscitivo *es un subconjunto propio* del conjunto de inferencias que el agente *debe* hacer para ser *normativamente mínimamente* racional. La tesis normativa mínima es enunciada por Cherniak en los siguientes términos:

La persona debe hacer *todas* (y sólo aquellas) inferencias correctas y factibles de sus creencias que, de acuerdo con sus creencias, tenderían a satisfacer sus deseos (Cherniak 1986, p. 23).<sup>8</sup>

En el caso en que *p* implica *q*, si *S* cree que *p* y la inferencia de *p* a *q* es factible para *S* y aparentemente apropiada, la tesis normativa dice que *S* *debe* hacer esta inferencia para ser mínimamente normativamente racional, pero no dice nada acerca de lo que *S* *de hecho* va a hacer, ni acerca de lo que *debe* hacer *para* ser considerado como agente cognoscitivo. La tesis normativa mínima menciona las condiciones de factibilidad para determinar aquello que el sujeto *debe* hacer y describe, por tanto, un *ideal regulativo* que sirve como guía de las decisiones y las acciones de los sujetos. En tanto que esta tesis considera que no tiene sentido requerir que alguien haga algo que no puede hacer, su punto de partida son las maneras como *los sujetos* pueden *razonar para establecer las maneras en que* deben hacerlo.

La idea central es que no podemos afirmar *a priori* que todo agente cognoscitivo debe hacer las mismas inferencias. Por el contrario, para saber *cuáles son las inferencias que debe hacer un sujeto en un momento dado re-querimos información acerca de qué inferencias puede hacer exitosamente en la situación en la que se encuentra*. Así, "ser agente cognoscitivo" para un sujeto implica que haga algunas de las inferencias correctas que puede hacer dadas sus limitaciones, el lugar y el tiempo en que se encuentre, esto es, "ser agente cognoscitivo" es razonar de acuerdo con ciertas normas que se establecen tomando en cuenta la situación del sujeto, y el tiempo y el lugar en que se aplican.<sup>9</sup>

Ahora bien, desde este punto de vista la posibilidad de interpretar la conducta cognoscitiva de un sujeto depende de: 1) tomar en cuenta la evidencia que tiene el interpretado para sostener una creencia (evidencia que pueda serle accesible al intérprete al mismo tiempo) y; 2) reconocer que

todos los seres humanos tenemos limitaciones cognoscitivas y un tiempo limitado para llevar a cabo inferencias. El cumplimiento de estas condiciones *no* implica que, para ser agentes cognoscitivos, los sujetos deban llevar a cabo la mayoría de las inferencias correctas o sostener un mayor número de creencias verdaderas que falsas.<sup>10</sup> Más bien, deja en claro que la conexión que existe entre la posibilidad de atribuirle creencias a un sujeto y la posibilidad de describirlo correctamente como racional es mucho menos fuerte de lo que la tesis de racionalidad global exige. Veamos.

La tesis de racionalidad mínima considera que la atribución de racionalidad debe tomar en cuenta tanto la información empírica acerca de las capacidades cognoscitivas de los sujetos, como los constreñimientos prácticos que las situaciones específicas imponen a nuestro razonamiento, de manera que la pregunta acerca de cuáles son los requisitos que deben satisfacerse para ser considerados como agentes cognoscitivos es parcialmente una pregunta empírica. Aceptar esto nos permite, por un lado, desarrollar una teoría cognoscitiva que establezca lo que es "ser agente cognoscitivo" a partir del comportamiento *real* de los sujetos y, por tanto, una teoría cognoscitiva *que tenga poder predictivo y que sea útil en la práctica*; en otras palabras, nos permite explicar el comportamiento cognoscitivo de los sujetos sin apelar a un ideal normativo *independiente* de las capacidades cognoscitivas *reales* de los sujetos. Por otro lado, nos permite establecer un ideal regulativo que tome en cuenta una teoría psicológica; que nos permita prescribir y evaluar al mismo tiempo las maneras como los sujetos deben razonar para mejorar sus procedimientos cognoscitivos y que nos permita dar "consejos" acerca de las maneras como debemos proceder para hacer acciones aparentemente apropiadas. Un ideal de esta naturaleza sirve, a la vez, como guía en nuestros razonamientos y como meta a la cual debemos aspirar para ser considerados como agentes cognoscitivos.

Hemos visto que es posible desarrollar una tesis de racionalidad que tenga como punto de partida una noción débil de agencia cognoscitiva y que considere que la posibilidad de interpretar a un sujeto depende, por lo menos en parte, de la posibilidad de describir sus estados intencionales a partir de un estudio empírico de *sus* capacidades y limitaciones cognoscitivas reales. Como veremos en lo que sigue, la noción de racionalidad que subyace en esta tesis no requiere suponer la existencia de un único conjunto de normas del razonamiento correcto del cual todo sujeto deba usar al menos algunas para ser considerado como agente cognoscitivo.

### III. Racionalidad, agencia cognoscitiva e interpretación intencional

La tesis de racionalidad mínima, como vimos anteriormente, es una tesis *descriptiva* acerca de las capacidades reales de los sujetos para razonar, a la cual se asocia una tesis *normativa* mínima acerca de la aceptabilidad epistémica de las creencias. A diferencia de la tesis de racionalidad global, la de racionalidad mínima nos permite aceptar las siguientes dos tesis: 1) es necesario tomar en cuenta las maneras como los sujetos *de hecho* razonan y las maneras como *pueden* hacerlo para establecer las maneras como *deben* razonar y; 2) algún sujeto que sostuviese un conjunto de creencias sustantivamente diferente del nuestro, o que llevase a cabo inferencias particulares *típicamente* diferentes de las que nosotros haríamos en circunstancias similares, podría ser considerado como agente cognoscitivo.

Cherniak, sin embargo, considera que "puede haber esquemas genuinamente diferentes de los nuestros, pero *ininteligibles* para nosotros" (Cherniak 1986, p. 131)<sup>11</sup>, lo cual sugiere que no aceptaría que pueda haber sujetos, *inteligibles para nosotros*, cuyos razonamientos y maneras de proceder obedezcan a un conjunto de normas de razonamiento diferente del conjunto de normas propio de *nuestros* procedimientos epistémicos. Esto se debe, en gran medida, a que comparte con la tesis de racionalidad global la idea de que hay *un único* conjunto de normas, constitutivo de nuestro concepto de racionalidad, del cual cualquier agente cognoscitivo, en cualquier lugar, debe usar por lo menos alguna en *cualquier* momento.<sup>12</sup>

En lo que sigue sostendremos que lo único que se requiere para la atribución de creencias es suponer cierta semejanza entre agentes. Este supuesto puede articularse a través del principio de humanidad para la interpretación,<sup>13</sup> el cual como veremos, nos permite aceptar, con Cherniak, que la posibilidad de interpretar a un sujeto depende de tomar en cuenta sus capacidades y limitaciones cognoscitivas, pero también (y a diferencia de Cherniak) nos permitirá rechazar la noción de interpretación subyacente en el principio de caridad y, con ella, la idea de que es necesario suponer la existencia de un único conjunto único de normas del razonamiento correcto.

El principio de humanidad, formulado por Grandy y examinado por Stich (Stich 1990, p. 45)<sup>14</sup>, dice que para que un sujeto  $S_1$  pueda interpretar a un sujeto  $S_2$  basta que  $S_1$  y  $S_2$  sean similares en ciertos aspectos epistémicamente relevantes; es suficiente que  $S_1$  y  $S_2$  compartan algunas creencias, o lleven a cabo algunas inferencias similares en situaciones específicas, o compartan algunos principios básicos de razonamiento, o algunas catego-

rías básicas. La idea central es que  $S_1$  puede interpretar a  $S_2$  si y sólo si puede asignarle significado o intencionalidad a las afirmaciones o acciones de  $S_2$  con base en aquellas que  $S_1$  enunciaría o haría en la situación en la que  $S_2$  se encuentra.

De acuerdo con este principio, la posibilidad de caracterizar intencionalmente los estados mentales de un sujeto depende de la posibilidad de generar un modelo del tipo de conexión racional que dicho sujeto establece entre dichos estados, con su conducta y con el medio ambiente. Sin embargo, no siempre tenemos el conocimiento requerido de las creencias y deseos del sujeto en cuestión para generar dicho modelo, de manera que para llevar a cabo una interpretación intencional, el intérprete debe tomarse a sí mismo como modelo y suponer que las interacciones de los estados mentales del hablante son similares a las suyas propias.<sup>15</sup> La idea central es que para poder afirmar 'S cree que  $p$ ' es suficiente suponer que la creencia de S en  $p$  está relacionada con el resto de sus creencias, deseos, percepciones, con su medio ambiente y con su comportamiento de *maneras semejantes* a las maneras en que están relacionados *mis* estados mentales. La semejanza requerida debe ser tal que  $S_1$  identifique el estado de creencia que  $S_2$  expresa al enunciar ' $p$ ' y lo explique en términos de las relaciones racionales que  $S_1$  establecería si enunciara ' $p$ '.

Si lo anterior es correcto, entonces podemos aceptar que, para llevar a cabo una interpretación intencional de un sujeto, *no* es necesario suponer, como lo hace el principio de caridad, que las afirmaciones (sinceras) de los agentes son verdaderas, ni que estos siguen, *la mayor parte del tiempo*, la mayoría de los principios del razonamiento correcto que comparten con nosotros. La interpretación intencional, desde este punto de vista, no es la asignación de condiciones de verdad a las oraciones del hablante, sino la caracterización de los posibles estados mentales que tiene el sujeto en cuestión a partir del supuesto de que dichos estados cognoscitivos son semejantes a estados reales y posibles nuestros.<sup>16</sup> De esta noción de interpretación se desprende la idea de que es preferible atribuirle a un sujeto una creencia falsa pero explicable que una creencia verdadera pero misteriosa.

Ahora bien, los estándares de semejanza relevantes para la descripción intencional son, desde esta perspectiva, relativos a fines específicos y dependen de los sujetos involucrados y del momento en que se hace dicha descripción; esto es, la posibilidad de interpretar a un sujeto es una cuestión de grados y de respectos. La idea central es que la posibilidad de entender las afirmaciones o las acciones de un sujeto y, por tanto, la posibilidad de acceder a sus conceptos y de describirlos intencionalmente se ale-

ja si el hablante no muestra, en circunstancias similares, cierta semejanza relevante con el intérprete. Así, si hubiese criaturas que utilizasen procedimientos cognoscitivos muy diferentes de los nuestros, no serían intencionalmente describibles por nosotros. Sin embargo, la imposibilidad de interpretar los estados intencionales de un sujeto no significa que éste no tenga creencias. Los procesos cognoscitivos de un sujeto *S* pueden ser muy diferentes de los míos (en ciertos aspectos) sin que ello implique que no hay algún otro sujeto *P* que puede interpretar a *S* y que puede hacerme acceder a los conceptos de *S* a través de aquellos que *P* comparte conmigo.

El principio de humanidad, a diferencia del principio de caridad, nos permite explicar porqué, en ocasiones, no es posible describir intencionalmente los estados mentales de aquellos sujetos que son muy diferentes de mí y explica cómo pueden, esos sujetos, ser considerados como agentes cognoscitivos: yo *no* puedo dar cuenta de las conexiones racionales que *S* establece entre sus estados mentales porque no son semejantes a las que yo establezco, pero, en tanto que la semejanza es una cuestión de grados, *P* puede hacerme accesibles las afirmaciones de *S* y explicarme las conexiones racionales que *S* establece entre sus estados mentales, de manera que, a través de la semejanza con *P*, yo puedo considerar a *S* como agente cognoscitivo. Si esto es correcto, entonces la atribución de agencia cognoscitiva siempre es relativa a la semejanza entre dos sujetos particulares y la evaluación de los procesos cognoscitivos variaría de sujeto a sujeto.

Estamos de acuerdo con este planteamiento en que no se requiere suponer una conexión conceptual entre descripción intencional y racionalidad *al mismo tiempo* que una conexión conceptual entre racionalidad y verdad masiva. Sin embargo, no estamos de acuerdo con la idea de que los estándares de semejanza relevantes sean relativos a individuos específicos. Si esto fuera correcto, entonces podría haber tantas concepciones de aceptabilidad epistémica como sujetos con estados mentales, lo cual implicaría que *no* hay estándares de evaluación epistémica generalizables a dominios de conocimiento o a comunidades de personas: cada sujeto podría tener una noción de racionalidad particular que podría ser considerada como "aceptable" si resultase eficaz en el medio ambiente en que se usa. La idea de que los criterios de evaluación y atribución de creencias son relativos a la semejanza entre dos sujetos particulares nos lleva a justificar, como racional, casi cualquier conducta cognoscitiva, lo cual nos enfrenta a un fuerte individualismo que, en última instancia, es conducente a un relativismo difícilmente sostenible.

En lo que sigue haremos ver que algunas modificaciones al principio de humanidad nos permiten aceptar (como correctas) tanto las explicaciones en torno a la atribución de agencia cognoscitiva, como la noción débil de racionalidad en que dicho principio se sostiene sin aceptar el individualismo al que éste podría llevarnos. Veremos que otra manera de entender las noción de semejanza nos permite aceptar que hay criterios de atribución y de evaluación de creencias *generalizables* a dominios de conocimiento.

#### IV. *El principio de humanidad modificado*

El principio de humanidad *modificado* que aquí se propone puede establecerse de la siguiente manera: para que un sujeto  $S_1$  pueda llevar a cabo una interpretación intencional de un sujeto  $S_2$  basta que  $S_1$  *entienda* las afirmaciones y el comportamiento de  $S_2$ , donde entender implica tener una idea aproximada del tipo de proceso cognoscitivo prescrito en la cultura donde está inmerso  $S_2$ . Para que esto sea posible es necesario que  $S_1$  y  $S_2$  sean semejantes en el sentido de que ambos sean epistémicamente responsables. Examinemos las condiciones que este principio impone a la interpretación intencional.

El principio de humanidad *modificado* comparte con los principios de humanidad y caridad el supuesto de que para poder llevar a cabo una interpretación del habla de un sujeto es necesario *entender* sus emisiones, esto es, es necesario poder dar cuenta de ellas y de su conducta como afirmaciones y comportamiento *racional*. Sin embargo, la noción de entendimiento que subyace en él marca profundas diferencias con estos dos principios. La teoría de la interpretación que acompaña al principio de caridad sostiene que el comportamiento de los sujetos es la base de la cual inferimos lo que quieren decir y, a partir de dicha evidencia, empatamos los sonidos que emiten con una explicación teórica del mismo que constituye nuestro entendimiento (Bar-On 1999, p. 4). Sin embargo, como hicimos ver anteriormente, las generalizaciones que utilizamos para correlacionar la conducta del sujeto con las circunstancias observadas no nos dan una explicación satisfactoria de nuestra capacidad predictiva respecto al comportamiento futuro de los sujetos. Por ello y para sostener las explicaciones propuestas por el principio de humanidad es necesario que la noción de entendimiento tenga una dimensión pragmática que el principio de caridad deja fuera.

Una interpretación intencional, desde este punto de vista, debe poder dar cuenta de la corrección o incorrección de la conducta cognoscitiva del

hablante con base en los constreñimientos que guían y estructuran *su uso* de conceptos. A diferencia del principio de humanidad, el *modificado* considera que dichos constreñimientos -o las guías que rigen nuestra estructuración de conceptos y la elección y uso de ciertos procesos cognoscitivos y no de otros- se establecen y tienen *status* epistémico en el marco de una cultura o de un dominio de conocimiento particular. La idea central es que el entendimiento no debe sustentarse en la proyección de *nuestros* estándares de corrección, sino en el conocimiento que tengamos o podamos adquirir de los estándares de corrección que prevalecen en la cultura o el dominio específico de conocimiento en que se lleva a cabo la evaluación.

Como mencionamos anteriormente, el principio de humanidad *modificado*, al igual que el de humanidad, requiere que los sujetos, *algunas veces*, actúen de manera similar a como nosotros lo haríamos en determinadas circunstancias y, por tanto, requiere semejanzas entre el intérprete y el interpretado. Sin embargo, a diferencia del principio de humanidad, la semejanza requerida por el *modificado* debe ser tal que: 1) las expectativas que tiene  $S_1$  con respecto a las acciones o a las afirmaciones de  $S_2$  sean aproximadamente adecuadas, en el sentido que permitan predecir, con cierto grado de confianza, los rumbos de acción que  $S_2$  *deberá* tomar y; 2) cuando las expectativas no sean cumplidas,  $S_1$  pueda identificar y *explicar* las desviaciones de la conducta esperada a partir de la identificación del tipo de constreñimiento que guía el razonamiento de  $S_2$ . Veamos.

El segundo de estos requisitos se sustenta en la idea de que la ocurrencia de errores impone ciertos constreñimientos a los rumbos de acción que, de acuerdo con las creencias del sujeto, tenderán a satisfacer sus deseos. En tanto que dichos constreñimientos pueden darnos una idea aproximada del tipo de proceso cognoscitivo epistémicamente permitido *en la cultura a la que pertenece el sujeto en cuestión*, dada su situación, y sus capacidades y limitaciones, nuestras expectativas respecto al resultado correcto de ciertos procesos cognoscitivos y el diagnóstico del tipo de error que dicho sujeto es susceptible de cometer al llevar a cabo ciertas inferencias o al elegir rumbos de acción deben partir de cierta empatía hacia el hablante y *no* de la imposición de *nuestros* criterios de corrección.

Ahora bien, respecto al primero de los requerimientos de semejanza, la idea es que las expectativas de  $S_1$  con respecto a  $S_2$ , si  $S_2$  es interpretable por  $S_1$ , siempre incluyen la de que  $S_2$  es racional. Si aceptamos que una acción o afirmación racional se caracteriza por estar constreñida por normas del razonamiento, y si aceptamos que dichas normas se establecen toman-

do en cuenta las maneras como los sujetos *de hecho* razonan y las maneras como *pueden* hacerlo, entonces podemos afirmar que la conducta cognoscitiva de un sujeto interpretable por nosotros está sujeta a ciertas normas, aunque no necesariamente a *nuestros* criterios de corrección epistémica. Si esto es correcto, entonces no es necesario suponer que los criterios de evaluación deben ser fijos y generales y, por tanto, no es necesario suponer que hay un único conjunto de principios del razonamiento correcto. Sin embargo, es necesario suponer que la conducta cognoscitiva de los sujetos es regida por constreñimientos que se establecen en el marco de una cultura o de un dominio de conocimiento específico y que fungen como criterios de evaluación epistémica. La idea central es que el tipo de interacciones racionales que establece un sujeto entre sus estados mentales y su conducta depende, en gran medida, de su participación en una cultura<sup>17</sup> o en un dominio de conocimiento específico.

Si lo anteriormente argumentado es correcto, entonces pueden haber diferentes criterios de evaluación epistémica e incluso podríamos aceptar que hay conjuntos de normas del razonamiento correcto diferentes entre sí. Sin embargo, es necesario aceptar que, sean cuales fueren los criterios de corrección, una afirmación o acción racional es aquella para la cual pueden darse razones y que, de esta manera, satisface dichos criterios y sirve como razón para otros enunciados o acciones. Así, y dado que nuestras expectativas incluyen la de que el sujeto actuará racionalmente, el rango de acciones predecibles y *relevantes* para el requisito de semejanza tiene restricciones importantes. Por ejemplo, hay acciones o afirmaciones que no necesariamente deben evaluarse como racionales o irracionales, ya que si bien son predecibles, no son acciones relevantes para la satisfacción de este requisito: si un loco compulsivo se lava las manos cada cinco minutos y hace dos minutos que se las lavó por última vez, tenemos un alto grado de certidumbre en la predicción de lo que hará en los siguientes minutos, sin embargo, su acción no será evaluada como racional o irracional, ya que la evaluación requerida por el principio de humanidad *modificado* no debe estar sustentada en la descripción de regularidades observadas, sino en la posibilidad de que las acciones puedan servir como razones para futuras acciones. También puede haber casos en que sujetos semejantes a nosotros -que rigen su conducta por normas de racionalidad<sup>18</sup>- utilicen normas que no se ajusten a las nuestras. Para poder explicar las desviaciones de la conducta esperada en casos como éste debemos indagar el tipo de procedimiento cognoscitivo sancionado por la comunidad a la que pertenece dicho sujeto. La



idea central es que el *entendimiento* es una precondition de la evaluación o el juicio.

Si lo anteriormente expuesto es aceptado, entonces el principio de humanidad *modificado* daría las siguientes directivas para interpretar a un hablante:

*A menos que haya evidencia suficiente para lo contrario, debemos interpretar a los sujetos como si fueran epistémicamente responsables, esto es, como si utilizaran los mejores procesos cognoscitivos que les son accesibles para llevar a cabo acciones aparentemente apropiadas (aun si no tiene éxito en alcanzarlos).<sup>19</sup>*

En otras palabras, debemos suponer que los hablantes tienen *buenas razones* para afirmar o hacer lo que dicen o hacen, donde una "buena razón" para algo no es sólo aquello que un sujeto particular considera que *debe* hacer, sino también aquello que es sancionado por una comunidad como capaz de conferir justificación a la afirmación (o a la acción). En este sentido es necesario afirmar que los estándares a través de los cuales se juzga cuándo un sujeto es epistémicamente responsable son socialmente articulados.

La idea subyacente en el principio de humanidad *modificado* es que la interpretación intencional es parte de un proceso social en el cual diferentes sujetos interaccionan y buscan entenderse mutuamente. Así, la posibilidad de atribuirle creencias a un sujeto depende de nuestra capacidad para *entender* las acciones y las afirmaciones de los sujetos como parte de una red de razones articulada en una estructura social que establece los criterios de evaluación y que nos proporciona una idea del tipo de proceso cognoscitivo que un sujeto puede y debe utilizar en determinadas circunstancias. En otras palabras, para que  $S_1$  pueda interpretar a  $S_2$ ,  $S_1$  debe entender a  $S_2$  en el sentido de tener una buena idea de aquellos procesos prescritos en la cultura donde está inmerso  $S_2$ , de manera tal que pueda juzgarlo y evaluar la razonabilidad de sus afirmaciones.

Si la argumentación anterior es correcta, entonces podemos aceptar que el principio de humanidad *modificado*, a diferencia del de caridad, no requiere suponer la existencia de *un único* conjunto de normas de racionalidad, constitutivo de nuestro concepto de racionalidad, tal que todo sujeto en todo tiempo y lugar deba usar por lo menos alguna de esas normas y tal que, si una presunta norma no pertenece a ese conjunto, entonces está epistémicamente prohibido usarla. Lo único que se requiere para la atribución

de creencias y, en general, para la atribución de agencia cognoscitiva es *entender* el comportamiento cognoscitivo del sujeto en cuestión.

Ahora bien, nuestra habilidad para entender las acciones o afirmaciones de los otros sujetos no requiere que estos sean racionales en *nuestro* sentido de racionalidad; más bien descansa en el supuesto de que los procesos cognoscitivos que utilizan para resolver problemas, o para elegir acciones aparentemente apropiadas, o para hacer juicios, sugieren las maneras en que esos sujetos *son* racionales. Para entender las afirmaciones de sujetos diferentes de nosotros, en el sentido epistémicamente relevante antes mencionado, es necesario tener una idea aproximada del tipo de afirmaciones que puede enunciar, o acciones que puede llevar a cabo, o errores que puede cometer en circunstancias específicas; en otras palabras, necesitamos "ponernos *sinceramente* en los zapatos" del sujeto y prever el tipo de rumbos de acción que él es susceptible de tomar. Esto implica que *no* pueden hacerse generalizaciones *universales* -aplicables a todos los seres humanos en todo tiempo y lugar- acerca de lo que es un "buen" procedimiento cognoscitivo, pero no necesariamente implica que los criterios de evaluación deban ser diferentes para *cada sujeto* y, por tanto, que no puedan hacerse generalizaciones *de ningún tipo* con respecto a los procedimientos cognoscitivos.<sup>20</sup> Si aceptamos que es posible *entender* las afirmaciones y las acciones de agentes que usan normas del razonamiento diferentes de las nuestras, entonces también es posible enunciar juicios respecto a la manera como esos sujetos pueden ser "mejores" agentes cognoscitivos y, por tanto, llevar a cabo evaluaciones *generalizables* acerca de lo que es un "buen" proceso cognoscitivo.

### Conclusión

Si aceptamos una teoría cognoscitiva como la que se desprende de la tesis de racionalidad mínima, entonces podemos aceptar una noción de agencia cognoscitiva débil, de acuerdo con la cual la posibilidad de interpretar a un sujeto depende de tomar en cuenta sus capacidades y limitaciones cognoscitivas. Esta noción de agencia cognoscitiva abre la posibilidad de establecer una teoría de la interpretación que no requiera suponer que la conexión que existe entre racionalidad y creencia debe establecerse *a priori*. La pregunta acerca de cuán racionales tienen que ser los sujetos para ser considerados como agentes cognoscitivos es parcialmente una pregunta empírica y, en esa medida, la atribución de creencias debe tomar en cuenta cuestio-

nes de factibilidad y las condiciones prácticas en que se encuentra el sujeto a quien se busca interpretar.

Por otro lado, puesto que, como hemos argüido, lo único que debemos suponer para llevar a cabo una interpretación intencional y para poder atribuirle creencias a un sujeto y, en general, para considerarlo como agente cognoscitivo es que éste es *epistémicamente responsable* y dada la noción de responsabilidad epistémica aquí expuesta, podemos aceptar que la evaluación de la razonabilidad de las creencias de un sujeto no es algo que pueda llevarse a cabo *independientemente* de su constitución psicológica, del lugar y el tiempo en que está inmerso y de las circunstancias específicas en que el sujeto lleva a cabo una inferencia o una acción particular. Esto no implica, sin embargo, que no haya normas de evaluación epistémica: lo único que implica es que no hay nada que nos asegure que las normas que usamos para evaluar el comportamiento epistémico de los sujetos deben ser aplicables a todos los demás sujetos en todo tiempo y lugar. En otras palabras, el principio de humanidad *modificado* nos permite aceptar que no es necesario suponer que existe un único conjunto de normas del razonamiento correcto del cual todo agente cognoscitivo debe usar alguna para ser considerado como tal.

Una noción de agencia cognoscitiva como la que aquí busca defenderse, acompañada por un principio de interpretación como el de humanidad *modificado*, abre la puerta a un tipo de epistemología naturalizada que utilice tanto los métodos, como los resultados de las ciencias empíricas para determinar, por lo menos en parte, los ideales regulativos que nos sirvan en la evaluación de creencias y en la selección de teorías. Una epistemología de este tipo resaltaría la importancia de conocer tanto las características relevantes del equipo cognoscitivo de los agentes como los contextos específicos en que estos se desarrollan. De acuerdo con la postura que, en última instancia, quisiera defender, la racionalidad es función de un conjunto de capacidades y habilidades heurísticas que van desarrollándose de acuerdo con el tipo de problemas al que se enfrentan los agentes cognoscitivos a lo largo del tiempo: los mecanismos que cada sujeto elige para darles solución dependen, hasta cierto punto, del tipo de mecanismo que se utiliza más en el ambiente en que dicho sujeto está inmerso y del éxito o fracaso que los mecanismos hayan mostrado tener a lo largo de la historia. Esto, sin embargo, es materia de discusión para otro trabajo.

## Notas

- † Este trabajo se hizo con el apoyo del proyecto de investigación del CONACYT 30966. Agradezco a Sergio Martínez Muñoz y a Claudia Lorena García Aguilar las ideas que aportaron, los valiosos comentarios y el tiempo que tomaron en hacer correcciones a versiones preliminares de este trabajo. Asimismo quisiera agradecer las sugerencias y correcciones que hicieron Gustavo Córdoba y María Inés Pazos.
- 1 Si bien las tesis de "racionalidad nula" y de "racionalidad ideal" son planteamientos importantes para los problemas que aquí se discutirán, no abordaremos ninguna de ellas por falta de espacio y porque no son directamente relevantes para el argumento central del trabajo, ni para la tesis que aquí sostendremos. Para una discusión detallada en torno a estas posiciones y sus problemas centrales, Cf. Cherniak (1986, capítulo 1).
  - 2 Es importante mencionar que en este trabajo no daremos una argumentación detallada a favor o en contra de esta posición. Además de que un número importante de autores ya lo han hecho (Cf. Davidson 1990, Bar-On 1994, entre otros), aquí nos interesará exponerla brevemente para hacer ver que si bien es una alternativa prevalecte en la epistemología contemporánea, no es la única tesis sobre la racionalidad plausible.
  - 3 Esta argumentación presupone la distinción competencia-ejecución, de acuerdo con la cual la competencia racional de un sujeto es el conocimiento tácito que éste tiene de los principios del razonamiento y su habilidad para razonar correctamente bajo circunstancias adecuadas; y la ejecución es la manera concreta de implementar dichos principios y puede ser influida por factores de interferencia, tales como el olvido o la distracción.
  - 4 Una diferencia central de esta tesis con alguna otra que suponga un requisito de racionalidad ideal es que, para esta última, un sujeto puede ser considerado como agente cognoscitivo sólo si su capacidad de razonar *siempre* lo lleva a realizar inferencias correctas (o, en otras palabras, su capacidad de razonar *nunca* se desvía de las normas del razonamiento correcto).
  - 5 A diferencia de una tesis sostenida en un requisito de racionalidad ideal, esta tesis no sugiere que, para ser considerados como agentes cognoscitivos, debemos tener una capacidad de memoria y de computación infinita; lo único que exige es que la mayor parte del tiempo actuemos conforme a los principios de razonamiento correcto.
  - 6 Cf. C.A. Hooker (1994, p. 207), donde el autor afirma que "(...) real behaviour is *perspicuously* represented as some deviation from simplifying idealized behaviour, while for degenerate idealizations it cannot in general be so represented".
  - 7 Es común que en situaciones de riesgo o en situaciones donde hay que responder rápidamente los seres humanos utilicen mecanismos heurísticos o estrategias útiles que les permiten llevar a cabo acciones aparentemente apropiadas dado el tiempo que tienen disponible y dadas sus capacidades cognoscitivas limitadas. Los resultados de utilizar estos mecanismos de razonamiento son, en general, diferentes de aquellos que arrojaría el uso de los supuestos principios de razonamiento correcto. Cherniak se refiere al uso de "heurísticas rápidas pero sucias" que nos permiten llegar a conclusiones correctas sin llevar a cabo razonamientos formales (los cuales nos llevarían a las mismas conclusiones de manera más segura). Si bien estas estrategias no son infalibles ni apli-

cables en todo contexto, y sólo son confiables en ciertos medios ambientes y dentro de ciertos parámetros, la elección de tales mecanismos de inferencia *es racional* y la descripción de su uso debería jugar un papel en la atribución de agencia cognoscitiva. Cf. Cherniak (1986, p. 78).

<sup>8</sup> El subrayado es mío.

<sup>9</sup> La tesis de racionalidad mínima afirmarí­a incluso que pueden haber situaciones en las cuales dos sujetos que comparten las mismas creencias, los mismos deseos e intereses *no deban* llevar a cabo las mismas inferencias, o que las acciones que son apropiadas en unos casos no lo sean en otros.

<sup>10</sup> Es importante enfatizar que la tesis de racionalidad mínima (y la posición de este trabajo) *no* busca definir lo que es "ser racional" como aquello que uno hace la mayoría de las veces. La incorporación en una teoría normativa de las maneras como los sujetos *pueden* razonar busca evitar que la definición de "ser racional" dependa de qué tan a menudo son puestos a prueba los sujetos, o de las opciones que se les presentan para manifestar su irracionalidad, ya que esta última no es función de la cantidad de veces que cometemos errores.

<sup>11</sup> El subrayado es mío.

<sup>12</sup> La diferencia más importante de la tesis de racionalidad mínima respecto a la tesis de racionalidad global es que *no* supone que este conjunto de normas nos dé los criterios para considerar a un sujeto como agente cognoscitivo, más bien, nos da los criterios para decir como *debe* razonar un sujeto para ser "mejor" agente cognoscitivo.

<sup>13</sup> Este principio de la interpretación nos permite explicar porqué hay una conexión conceptual entre racionalidad y creencia. El esquema de la explicación es como sigue: 1) Para considerar a un sujeto S como agente cognoscitivo basta que S sea similar a mí en ciertos aspectos epistémicamente relevantes; 2) puedo considerar a un sujeto como agente cognoscitivo sólo si puedo atribuirle creencias y otros estados intencionales; 3) si 1 & 2, entonces para poder atribuirle creencias a un sujeto es suficiente que ese sujeto sea similar a mí en ciertos aspectos epistémicamente relevantes; 4) si 3, entonces *el principio de humanidad*; 5) creo que soy racional y, por tanto, que los sujetos similares a mí (en aspectos epistémicamente relevantes) también lo son; 6) si 4 & 5, entonces puedo atribuirle creencias a un sujeto sólo si éste es racional; 7) si 2 & 6, entonces los agentes cognoscitivos son racionales.

<sup>14</sup> Alvin Goldman, Robert Gordon y otros autores sustentan lo que se ha llamado "teoría de la simulación" en este principio de interpretación. La idea central es que las interpretaciones intencionales que llevamos a cabo están basadas en una *simulación* de la situación de los interpretados, esto es, en inferencias que hacemos de ciertos supuestos acerca de los estados mentales previos del interpretado a sus nuevos estados. Cf. A. Goldman (1995, p. 91). Es importante mencionar que Stich ha sido muy crítico de esta posición. Para una discusión detallada al respecto véase Martin Davies y Tony Stone (eds.) (1995).

<sup>15</sup> Nótese que el principio de humanidad, al igual que el de caridad, busca maximizar el acuerdo. Sin embargo, a diferencia de éste último, el de humanidad no supone que la interpretación debe preservar la verdad de las oraciones emitidas por el hablante, además de que considera que la historia del hablante y sus circunstancias son relevantes para llevar a cabo una interpretación. En este sentido, la noción de interpretación subya-

- cente en este principio no consiste en la asignación de condiciones de verdad a las oraciones emitidas por el hablante en el lenguaje del intérprete. Cf. R. Grandy (1973).
- 16 Goldman y otros autores afirman que para poder llevar a cabo una interpretación intencional es necesario que el intérprete le atribuya al hablante sus mismas categorías fundamentales o, al menos, las mismas operaciones de formación de conceptos. Cf. A. Goldman (1995, p. 91).
- 17 La cultura funge como depositaria de una gran colección de procesos cognoscitivos que han dado soluciones parciales a problemas específicos y de los cuales los sujetos particulares deben elegir aquellos que los llevarán a hacer acciones aparentemente apropiadas. Cf. E. Hutchins (1996, p. 177).
- 18 En este trabajo no se intenta articular una teoría sobre la normatividad epistémica por lo cual no hemos dado una caracterización de aquello que es distintivo de las normas de racionalidad.
- 19 Este requisito nos permite afirmar que la noción de racionalidad que subyace en el principio de humanidad *modificado*, a diferencia de la que subyace en el de humanidad, sirve como guía de las maneras como deberíamos actuar en casos en que no sabemos qué es más conducente a la realización de acciones aparentemente apropiadas. Esto se debe a que esta noción *no* está sustentada en la medida del éxito epistémico de los sujetos particulares.
- 20 Un pragmatismo individualista como el que Stich, en última instancia, busca defender, tendría este tipo de implicación. Cf. Stich (1990, capítulo 6).

## BIBLIOGRAFIA

- Bar-On, Dorit: 1994, 'Conceptual Relativism and Translation', in G. Preyer et al. (eds.): *Language, Mind and Epistemology*, Holanda, Kluwer Academic Publishers, pp. 145-170.
- : 1999, 'Between Pluralism and Relativism', artículo sin publicar presentado en el *XIV Congreso Interamericano de Filosofía*, Puebla, México.
- Brandom, Robert: 1995, 'Knowledge and the Social Articulation of the Space of Reasons', *Philosophy and Phenomenological Research* LV/4, 895-908.
- Cherniak, Christopher: 1986, *Minimal Rationality*, Cambridge, Massachusetts, The MIT Press.
- Cohen, Jonathan: 1986, *The Dialogue of Reason*, Oxford, Clarendon Press.
- Davidson, Donald: 1970, 'Semántica para lenguajes naturales', in *De la verdad y de la interpretación: contribuciones a la filosofía del lenguaje*, traducción de Guido Filippi, Barcelona, Gedisa, 1990, pp. 73-81.
- : 1973, 'Interpretación radical', in *De la verdad y de la interpretación: contribuciones a la filosofía del lenguaje*, pp. 137-150.
- : 1974, 'De la idea misma de un esquema conceptual', in *De la verdad y de la interpretación: contribuciones a la filosofía del lenguaje*, pp. 189-203.
- : 1975, 'Pensamiento y habla', in *De la verdad y de la interpretación: contribuciones a la filosofía del lenguaje*, pp. 164-178.
- : 1985, 'Incoherence and Irrationality', *Dialectica* 39/4, 345-354.

- Davies, Martin and Stone, Tony (eds.): 1995, *Folk Psychology*, Cambridge, Massachusetts, Blackwell.
- Goldman, Alvin I.: 1995, 'Interpretation Psychologized', *Folk Psychology*.
- Gordon, Robert: 1995, 'Folk Psychology as Simulation', *Folk Psychology*.
- Grandy, Richard: 1973, 'Reference, Meaning, and Belief', *The Journal of Philosophy* LXX, no. 14, 439-452.
- Hooker, C.A.: 1994, 'Idealisation, Naturalism, and Rationality: Some Lessons From Minimal Rationality', *Synthese* 99, 181-231.
- Huthcins, Edwin: 1996, *Cognition in the Wild*, Cambridge, Massachusetts, The MIT Press.
- Kahneman, Slovic, Tversky (eds.): 1982, *Judgement and Uncertainty: Heuristics and Biases*, Cambridge, Massachusetts, Cambridge University Press.
- Kim, Jaegwon: 1993, 'What Is Naturalized Epistemology?', in Louis P. Pojman (ed.): *The Theory of Knowledge*, Belmont, California, Wadsworth Inc, pp. 329-340.
- Stein, Edward: 1996, *Without Good Reason*, Oxford, Clarendon Press.
- Stich, Stephen: 1990, *The Fragmentation of Reason*, Cambridge, Massachusetts, The MIT Press.
- , Nichols, Shaun: 1995, 'Folk Psychology: Simulation or Tacit Theory', *Folk Psychology*.

Angeles Eraña tiene una maestría en Filosofía de la Ciencia por la Universidad Nacional Autónoma de México. Autora de artículos sobre problemas de normatividad epistémica y diversidad cognoscitiva y miembro de la Asociación Filosófica de México. Actualmente termina la redacción de su tesis doctoral y es Estudiante Asociada del Instituto de Investigaciones Filosóficas, UNAM.