

# Realismo y constructivismo en la Teoría moral kantiana: el ejemplo de la Ética del discurso \*

CRISTINA LAFONT  
Northwestern University, Evanston

**RESUMEN.** En este artículo se argumenta contra la interpretación metaética del kantianismo moral como una forma de constructivismo o antirrealismo moral. Dado que los kantianos no comparten el expresivismo característico del antirrealismo moral estándar, el constructivismo kantiano parece llevar a una posición inherentemente inestable que sólo puede desarrollarse o bien en un realismo consistente con el cognitivism moral kantiano o en un decidido antirrealismo moral. Tomando la ética del discurso de Habermas como ejemplo, aquí se contrasta una interpretación realista con una interpretación antirrealista del principio moral de universalización con la intención de mostrar que sólo la primera es compatible con el cognitivism moral característico de las éticas kantianas, mientras que la segunda lleva inevitablemente a un planteamiento decisionista y, con ello, a un claro relativismo moral.

**ABSTRACT.** In this paper I argue against the metaethical view of moral Kantianism as a form of constructivism or antirealism. Given that Kantians do not share the expressivism characteristic of standard moral antirealism, Kantian constructivism seems to be an inherently unstable position, which can only be fully developed into either a consistently antirealist or a consistently realist approach. Taking Habermas' discourse ethics as an example, I contrast a realist with an antirealist interpretation of the principle of universalization, and try to show that only the former is compatible with the moral cognitivism characteristic of Kantian moral theory, whereas the latter unavoidably leads to a decisionist (i. e., relativist) approach.

Un rasgo llamativo de los debates contemporáneos en metaética es la proliferación de toda suerte de realismos, antirrealismos y quasi-realismos morales, así como de una variedad de combinaciones de ellos prácticamente innumerable. En términos generales, dichos debates metaéticos pueden entenderse como versiones modernas del viejo debate entre las corrientes empirista y racionalista de la filosofía moral (en el que los primeros típicamente defienden el antirrealismo y los segundos el realismo). La filosofía moral kantiana, sin embargo, parece poco afectada por estos debates. Al menos hasta el momento los kantianos no han participado muy activamente en ellos. Quizás una de las razones

---

\* En este texto se basó mi lección inaugural (*Antrittsvorlesung*) en la Universidad de Francfort que tuvo lugar el 27 de junio de 2001. Traducción de Gerard Vilar.

de este desinterés sea el que parece incuestionable que la teoría moral kantiana es una rama del constructivismo, es decir, del antirrealismo<sup>1</sup>. Esta impresión seguramente se ve reforzada por el hecho de que los representantes contemporáneos más importantes de la filosofía moral kantiana, Habermas y Rawls, entienden sus teorías morales respectivas como constructivistas o antirrealistas. Sin embargo, ni Kant ni Rawls ni Habermas comparten los supuestos expresivistas característicos del antirrealismo moral estándar, es decir, la tesis de que al hacer juicios morales ni siquiera pretendemos hacer afirmaciones acerca de lo que es objetivamente correcto o incorrecto, sino que simplemente damos expresión a nuestras actitudes no cognitivas. Los kantianos no son expresivistas, sino cognitivistas morales, es decir, defienden explícitamente la validez objetiva de los juicios morales. En la medida en que el expresivismo puede considerarse la característica paradigmática del antirrealismo moral, en el caso del kantianismo moral puede tratarse a lo sumo de un *antirrealismo anómalo*. Por un lado, parece compartir con el antirrealismo el supuesto de que los juicios morales no son estrictamente verdaderos o falsos en la medida en que su validez es una función de nuestras creencias y actitudes. Pero, por otro lado, comparte con el realismo la afirmación de que los juicios morales son objetivos, esto es, universalmente válidos. Sin embargo, dadas las obvias diferencias en creencias y actitudes entre los seres humanos, el sentido en que la moralidad sería supuestamente objetiva bajo supuestos antirrealistas resulta más que misterioso. Es difícil imaginarse la posibilidad de mostrar un sentido en el que nuestros juicios morales podrían ser universalmente válidos que al mismo tiempo no implicara excluir que dicha validez sea meramente función de las actitudes o creencias de unos u otros.

Debido a esta anomalía, el constructivismo kantiano parece ser una posición esencialmente inestable<sup>2</sup>. Desarrollar sus implicaciones metaéticas parece que obligaría o bien a defender una posición coherentemente realista o bien una posición coherentemente antirrealista. Se puede seguir una estrategia coherentemente antirrealista y afirmar que la corrección moral es exclusivamente función de nuestras creencias y actitudes. Como intentaré mostrar en lo que sigue, esta estrategia es incompatible con el cognitivismo moral y conduciría inevitablemente a una posición decisionista, es decir, relativista. Dado que el cognitivismo moral es un rasgo esencial de la filosofía moral kantiana, esta estrategia relativista conduciría a una teoría moral que ya no sería reconociblemente kantiana. O bien se puede permanecer firme en la defensa del cognitivismo moral y reconocer que la corrección moral de las normas no es función exclusivamente de nuestras creencias y actitudes. Sin embargo, al

<sup>1</sup> Para una interesante excepción a esta tendencia véase M. Smith, *The Moral Problem*, Blackwell, 1995.

<sup>2</sup> Muchos críticos del «constructivismo kantiano» de Rawls han apuntado la inestabilidad de esta posición que pretende no ser ni realista ni antirrealista. Véase O. O'Neill, *Constructions of Reason*, Cambridge UP, 1989, pp. 206-218; también D. Brink, *Moral Realism and the Foundations of Ethics*, Cambridge UP, 1989, pp. 303-322.

hacer eso ya se ha concedido todo lo que una estrategia realista requiere. Lo que quedaría por mostrar en este caso es que los supuestos realistas implícitos en dicha estrategia son compatibles con los rasgos constructivistas generales característicos de la filosofía de Kant.

Aunque estoy convencida de que lo son, no voy a intentar mostrarlo en este contexto. Aquí no voy a centrarme en la teoría moral de Kant, sino en uno de sus desarrollos contemporáneos, a saber, la ética del discurso. No obstante, la discusión de este planteamiento que ofreceré a continuación puede verse como un primer paso en el proyecto más general de esclarecer el núcleo realista implícito en el kantianismo moral. El objetivo que hay tras este proyecto no es negar los muchos rasgos constructivistas inherentes a la teoría moral kantiana, sino sólo mostrar que en relación a la cuestión específica que constituye el núcleo del debate entre realismo y antirrealismo en metaética, a saber, en virtud de qué pueden ser objetivamente válidos los juicios morales, sólo una respuesta realista es compatible con el cognitivismo moral. Una vez que se acepta esto resulta claro que la teoría moral kantiana representa efectivamente una alternativa genuina tanto al racionalismo como al empirismo tradicionales precisamente porque *combina* elementos realistas y constructivistas de la forma más plausible posible. En este sentido, el análisis que sigue de la ética del discurso persigue mostrar que la teoría moral kantiana no debería interpretarse como una rama anómala del antirrealismo, sino más bien como el único tipo de realismo moral netamente plausible.

### *Realismo y antirrealismo en la ética del discurso*

Aun cuando la ética del discurso es una teoría filosófica que ha sido profusamente desarrollada y discutida en los últimos treinta años, la interpretación exacta de sus tesis básicas parece seguir siendo poco clara. Ello es así no sólo por lo que hace a sus fines y posibles desarrollos, como evidencian las intrincadas discusiones entre sus fundadores, K.-O. Apel y J. Habermas. La opacidad afecta incluso a su núcleo mismo, es decir, a la conexión interna entre condiciones discursivas y corrección moral que dicha teoría postula. En su reciente libro, *Verdad y Justificación*<sup>3</sup>, Habermas aborda explícitamente la tarea de determinar la naturaleza exacta de dicha conexión, y en esa misma medida también especificar lo que él considera es la lectura correcta de la ética del discurso. Lo que motiva esta tarea es la reciente retractación por parte de Habermas de la teoría discursiva de la verdad, así como la convicción de que debe seguirse una estrategia realista para explicar la verdad, pero no para explicar la corrección moral. De acuerdo con esto, Habermas defiende

<sup>3</sup> J. Habermas, *Wahrheit und Rechtfertigung*, Frankfurt, Suhrkamp, 1999 (en adelante citado como *WR*).

ahora una interpretación explícitamente antirrealista (o constructivista) de la ética del discurso<sup>4</sup>. Aunque estoy completamente de acuerdo con la nueva posición de Habermas en relación a la verdad, su lectura antirrealista de la ética del discurso destruye, en mi opinión, el núcleo mismo de dicha teoría, a saber, la defensa del cognitivismo moral. A continuación voy a distinguir tres posibles lecturas de la ética del discurso y a argumentar que sólo la lectura que resulta de aplicar una estrategia realista permite entender la ética del discurso como una defensa del cognitivismo moral.

La ética del discurso se puede caracterizar como un intento de analizar la conexión interna entre corrección moral y aceptabilidad racional, es decir, como un intento de explicar «el punto de vista moral basado en los presupuestos comunicativos generales de la argumentación»<sup>5</sup>. El objetivo es mostrar cómo las condiciones discursivas de aceptabilidad racional, que en cuanto tales son moralmente neutras (es decir, están motivadas por razones cognitivas y no morales), imponen restricciones a las posibles respuestas a cuestiones morales. Ello es así al menos en la medida en que *estas cuestiones son tratadas como cognitivas* (y, por consiguiente, como necesitadas de justificación) *por los participantes* en los discursos morales. La práctica de justificar las normas morales sujeta sus posibles resultados a *condiciones generales necesarias* de aceptabilidad racional. Si es plausible considerar las cuestiones morales como cuestiones cognitivas, y si la explicación discursiva de la noción de aceptabilidad racional es correcta, entonces las condiciones discursivas de la aceptabilidad racional contenidas en dicha explicación son condiciones necesarias para la validez de nuestras pretensiones de conocimiento en relación a la corrección moral de las normas, exactamente del mismo modo que lo son para nuestras pretensiones cognitivas en relación a la verdad de los enunciados.

Así pues, la ética del discurso ofrece una regla de argumentación inherente a los discursos morales en los que está en cuestión la *justificación* de normas. Esta regla de argumentación es el llamado principio de universalización (U), principio que expresa la condición que cualquier norma debe satisfacer a fin de ser válida, esto es, de estar *justificada*:

(U) que las consecuencias y los efectos colaterales que se siguen de su acatamiento *general* para la satisfacción de los intereses de *cada* persona (presumiblemente) pueden ser aceptados por *todos* los afectados (así como preferidos frente a las posibilidades alternativas de regulación)<sup>6</sup>.

<sup>4</sup> Véase «Richtigkeit versus Wahrheit. Zum Sinn der Sollgeltung moralischer Urteile und Normen», en *WR*, pp. 271-318.

<sup>5</sup> En *Erläuterungen zur Diskursethik*, Frankfurt, Suhrkamp, 1991, p. 119; versión castellana *Aclaraciones a la ética del discurso*, Madrid, Trotta, 2000, p. 127 (en adelante citado como *EDE*).

<sup>6</sup> «Diskursethik:», en *Moralbewusstsein und kommunikatives Handeln*, Frankfurt, Suhrkamp, 1983, pp. 75-76; «Ética del discurso: Notas sobre un programa de fundamentación», en *Conciencia moral y acción comunicativa*, Barcelona, Península, 1985, p. 86 (en adelante citado como *ED*).

Un análisis de los diferentes elementos contenidos en esta regla de argumentación o principio moral nos permite verla como una interpretación discursiva del punto de vista moral. El principio de universalización (U) está constituido por dos componentes diferentes:

Por una parte, dicho principio expresa las condiciones discursivas generales de aceptabilidad racional inherentes a toda práctica de justificación. La pretensión de que una norma es moralmente correcta (o un enunciado verdadero) puede considerarse justificada si pudiera ser aceptada como convincente *por todos* en las condiciones de una situación ideal de diálogo. La interpretación discursiva de la aceptabilidad racional, cuando se aplica a normas en particular, se traduce en el principio discursivo:

D: Válidas son aquellas normas (y sólo aquellas normas) a las que todos los que puedan verse afectados por ellas podrían dar su asentimiento como participantes en discursos racionales<sup>7</sup>.

Por otro lado, más allá de las condiciones discursivas generales, el principio moral (U) tiene que especificar además en qué consiste el punto de vista moral. Sostener que una norma es moralmente correcta significa que está en igual interés de todos. Como Habermas explica en *Facticidad y validez*:

El principio moral sólo resulta de una especificación del principio general de discurso para aquellas normas de acción que *sólo* pueden justificarse desde el punto de vista de si tienen en cuenta por igual los intereses de todos (p. 173).

Esta característica general de la ética del discurso se deriva de dos consideraciones básicas que sirven de apoyo a toda tentativa de dar una explicación discursiva del punto de vista moral. Habermas las resume en su «Ética del discurso: Notas sobre un programa de fundamentación»:

La ética del discurso se mantiene o se desmorona con estos dos supuestos: *a*) que las pretensiones de validez normativas tienen sentido *cognitivo* y se pueden tratar *como* pretensiones de verdad, y que *b*) la *fundamentación* de normas y mandatos requiere la realización de un *discurso* real que en último término no puede darse monológicamente, esto es, en la forma de un hipotético proceso de argumentación que ocurra en la mente individual (*DE*, p. 78; las cursivas son mías).

Primero de todo, es importante remarcar que los dos supuestos mencionados por Habermas no se hallan al mismo nivel. Es decir, el segundo supuesto sólo tiene sentido si el primero es correcto. Cualquier propuesta de identificar las condiciones decisivas (esto es, necesarias y suficientes) para la justificación de normas morales tiene sentido sólo una vez que se ha establecido que tales normas pueden ser justificadas en general. Sólo si las cuestiones morales pueden

<sup>7</sup> En *Facticidad y validez*, Madrid, Trotta, 1998, p. 172.

considerarse cognitivas tiene sentido el intento de justificar qué respuestas a tales cuestiones son moralmente correctas y cuáles no lo son. De lo contrario, el problema de la justificación ni siquiera se plantearía.

A fin, pues, de explicar por qué los participantes en discursos morales consideran que los juicios morales precisan de justificación es necesario explicar en qué sentido (esto es, en virtud de qué) pueden ser considerados seriamente por éstos como *cognitivos*. Sólo si el análisis de las cuestiones morales fuera capaz de mostrar que hay algo que *conocer* acerca de las mismas, tendría entonces sentido preguntar por el modo más racional de descubrir la respuesta correcta a dichas cuestiones. En otras palabras, es necesario analizar en qué medida son comparables las pretensiones de validez verdad y corrección moral a fin de determinar hasta qué punto es realmente defendible el cognitivismo de la ética del discurso.

El supuesto de que hay algo que conocer acerca de un tipo dado de cuestión implica que su respuesta se basa en un hecho, es decir, que lo que puede conocerse es o bien de un modo o bien de otro, y, por consiguiente, no puede ser el caso que una respuesta y su opuesta sean ambas correctas. Si las cuestiones morales son realmente cognitivas, los participantes en los discursos morales tienen que suponer que las cuestiones morales (planteadas con suficiente claridad) sólo admiten «una única respuesta correcta», exactamente como suponen que es el caso para las cuestiones acerca de la verdad.

Este requerimiento parece basarse en la *validez incondicional* que los hablantes atribuyen a los juicios morales sobre normas. Dichos juicios son *absolutos* o, lo que es lo mismo, están codificados *binariamente* (es decir, no decimos que una norma es *más o menos* justa, sino sólo justa o injusta). También tienen un carácter *fijo*, en el sentido de que no admiten cambios en el tiempo (es decir, no decimos que una norma antes era justa y ahora es injusta o viceversa). Nuestro uso del par de términos justo/injusto parece ser estructuralmente idéntico a nuestro uso de los términos verdadero/falso.

Sin duda hay diferencias de significado obvias entre las dos pretensiones de validez desde el punto de vista del *contenido*. No obstante, ambas tienen algo en común que explica las similitudes estructurales en nuestro uso de las mismas: el supuesto cognitivista de «una única respuesta correcta» tiene sentido en ambos casos, precisamente en la medida en que conocemos *las condiciones necesarias y suficientes para la satisfacción de ambas pretensiones de validez*. A ello se debe que ambas pretensiones de validez sean desempeñables discursivamente. En la misma medida en que conocemos las condiciones necesarias y suficientes para la verdad de un enunciado p ('P' es verdadero si y sólo si p), conocemos también las condiciones necesarias y suficientes para la corrección moral de una norma N (N es moralmente correcta (justa) si y sólo si está igualmente en interés de todos). Captar el significado de dichas pretensiones de validez es lo que permite a los participantes *saber lo que tienen que buscar* en cada caso, según si se trata de un discurso teórico o práctico-moral. Como

Habermas explica en *Aclaraciones a la ética del discurso*: «Al igual que el modo asertórico puede explicarse mediante la existencia de los estados de cosas afirmados, el modo deontológico puede explicarse mostrando que las acciones requeridas están igualmente en interés de todos los posiblemente afectados» (EDE, pp. 130 y 138).

Los supuestos realistas inherentes a ambas pretensiones de validez explican por qué nuestra discusión acerca de ambos tipos de cuestiones tiene lugar bajo la premisa de «una única respuesta correcta». Nuestro conocimiento relativo a la verdad de los enunciados depende de la existencia de los estados de cosas afirmados en el mundo objetivo. Dependiendo de lo que sea el caso, el enunciado sólo puede ser *verdadero o falso*. De modo similar, nuestro conocimiento acerca de la corrección moral de las normas depende de la existencia de intereses comunes entre todos los individuos pertenecientes al mundo social. En virtud de lo que resulten ser los intereses generalizables entre todos los seres humanos<sup>8</sup>, una norma sólo puede ser *justa o injusta*.

Estos supuestos realistas cumplen un rol *formal* crucial en tanto que son la *condición de sentido de los discursos teórico y práctico* respectivamente. Pues del mismo modo que el supuesto de la existencia de estados de cosas en el mundo objetivo es la condición de posibilidad de una discusión significativa acerca de la verdad de los enunciados, el supuesto de la existencia de intereses generalizables entre los individuos en el mundo social es la condición de posibilidad de una discusión significativa acerca de la corrección moral de las normas sociales. Este supuesto de existencia es indispensable para los discursos prácticos no porque sea necesariamente el caso que existen tales intereses entre todos los seres humanos, sino más bien porque si llegáramos a la conclusión de que dichos intereses *no existen* y, por tanto, que dicho supuesto no tiene sentido (lo cual es obviamente una cuestión empírica, abierta), cuestionar la corrección moral de las normas también carecería de sentido. Si, *de facto*, no existieran intereses comunes entre todos los seres humanos, no entenderíamos qué sig-

<sup>8</sup> La presuposición de que existen intereses generalizables, compartidos por todos los seres humanos, va más allá de la presuposición de que pueden existir algunos intereses que *de facto* estén compartidos por todos los seres humanos en un momento dado. Para ser más exactos, lo que se presupone es que algunos de dichos intereses compartidos son *irrenunciables* para *todos* los seres humanos, es decir, son aquellos intereses a los que no es posible renunciar por ser necesarios para mantenerse como un ser racional. En sus *Lectures on the History of Moral Philosophy* (Harvard University Press, 2000), Rawls identifica una presuposición similar como crucial para el cognitivismo moral de Kant. Rawls argumenta que para poder explicar cómo la ley moral puede tener contenido objetivo, es decir, para explicar cómo el imperativo categórico puede especificar preceptos que sean básicamente los mismos para todos los agentes racionales parece necesario apelar a lo que Kant denomina en la *Metafísica de las costumbres* «verdaderas necesidades» humanas: «I understand Kant to say that we have certain true human needs, certain requisite conditions, the fulfillment of which is necessary if human beings are to enjoy their lives» (p. 174). Sólo bajo dicho supuesto tiene sentido pensar que lo que los seres humanos pueden querer racionalmente es lo mismo para todos. Como lo expresa Rawls, el test de la contradicción en la voluntad del imperativo categórico presupone «that we have such needs and that they are more or less the same for everyone» (p. 174).

nifica considerar una norma justa o injusta. Pero en la medida en que es razonable suponer que existen tales intereses, en la medida que nos parece que tiene sentido descubrir qué normas están «en igual interés de todos», sólo puede haber *una* respuesta correcta. Es decir, toda norma o bien salvaguarda dichos intereses de todos por igual, o no lo hace.

Con la ayuda de este supuesto es posible entender la oposición entre el realista kantiano y el antirrealista. Pues dicho supuesto permite determinar exactamente qué es lo que un realista kantiano supone que existe y lo que un antirrealista supone que no existe. De acuerdo con la breve caracterización ofrecida más arriba, la corrección moral de una norma está en función de si preserva los intereses generalizables de todos los seres humanos o no; si lo hace, la norma es moralmente correcta, si no lo hace, es incorrecta. Los intereses generalizables de todos los seres humanos son aquellos intereses *irrenunciables* para *todos* ellos, es decir, aquellos intereses de los que los seres humanos no pueden racionalmente prescindir, pues son necesarios para mantenerse como seres racionales. De acuerdo con ello, un antirrealista o relativista moral está genuinamente en desacuerdo con este supuesto del discurso moral cuando afirma que no hay literalmente nada que pueda considerarse un interés generalizable entre todos los seres humanos racionales. Pues lo que es racional que los seres humanos quieran depende esencialmente de los deseos y preferencias que *de facto* tengan y dichos deseos no son compartidos (al contrario, normalmente se oponen entre sí). Precisamente el desacuerdo con este supuesto realista del discurso moral es lo que motiva el enfoque revisionista con relación al discurso moral ordinario característico del expresivismo moral. Visto desde esta perspectiva, el realista moral kantiano no está postulando ninguna ontología excéntrica en el sentido de Mackie. Las entidades contenidas en dicho supuesto son reconocidas igualmente por ambas posiciones, a saber, los diferentes intereses que los seres humanos tienen. En este sentido, el conflicto entre ambas posiciones no se debe a la introducción de entidades adicionales por parte del realista. Lo que este último defiende es que existe un conjunto de intereses que todos los seres humanos considerarían irrenunciables, y el antirrealista, por el contrario, defiende que tal conjunto no existe en absoluto (esto es, sostiene que el producto de la intersección de dichos intereses es el conjunto vacío, por decirlo así).

En «Aclaraciones a la ética del discurso», Habermas ofrece un explicación muy parecida del cognitivismo moral de la ética del discurso en su crítica al constructivismo de Rawls<sup>9</sup>. Sin embargo, aunque es posible encontrar una explicación muy similar a la ofrecida aquí de las semejanzas entre las pre-

<sup>9</sup> Habermas explica la conexión interna entre verdad y corrección moral del modo siguiente: «Una pretensión de validez dice que se cumplen las respectivas condiciones de validez de una expresión, sea esta última una afirmación o un mandato moral. Que se cumplen, no es algo que se pueda mostrar echando mano directamente a evidencias irrefutables, sino sólo por el camino de la satisfacción discursiva de la pretensión, sea esta última de verdad proposicional o de corrección normativa. Las condiciones de validez que se sustraen a un acceso directo se

tensiones de validez «verdad» y «corrección moral» dondequiera que Habermas defiende el estatus cognitivo de las cuestiones morales, lo cierto es que recientemente ha empezado a insistir cada vez más en las diferencias entre ambas pretensiones de validez. Analizar estas diferencias es fundamental a la hora de determinar qué interpretación de la ética del discurso es la que Habermas favorece actualmente y, por consiguiente, en qué medida es defendible la asunción del cognitivismo moral si se aceptan dichas diferencias.

Al retractarse de su concepción epistémica de la verdad en su artículo «El giro pragmático de Rorty»<sup>10</sup>, Habermas añade una nota de advertencia. Mientras la estrategia realista puede adoptarse para explicar la pretensión de validez «verdad», no ocurre otro tanto con «las pretensiones de validez morales u otras pretensiones normativas que presentan una orientación hacia el desempeño discursivo» (p. 264). Ello es así porque dichas pretensiones

«carecen de la propiedad de “trascender la justificación” que poseen las pretensiones de verdad debido a la presunción de un único mundo objetivo inserta en el uso comunicativo del lenguaje. Las pretensiones de validez normativas surgen de un mundo de relaciones interpersonales en el seno de un mundo social que no es independiente de “nuestro hacer” en el modo en que lo es el mundo objetivo. El tratamiento discursivo de las pretensiones normativas es, sin embargo, “análogo a la verdad” en la medida en que los participantes en el discurso práctico se orientan por la meta de una “única respuesta correcta” obligada, permitida o prohibida. El mundo social es intrínsecamente histórico, esto es, constituido ontológicamente de modo diferente al mundo objetivo. Por esta razón, en el caso del mundo social, la idealización de las condiciones justificatorias no puede incluir la “anticipación de una corroboración futura”, en el sentido de una refutación anticipada de futuras objeciones...; más bien se puede entender sólo en el sentido crítico de una reserva en relación al grado de convergencia, es decir, al estado de descentramiento efectivamente alcanzado por la comunidad de justificación. El desempeño discursivo de una pretensión de verdad implica que están satisfechas las condiciones de verdad interpretadas como condiciones de asertabilidad. En el caso de una pretensión de validez normativa el consenso alcanzado discursivamente fundamenta el que la correspondiente norma sea digna de reconocimiento, y en esa medida el consenso mismo contribuye a la satisfacción de las condiciones de validez de la norma. Mientras que la aceptabilidad racional es meramente indicativa de la verdad de una proposición, ofrece una contribución constructiva en relación a la validez de las normas» (WR, p. 261; la cursiva es mía).

Habermas insiste aquí en una interpretación epistémica de la pretensión de validez «corrección moral» como «aceptabilidad racional en condiciones

---

interpretan mediante las razones que se pueden aducir en el discurso. Y el tipo de razones relevantes para la satisfacción discursiva de una pretensión de validez arroja luz sobre el sentido específico de la pretensión de validez misma elevada en cada caso. *Al igual que el modo asertórico se puede aclarar mediante la existencia de los estados de cosas afirmados, así también el modo deontológico puede serlo mostrando que las acciones mandadas van en interés por igual de todos los posiblemente afectados»* (EDE, pp. 130 y 138; la cursiva es mía).

<sup>10</sup> En WR, pp. 230-270.

ideales». En este sentido, y contrariamente a la interpretación ofrecida anteriormente, el principio de universalización (U) no establece las condiciones epistémicas bajo las cuales está *justificada* nuestra pretensión de corrección moral en relación a una cierta norma. Lo que dicha regla establece es que esas condiciones epistémicas (esto es, el consenso discursivo entre los afectados) son las condiciones para que la norma sea *moralmente correcta*.

En todas las exposiciones de las tesis centrales de la ética del discurso el principio de universalización (U) es introducido como una condición para la validez de las normas (desde el punto de vista moral). Pero siempre permanece ambiguo a *qué tipo de validez* se está haciendo referencia exactamente. En la medida que se supone que (U) es una regla de argumentación para la *justificación* de normas morales en los discursos prácticos, parece natural pensar, como hice antes, que las normas que satisfacen (U) son válidas en el sentido de que están *justificadas*. Pero si se acepta la interpretación antirrealista ofrecida en la cita anterior, es decir, si la pretensión de corrección moral no trasciende la justificación, entonces el hecho de que una norma esté justificada (esto es, que satisfaga (U)) es *todo lo que significa ser «moralmente correcta»*. La corrección moral no significa nada más que aceptabilidad racional en condiciones ideales. Habermas hace explícitamente esta afirmación en *Verdad y justificación* cuando escribe:

Un acuerdo acerca de normas o acciones alcanzado discursivamente en condiciones ideales posee más que una fuerza autorizadora *garantiza* la corrección de los juicios morales. Asertabilidad idealmente garantizada *es* lo que llamamos validez moral ... A diferencia del caso de una pretensión de verdad que trasciende la justificación, la asertabilidad idealmente garantizada de una norma no hace referencia más allá de los límites del discurso a algo que «exista» independientemente de ser digno de reconocimiento. El que la «corrección» no trasciende la justificación se debe a un argumento conceptual: puesto que la «validez» de una norma consiste en su ser aceptada, esto es, reconocida como válida en condiciones de justificación ideales, la «corrección» es un concepto epistémico (*WR*, pp. 297-298).

Como vimos anteriormente, el segundo supuesto mencionado por Habermas como crucial para una defensa exitosa de la ética del discurso sólo le compromete a la tesis de que «la *justificación* de normas y mandatos requiere que se realice un *discurso* real» (*DE*, p. 68). No obstante, la interpretación ofrecida más arriba de la ética del discurso compromete a un supuesto *adicional*, a saber, que la justificación discursiva de normas es en lo que consiste su corrección moral. Consiguientemente, los discursos morales tendrían que garantizar, no sólo la *justificación*, sino también la *corrección* moral. Sin embargo, a fin de defender la tesis fuerte de que la noción de aceptabilidad racional *explica* la noción de corrección moral, la primera noción no puede ser entendida en un sentido operacional, sino que debe entenderse como una noción de éxito. Esto es, dicha noción debe incluir la corrección moral como condición:

la aceptabilidad racional discursiva no es la mejor vía epistémica para *inferir* la corrección moral de una norma, sino más bien lo que *establece su corrección moral*. Ésta es la estrategia seguida en «Aclaraciones a la ética del discurso», donde Habermas muestra en qué consiste la idealización de la aceptabilidad racional (discursiva):

El discurso moral práctico representa la extensión ideal de cada comunidad de comunicación desde dentro. En este forum *sólo aquellas normas propuestas que expresan un interés común de todos los afectados pueden lograr un asentimiento justificado* (EDE, pp. 113 y 121).

Esta caracterización del discurso moral nos permite entender la aceptabilidad racional en condiciones ideales de dos maneras diferentes. De un lado, podemos entenderla en un sentido *realista*. Esto significaría que en tales condiciones ideales, las partes afectadas se pondrían de acuerdo únicamente en aquellas normas que *de hecho* expresan un interés común, compartido por todos ellos. De otro lado, podríamos interpretarla en un sentido *antirrealista*, es decir, *decisionista*. Ello significaría que en las antedichas condiciones ideales, las normas que las partes afectadas acordarían, *sean las que fueren*, merecerían ser calificadas de estar «en igual interés de todos» debido *únicamente al hecho de que fueron acordadas de ese modo*.

En el sentido antirrealista, sin embargo, es imposible defender el cognitivismo y la validez incondicional de las pretensiones de corrección moral. Si «ser de interés común a todos los afectados» significa ser el resultado del acuerdo discursivo entre éstos, una vez que dichos afectados han *decidido de hecho* cuáles son las normas en las que están de acuerdo, ya no tiene sentido hablar de una posible revisión o de crítica de las mismas. Como Tugendhat observa en su crítica de la ética del discurso:

Lo que finalmente resulta decisivo es el acuerdo fáctico, y no tenemos derecho a despreciarlo argumentando que no fue racional... Aquí tenemos un acto irreduciblemente pragmático, y ello precisamente porque no es un acto de la razón, sino un acto de la voluntad, un acto de elección colectiva. El problema que afrontamos no es un problema de justificación, sino de participación en el poder, de quién toma las decisiones acerca de lo que está permitido y lo que no <sup>11</sup>.

En tal caso, si una norma es justa o no depende del acuerdo fáctico. Por consiguiente, dicha norma no es incondicionalmente válida, sino válida sólo relativamente a una específica comunidad en un tiempo específico. Si esto es así, como Tugendhat argumenta convincentemente, el significado de los discursos morales no puede ser cognitivo, pues lo que tales discursos permiten no es la *imparcialidad del juicio*, sino más bien la *formación de la autonomía*

<sup>11</sup> E. Tugendhat, *Problemas de ética*, Barcelona, Crítica, 1988, pp. 138 y ss. (Phil. Aufsätze, 311).

*de la voluntad*. Contra la interpretación de Tugendhat, Habermas ha insistido siempre en el significado cognitivo de los discursos morales, puesto que éstos permiten conocimiento moral (y, por ende, un consenso racionalmente motivado) y no meramente decisiones bajo condiciones de juego limpio (es decir, compromisos)<sup>12</sup>. Semejante interpretación antirrealista del rol de los discursos morales haría imposible cualquier defensa del supuesto anterior relativo al cognitivismo moral.

Teniendo en cuenta que la ética del discurso es, ante todo, un intento de defender el cognitivismo moral, la caracterización mencionada anteriormente de las condiciones ideales de los discursos morales sólo puede conservar un sentido cognitivista si la corrección moral se entiende en el primer sentido (es decir, el sentido realista) y se intenta *explicar en términos epistémicos* mediante la noción de «aceptabilidad racional en condiciones ideales» (en lugar de reducirla a «aceptación fáctica en condiciones de juego limpio»). Como se ha dicho anteriormente, bajo esta caracterización la aceptabilidad racional en condiciones ideales no puede interpretarse en un sentido operacional, sino sólo como una noción de éxito. Esto es, una norma sólo está justificada (es racionalmente aceptable en condiciones ideales) si de hecho es moralmente correcta. Sin embargo, la introducción de esta restricción ulterior respecto a lo que puede considerarse racionalmente aceptable bajo condiciones ideales tiene consecuencias muy destructivas para la ética del discurso.

Con respecto a la fuerza explicativa de semejante enfoque, la primera dificultad es que, al seguir dicha estrategia, la ética del discurso se volvería completamente vacía. Esto sería así al menos por dos razones. Por un lado, tal explicación discursiva de la «corrección moral» sería trivialmente correcta en la medida que la aceptabilidad racional en condiciones ideales *incluye la corrección moral como condición*. Así, el principio moral de universalización afirma que una norma es moralmente correcta si en condiciones ideales pudiera ser aceptada por todos los afectados en base a razones que establecen su corrección moral. Por otro lado, sin embargo, no hay posibilidad alguna de dar una interpretación discursiva de la tercera de esas restricciones. Como sabemos, en este planteamiento los discursos morales son caracterizados en términos formales y procedimentales. De acuerdo con ello, las condiciones discursivas pueden garantizar como máximo la corrección del *procedimiento*, pero no la del *contenido* (es decir, la corrección de los argumentos empleados, de los criterios sustanciales de justificación aceptados en un tiempo dado, etc.). Por esta razón, la idealización de la aceptabilidad racional que haría posible garantizar la corrección moral *desde el punto de vista del contenido* no puede explicarse en términos discursivos. No tenemos la menor idea de qué condiciones serían aquellas bajo las cuales los participantes en discursos morales serían no sólo intersubjetivamente convincentes, sino también infalibles. La anticipación de corrección requeriría que los participantes en los discursos morales estuvieran en

<sup>12</sup> Véase *DE*, pp. 78 y ss., y 88 y ss.

una situación epistémica que trascendiera su condición humana. Esto es, tendrían que tener un conocimiento infinito de las consecuencias y efectos colaterales de una norma, de la amplitud de los procesos de aprendizaje posibles en la argumentación moral y con relación a los intereses humanos mismos, etc.

Sin embargo, con independencia de si tal idealización podría tener contenido alguno, el problema de la ética del discurso seguiría ahí. Pues una vez que tal anticipación fuera permitida —aun cuando sólo contrafácticamente— como un requisito normativo plausible de nuestras pretensiones de corrección moral todas las restricciones que proceden del sentido operacional (es decir, discursivo) de la aceptabilidad racional serían *superfluas*. Si lo aceptable racionalmente en condiciones ideales no es lo justificado, en el sentido de lo que resulta convincente para todos, sino lo *efectivamente correcto*, entonces no está claro cuál sería el propósito de someterse, adicionalmente, a requerimientos discursivos. Una vez que uno tiene razón, uno no necesita *además* ser intersubjetivamente convincente o cualquier otra cosa. Las condiciones de aceptabilidad racional capaces de garantizar la corrección de nuestros juicios morales, cualesquiera que fueran, harían innecesarias las condiciones discursivas de aceptabilidad racional para la práctica de revisar tales juicios. Dicho en otras palabras, las condiciones discursivas son necesarias por razones epistémicas que no se darían si la aceptabilidad racional (en el sentido de una noción de éxito) fuera, en efecto, la anticipación inherente a nuestras pretensiones de corrección moral: si la infalibilidad pudiera garantizarse epistémicamente, cualquier requerimiento discursivo adicional sería superfluo.

En este sentido, los requerimientos discursivos para la aceptabilidad racional pueden justificarse como condición *necesaria* para una práctica razonable de revisión de nuestros juicios morales sólo en la medida en que ninguna condición *suficiente* para garantizar la corrección moral nos es epistémicamente accesible. Precisamente porque ningún criterio epistémico puede garantizar *per se* el éxito de dicha práctica necesitamos un criterio operacional para evaluar *aquí y ahora* si una norma dada se puede considerar moralmente correcta o no. Precisamente tal regla *operacional* de argumentación es lo que la ética del discurso pretende ofrecer con su principio de universalización «(U)». Como Habermas observa:

Como regla de argumentación el principio de universalización tiene que poseer un sentido racional, y, por tanto, operacional, para sujetos finitos que juzgan desde sus respectivos contextos. Por ello únicamente es lícito exigir que al fundamentar una norma se tengan en cuenta las consecuencias y los efectos secundarios que, en base a la información y a las razones disponibles en un momento dado, sea *previsible* que se deriven del seguimiento universal de dicha norma para los intereses de todos y cada uno (*EDE*, pp. 139 y 146 y ss.).

En este sentido operacional, sin embargo, los resultados *obviamente no son infalibles*. Ahora bien, si el procedimiento permite el error, sus resultados no pueden considerarse moralmente correctos *per se*. Satisfacer las condiciones

procedimentales no *constituye el significado de «moralmente correcto»*. Simplemente ofrece el mejor apoyo epistémico al supuesto de que las normas en cuestión son *de hecho* moralmente correctas. Tal procedimiento no puede *garantizar* la corrección moral de una norma —nada puede hacerlo—. Pero sí puede *autorizarnos* a afirmar la corrección de una norma mientras que no aparezca ningún contraargumento (sea debido a nuevas experiencias, consecuencias, o efectos colaterales o a cualesquiera procesos de aprendizaje en general). Como Habermas observa correctamente: «la pretensión de validez de una norma que ha pasado el test de universalización lleva un “subíndice temporal y cognitivo”» (*EDE*, pp. 139 y 147). Si esto es así, la única pretensión de validez a la que se puede estar haciendo referencia aquí es la pretensión de aceptabilidad (o justificación) racional. Pues la pretensión de corrección moral, debido a su validez incondicional, es incompatible con cualquier subíndice temporal.

En esta dificultad se muestra que la lectura de la ética del discurso que Habermas está intentando defender es esencialmente inestable. O bien se desarrolla en una posición coherentemente realista o en una posición coherentemente antirrealista. Como indiqué al principio, siguiendo una estrategia antirrealista consistente habría que afirmar que la corrección moral no puede trascender la justificación porque es exclusivamente una función de nuestras actitudes epistémicas. En este caso, los resultados de los discursos *fácticos* que satisfacen las condiciones discursivas requeridas establecen la corrección moral de las normas, cualesquiera que dichas normas sean. Como ya vimos, esta estrategia conduce inevitablemente a un enfoque decisionista<sup>13</sup>. La otra opción es reconocer que la justicia o corrección moral de las normas no es una función exclusivamente de nuestras actitudes epistémicas (incluso en condiciones discursivas ideales), puesto que ésta depende también, como Habermas apuntaba antes, de «la información y las razones disponibles en un momento determinado». Sin embargo, con este reconocimiento ya se ha concedido todo lo que requiere una estrategia realista. Pues en este caso *la corrección moral trasciende inevitablemente la justificación*. Bajo esta interpretación, la justicia de las normas depende de condiciones epistémicas y *no epistémicas*. Requiere la aceptación de todos los afectados como condición necesaria para que podamos reconocer una norma como justa, pero la aceptación misma depende de una valoración correcta de las circunstancias no epistémicas que constituyen la condición suficiente para que una norma sea justa.

Hay muchas *circunstancias* no epistémicas que forman parte de las razones sustantivas que empleamos para justificar la justicia de las normas, específicamente aquellas relacionadas con las consecuencias reales y los efectos co-

<sup>13</sup> Habermas se aproxima peligrosamente a un enfoque decisionista en *Verdad y Justificación* cuando remonta el supuesto de «una única respuesta correcta» implícito en nuestros juicios morales a una «decisión... de codificar binariamente la corrección moral como una pretensión de validez análoga a la verdad» (*WR*, p. 316). Aunque él mismo puntualiza que «hablar acerca de una “resolución” o “decisión” apunta en la dirección equivocada» (*ibid.*, p. 317), no ofrece ninguna explicación alternativa de esta presuposición crucial del enfoque cognitivista de la moral.

laterales que de hecho se seguirán de la adhesión general a una norma particular. Ya sólo debido a dichas circunstancias es justificado calificar los discursos morales de cognitivos. Pero hay muchos otros factores que también explican por qué, en relación a nuestros juicios morales, nuestras prácticas discursivas están sujetas a condiciones objetivas que pueden trascender nuestra situación epistémica dada.

En primer lugar, parece que en relación a la cuestión de cuáles son nuestros intereses irrenunciables creemos que la respuesta depende menos de «nuestro hacer» que de nuestra condición humana. En este sentido, parece que reconocemos una (posible) discrepancia entre nuestros intereses irrenunciables y la evidencia que podamos tener de ellos en un momento determinado. Si seguimos buscando aquellas normas que de hecho están igualmente en interés de todos, a pesar de todas las dificultades para superar nuestros desacuerdos epistémicos, sólo puede ser porque, dada cualquier sospecha relativa a si una norma es justa, *la opción de cambiar los intereses afectados en lugar de la norma no es realmente viable*.

Pero incluso bajo el supuesto de que el acceso a *nuestros intereses irrenunciables* nos resultara completamente transparente, ello no constituiría garantía alguna relativa al conocimiento de los intereses irrenunciables de los demás. Sin embargo, de dicho conocimiento depende obviamente la condición de estar «igualmente en interés de todos y cada uno» requerida para la corrección moral.

Por otra parte, incluso admitiendo que nunca pudiéramos equivocarnos con relación a nuestros intereses irrenunciables, eso no podría garantizar un conocimiento infalible relativo a las consecuencias objetivas que una norma tendría a largo plazo y en circunstancias cambiantes (es decir, normalmente impredecibles) para todos los posibles afectados. La discriminación no es siempre una consecuencia de la voluntad reprimida de los afectados por ella. Puede también deberse a nuestra incapacidad para prever los efectos colaterales de una norma a largo plazo, o incluso a nuestra incapacidad para imaginar una norma más satisfactoria, a pesar de tener las mejores intenciones. No parece absurdo admitir la posibilidad de que descubriéramos que una norma a la que asentimos racionalmente (sobre la base de nuestra situación epistémica dada), ha resultado ser moralmente incorrecta (es decir, *de hecho* injusta, discriminatoria, etc.).

Pero una vez que se reconoce que el conocimiento de hechos es un componente esencial de las razones sustantivas que se emplean en el discurso moral para determinar la justicia de las normas no hay ninguna razón obvia para insistir en la única pretensión que distingue la lectura realista de la lectura antirrealista de la ética del discurso, a saber, la pretensión de que nuestras actitudes epistémicas en condiciones discursivas ideales son infalibles. Podemos abandonar esta dudosa pretensión y mantener una actitud falibilista con respecto a nuestros juicios morales tal como hacemos con respecto a cualquier otro juicio cognitivo.