

Self-Constitution in the Ethics of Plato and Kant

Christine M. Korsgaard

Autoconstitución en la ética de Platón y Kant*

Traducción de **Javier Fuentes González** y **Eva Monardes Pereira**

Universidad de Bonn — Universidad de Chile

jfuentesg10@gmail.com — emonardes@ug.uchile.cl

Publicado originalmente como Korsgaard, C. M. (1999). Self-Constitution in the Ethics of Plato and Kant. *The Journal of Ethics*, 3(1), 1–29.

Los traductores agradecemos sinceramente a Christine Korsgaard por la ayuda que nos brindó para obtener los derechos de traducción del presente artículo.

* He discutido este artículo, o el manuscrito inédito más extenso del cual procede, con los auditores de la reunión inaugural de The Society for Ethics, la Universidad de Ámsterdam, la Universidad de Constanza, la Universidad Humboldt de Berlín, la Universidad de Pittsburgh, la Universidad de Virginia, la Universidad de Salzburgo, la Universidad de Toronto, la Universidad York de Toronto y la Universidad de Zúrich. Agradezco a todos los auditores su interés, sus perspicaces comentarios y sus desafiantes preguntas. Me gustaría agradecer a Charlotte Brown, Barbara Herman, Govert den Hartogh, Anton Leist, Richard Moran, Amelie Rorty y Theo van Willigenburg por leer y comentar el manuscrito.



RESUMEN

Platón y Kant plantean un modelo “constitucional” del alma, en el que la razón y el apetito o la pasión tienen diferentes roles estructurales y funcionales en la generación de la motivación, en contraposición al común “modelo combativo,” en el que son mostrados como fuentes de motivación independientes que luchan por el control. Desde el punto de vista del modelo constitucional, podemos explicar qué hace que una acción sea diferente de un evento. Lo que hace que una acción sea atribuible a una persona y, por lo tanto, lo que la hace una acción, es que surge desde la constitución de la persona y, por lo tanto, desde la persona como un todo, y no desde una fuerza que actúa sobre o en ella. Esto, a su vez, implica una explicación de lo que hace buena a una acción: lo que hace que una acción sea buena es que se delibere sobre ella y se la elija de una manera que unifique a la persona en un sistema constitucional. A través de la acción deliberativa nos constituimos como agentes unificados. La justicia platónica y el imperativo categórico de Kant se muestran como estándares normativos para la acción porque son principios de autoconstitución.

Palabras clave: acción, autonomía, imperativo categórico, constitución, deliberación, estándar interno, justicia, Kant, normatividad, Platón, universalizabilidad, voluntad

ABSTRACT

Plato and Kant advance a “constitutional” model of the soul, in which reason and appetite or passion have different structural and functional roles in the generation of motivation, as opposed to the familiar “Combat Model” in which they are portrayed as independent sources of motivation struggling for control. In terms of the constitutional model we may explain what makes an action different from an event. What makes an action attributable to a person, and therefore what makes it an action, is that it issues from the person’s constitution, and therefore from the person as a whole, rather than from some force working on or in the person. This in turn implies an account of what makes an action good: what makes an action good is that it is deliberated upon and chosen in a way that unifies the person into a constitutional system. Through deliberative action we constitute ourselves as unified agents. Platonic justice and Kant’s categorical imperative are shown to be normative standards for action because they are principles of self-constitution

Keywords: action, autonomy, categorical imperative, constitution, deliberation, internal standard, justice, Kant, normativity, Plato, universalizability, will

I. Introducción

Una de las secciones más famosas del *Tratado* de David Hume comienza con estas palabras:

Nada es más habitual en la filosofía, e incluso en la vida común, que hablar del combate entre pasión y razón, dar la preferencia a la razón y afirmar que los hombres sólo son virtuosos en la medida en que se conforman a sus dictados. Toda criatura racional, se dice, está obligada a regular sus acciones por medio de la razón; y si cualquier otro motivo o principio desafía la dirección de su conducta, debe oponerse a él hasta que sea completamente sometido, o al menos llevado a una conformidad con ese principio superior.¹

Según como Hume entiende estas afirmaciones, razón y pasión son dos fuerzas en el alma, cada una es una fuente de motivos para actuar, y la virtud consiste en que la persona esté de acuerdo con la razón. ¿Por qué debería la persona hacer eso? Hume nos dice que en la filosofía:

La eternidad, la invariabilidad y el origen divino de [la razón] han sido mostrados de la mejor manera: se ha insistido con la misma fuerza en la ceguera, la inconstancia y el engaño de [la pasión].²

Hume se propone “mostrar la falacia de toda esta filosofía,” pero en su demostración no rechaza exactamente lo que llamaré el modelo combativo. Simplemente sostiene que la razón no es una fuerza y, por lo tanto, que no hay combate.

Creo que hay algunas cuestiones que Hume debería haber planteado en primer lugar, pues el modelo combativo tiene muy poco sentido. Desde la perspectiva de tercera persona, efectivamente a veces explicamos las acciones de una persona como resultado de que un motivo sea “más fuerte” que otro, por ejemplo, cuando la persona tiene pasiones en conflicto. Pero, ¿es la diferencia entre razón y pasión más o menos la misma que la diferencia entre una pasión y otra? ¿Y son las acciones de una persona meramente el resultado del juego, o más bien del combate, entre estas fuerzas dentro de ella? ¿En qué se diferenciarían, entonces, las acciones de los rubores, o de los espasmos, o incluso de los procesos biológicos?

¹ David Hume, *A Treatise of Human Nature*, 2da. edición, editado por L. A. Selby-Bigge y P. H. Nidditch (Oxford: Clarendon Press, 1978), Libro II, Parte III, Sección III, p. 413. [*Hemos respetado el formato de citación del artículo original y hemos traducido las citas de autores clásicos también a partir de éste. (N. de los T.)*]

² Hume, *A Treatise of Human Nature*, Libro II, Parte III, Sección III, p. 413.

Ahora bien, podemos intentar resolver este último problema trayendo de vuelta al panorama a la persona, al agente — la acción es diferente de otros movimientos físicos porque la persona *elige* seguir ya sea la razón o la pasión. Pero esto hace al modelo combativo aún más desconcertante. Pues, ¿cuál es la esencia de esta persona, en la que tanto la razón como la pasión son fuerzas, *ninguna* de las cuales se identifica con la persona misma, y entre las cuales ésta debe elegir? Y si la persona no se identifica ni con la razón ni con la pasión, entonces, ¿cómo — según qué principio — puede elegir entre ellas? Los filósofos que Hume describe aquí parecen imaginar que la persona elige entre la razón y la pasión valorando sus méritos — la razón es divina y fiable, la pasión ciega y engañosa. Pero aquello ciertamente presupone que la persona *ya* se identifica con la razón, que es lo que valora méritos. Pero, ¿cómo podría, entonces, la persona elegir la pasión por sobre la razón? El modelo combativo no nos da una imagen clara de la *persona* que elige entre razón y pasión.

La tradición nos proporciona otro modelo para la interacción entre razón y pasión en el alma que tiene más sentido, porque les asigna diferencias funcionales y estructurales.³ Lo llamo el modelo constitucional, porque su aparición más clara está en la *República* de Platón, donde el alma humana se compara con la constitución de una *polis* o ciudad-estado. Creo que el modelo constitucional tiene importantes implicancias para la filosofía moral, y mi tarea en este artículo es explicitar estas implicancias. Específicamente, el modelo constitucional implica una cierta concepción de lo que es una *acción*, lo que a su vez tiene implicancias sobre lo que hace que una acción sea buena o mala. Es un tanto difícil articular claramente estas implicancias antes de la discusión, pero la idea principal es ésta: lo que distingue a la acción del mero comportamiento y de otros movimientos físicos es que *posee un autor* — es atribuible de un modo muy especial a la *persona* que la hace, y con ello quiero decir, a *toda* la persona. El modelo constitucional nos dice que lo que hace que una acción sea tuya de este modo es que surge desde y está de acuerdo con tu constitución. Pero también proporciona un estándar para la acción buena, un estándar que nos dice qué acciones son las más verdaderamente propias de una persona y, por lo tanto, qué acciones son las más verdaderamente *acciones*. Ahora bien, ésta es la parte

³ Se podría pensar que Hume también presenta un modelo constitucional, ya que su propio argumento sugiere que la función de la pasión es determinar nuestros fines y que la función de la razón es descubrir medios para fines. Sin embargo, en otro lugar he sostenido que Hume no cree realmente en un principio de *razón práctica* instrumental que nos instruya para tomar los medios para nuestros fines y que sería necesario para integrar las dos funciones (la determinación del fin y la identificación de los medios) en *un único sistema que produzca acciones*. Por ello, Hume es incapaz de desarrollar una *persona* a partir de estos escasos recursos. Lo que acabo de decir quedará más claro a medida que avance este ensayo, pues en realidad es, en cierto sentido, una versión breve de la discusión completa de este ensayo. Para el argumento según el cual Hume no cree en un principio de razón práctica instrumental, véase mi “The Normativity of Instrumental Reason” en Garrett Cullity y Berys Gaut (eds.), *Ethics and Practical Reason* (Oxford: Clarendon Press, 1997), pp. 215-254, especialmente pp. 220-234.

que es difícil de decir antes de la discusión: las acciones que son las más verdaderamente propias de una persona son precisamente aquellas acciones que la unifican del modo más completo y que, por lo tanto, la constituyen del modo más completo como su autora. Son aquellas acciones que provienen desde ella y le dan el tipo de unidad volitiva que debe tener si le vamos a atribuir la acción como a una persona completa. Lo que hace que una acción sea mala, por el contrario, es que surge en parte no desde la persona, sino desde algo que opera *en* ella o *sobre* ella, algo que amenaza su unidad volitiva. Resumo estas afirmaciones diciendo que, según el modelo constitucional, la acción es autoconstitución.

II. Platón

En el libro I de la *República*, Sócrates y sus amigos discuten la pregunta sobre qué es la justicia. La discusión es interrumpida por Trasímaco, quien afirma que la mejor vida es la vida injusta, la vida vivida por los fuertes, quienes imponen las leyes de la justicia a los débiles, pero ellos mismos desatienden aquellas leyes. Mientras más completamente injusto seas, dice Trasímaco, mejor vivirás, pues los carteristas y los ladrones, quienes cometen pequeñas injusticias, son castigados, mientras que los tiranos, quienes esclavizan a ciudades enteras y roban sus tesoros, llevan una vida gloriosa y son la envidia de todos (336b-334d).⁴ Sócrates, distraído por estas afirmaciones, abandona la discusión sobre qué es la justicia y aborda la pregunta sobre si es mejor la vida justa o la injusta.

Sócrates procede a construir tres argumentos concebidos para demostrar que la vida justa es la mejor. El que es central para mi propia discusión es el siguiente (351b-352c): Sócrates pregunta a Trasímaco si una banda de asaltantes y ladrones con un propósito injusto común sería capaz de lograr ese propósito si fueran injustos entre sí. Trasímaco está de acuerdo con que no podrían hacerlo. La justicia, como dice Sócrates, es lo que proporciona un sentido de propósito común a un grupo, mientras que la injusticia causa odio y guerra civil, y hace al grupo “incapaz de lograr nada como unidad” (352a). Trasímaco es entonces conducido a estar de acuerdo con que la justicia y la injusticia tienen el mismo efecto dondequiera que ocurran y, por lo tanto, el mismo efecto dentro del alma humana individual que dentro de un grupo. La injusticia, por lo tanto, hace a un individuo “incapaz

⁴ Las referencias a la *República* de Platón se insertan en el texto usando los números estándar de Stephanus colocados en los márgenes de la mayoría de las ediciones y traducciones de las obras de Platón. Las citas directas proceden de la traducción de G. M. A. Grube, revisada por C. D. C. Reeve, que puede encontrarse en *Plato: Complete Works*, editado por John M. Cooper (Indianápolis: Hackett Publishing Company, 1997).

de lograr nada, porque se encuentra en un estado de guerra civil y no de una única mente.” Mientras más completa sea esta condición, peor resulta, ya que, según Sócrates, “los que son del todo malos y completamente injustos son completamente incapaces de lograr nada” (352c).⁵

Ahora bien, no hay nada evidentemente incorrecto en este argumento, excepto, por supuesto, que va en contra del hecho de que parecemos ver personas injustas por todos lados, haciendo y logrando cosas a diestra y siniestra. Entonces, ¿de qué está hablando Sócrates? El argumento deja a la audiencia de Sócrates desconcertada e insatisfecha. Entonces, los hermanos de Platón, Glaucón y Adimanto, exigen a Sócrates que vuelva a la pregunta abandonada, qué es la justicia y qué efecto tiene en el alma. Es esta exigencia la que hace emprender a Platón su intento de identificar la justicia con un objeto más grande y visible, la ciudad ideal, y su famosa comparación entre la constitución de la ciudad y la constitución del alma.

Será de ayuda repasar los principales elementos de esa comparación. Platón identifica tres clases en la ciudad. En primer lugar, están los gobernantes, quienes hacen las leyes y políticas para la ciudad y se encargan de sus relaciones con otras ciudades. En segundo lugar, están los auxiliares, una especie de combinación entre fuerza militar y policial, quienes hacen cumplir las leyes dentro de la ciudad y también la defienden de los enemigos externos, siguiendo las órdenes de los gobernantes. Los gobernantes proceden de las filas de estos auxiliares, y a los dos grupos en conjunto se les llama guardianes. Y, por último, están los granjeros, artesanos, mercaderes, etc., quienes se ocupan de las necesidades de la ciudad.

Las virtudes de la ciudad ideal se identifican luego con ciertas propiedades de estas partes y sus relaciones entre sí. La sabiduría de la ciudad descansa en la sabiduría de sus gobernantes (428b-429a). Al principio no se nos dice mucho al respecto, excepto que los gobernantes de la ciudad ideal, a diferencia de los gobernantes de Trasímaco, gobiernan en vistas del bien de la ciudad como un todo y no del suyo propio. La valentía de la ciudad descansa en la valentía de sus auxiliares, que se identifica con su capacidad de preservar ciertas creencias, inculcadas a ellos por los gobernantes, sobre qué hay que temer ante la tentación, el placer, el dolor y el miedo mismo (429a-430c). La *sophrosyne* de la ciudad — su moderación o templanza — descansa en el acuerdo de todas las clases de la ciudad sobre quién

⁵ Los otros dos argumentos son el argumento de la “superación”, usado para establecer que la justicia es una forma de virtud y conocimiento (349a-350d), y el argumento de la función, usado para establecer que la persona justa es la más feliz (352d-354a).

debe gobernar y ser gobernado (430e-432b). Y su justicia descansa en el hecho de que cada clase de la ciudad hace su propio trabajo y nadie trata de entrometerse en el trabajo de los demás (433a y ss.).

Platón se encarga entonces de encontrar las mismas tres partes en el alma humana. El modelo constitucional, al igual que el modelo combativo, parte desde la experiencia del conflicto interior. Sócrates plantea como principio que, si encontramos en el alma actitudes o reacciones opuestas ante una única cosa al mismo tiempo, debemos suponer que el alma tiene partes (436b-c). Por ejemplo, el alma de una persona sedienta es impulsada a beber por su sed. Entonces, si el alma en ese preciso momento rechaza beber, tiene que ser con una parte diferente. Y ésta es una experiencia que la gente realmente tiene: hay gente sedienta que decide no beber. Esto sucede cuando juzgan que la bebida será mala para ellos. Como dice Sócrates:

¿No hay algo en su alma que les ordena beber y algo diferente que les prohíbe hacerlo, que prevalece sobre la cosa que les ordena?... ¿No entra en juego en tales casos lo que prohíbe como resultado de un cálculo racional? (439c-d)

Entonces, la razón y el apetito deben ser dos partes del alma distintas.

De hecho, no obstante, el énfasis por parte Sócrates en el conflicto es un tanto confuso, ya que, incluso si no hay conflicto, se pueden discernir dos partes del alma. Supóngase, en cambio, que la bebida no tiene nada malo y que la persona que está sedienta efectivamente bebe. En esta clase de casos, Sócrates dice,

el alma de alguien que tiene apetito por algo quiere aquello por lo que tiene apetito y toma para sí lo que es su voluntad tener, y... en la medida en que desea que se le dé algo, su alma, dado que desea que esto ocurra, asiente a ello como en respuesta a una pregunta (437c)

El alma no actúa directamente a partir del apetito, sino a partir de algo que respalda al apetito y le dice que sí. Incluso cuando el conflicto está ausente, podemos ver entonces que hay dos partes del alma.

Sócrates argumenta a continuación que existe una tercera parte, el *thymos* o espíritu, que es distinta tanto de la razón como del apetito, aunque es el aliado natural de la razón (439e-441c). Se muestra que es distinto del apetito en el hecho de que la ira y la indignación, que son manifestaciones del espíritu, se dirigen a menudo contra los apetitos mismos. Esto es ilustrado por la historia de

Leoncio, quien se disgustó consigo mismo por querer mirar algunos cadáveres y reprendió a sus propios ojos por sus apetitos malignos (439e-440a). Sócrates afirma que el espíritu siempre lucha del lado de la razón en caso de conflicto entre razón y apetito. Sin embargo, es distinto de la razón, ya que está presente en los niños pequeños y en los animales, que no tienen razón; y, además, a veces necesita ser controlado por la razón (440e-441c).

Con estos argumentos, Sócrates establece que el alma tiene las mismas tres partes que la ciudad. La razón corresponde a los gobernantes y su función es dirigir cosas por el bien de la persona completa. El espíritu corresponde a los auxiliares y su función es ejecutar las órdenes de la razón. Los apetitos corresponden al resto de los ciudadanos y su tarea es abastecer a la persona con lo que sea que necesite.

Ahora bien, si el alma tiene partes, la pregunta que va a surgir es qué las hace una, qué las unifica en una única alma. Y parte de la respuesta es que las partes del alma deben estar unificadas — *necesitan* estar unificadas, como las personas en una ciudad — para poder actuar. Específicamente, podemos ver las tres partes del alma como correspondiendo a tres partes de una acción deliberativa. La acción deliberativa comienza a partir del hecho de que tenemos ciertos apetitos y deseos. Somos conscientes de estos, y nos invitan a realizar determinadas acciones o a buscar ciertos fines. Sin embargo, dado que somos racionales, no actuamos automáticamente según nuestros apetitos y deseos, sino que, en cambio, decidimos si satisfacerlos o no. Como dijo Sócrates en un pasaje que vimos hace poco, es como si lo que hace el apetito fuese hacerle una petición a la razón, y la razón dice sí o no. Y, por último, está la ejecución de la decisión — efectivamente hacer lo que hemos decidido hacer. Ya que, por cierto, no siempre hacemos lo que hemos decidido hacer, sino que a veces somos distraídos del curso que nos hemos propuesto por el placer o el dolor o el miedo. Entonces, podemos identificar tres partes de una acción deliberativa, correspondientes a las tres partes del alma de Platón, a saber:

El apetito hace una propuesta.

La razón decide si actúa según ésta o no.

El espíritu ejecuta la decisión de la razón.

Esta línea de pensamiento apoya la analogía de Platón entre ciudad y alma. Pues una ciudad también emprende acciones deliberativas: no es sólo un lugar para vivir, sino más bien un tipo de

agente que realiza acciones y, por lo tanto, tiene una vida y una historia. Y podemos ver las mismas tres partes en una decisión política. La gente de la ciudad hace una propuesta: dicen que hay algo que necesitan. Piden carreteras, o una mejor asistencia médica, o más protección policial. Luego, los gobernantes deciden si actúan según la propuesta o no. Dicen “sí” o “no” al pueblo. Y luego los auxiliares ejecutan las decisiones del gobernante.

De hecho, el principal objetivo de una constitución política escrita es precisamente establecer el modo de acción deliberativa de la ciudad, los procedimientos según los que sus decisiones colectivas se deben tomar y ejecutar. Una constitución define una serie de roles y cargos que, en conjunto, constituyen un procedimiento para la acción deliberativa, diciendo quién realizará cada paso y cómo se hará. Establece las formas apropiadas de hacer propuestas (por ejemplo, a través de una petición, o la presentación de proyectos de ley, o lo que sea), de decidir si se actúa según estas propuestas (la función legislativa), y de ejecutar las decisiones resultantes (la función ejecutiva). Y en cada caso dice quién está autorizado a ejecutar los procedimientos que ha especificado. De este modo, la constitución hace posible que los ciudadanos funcionen como un único agente colectivo.

Y esto explica la desconcertante definición de justicia por parte de Sócrates. La justicia, dice, es “hacer el trabajo propio y no entrometerse en el que no es propio” (433a-b). Cuando Sócrates introduce este principio por primera vez en la discusión (369e y ss.), se refiere a la especialización del trabajo, y eso es sobre lo que el principio parece tratar.⁶ Si pensamos que la constitución establece los procedimientos para la acción deliberativa, y los roles y cargos que constituyen esos procedimientos, podemos ver cuál es el punto de Sócrates. Pues usurpar el cargo de otro en los procedimientos constitucionales para la acción colectiva es *precisamente* lo que entendemos por injusticia, o al menos es una de las cosas que queremos decir con ello. Por ejemplo, si la constitución dice que el presidente no puede hacer la guerra sin el acuerdo del congreso y sin embargo la hace, ha usurpado el rol del congreso en esta decisión, y eso es injusto. Si la constitución dice que cada ciudadano puede emitir un voto en las elecciones, y mediante algún fraude logras votar más de una vez, estás disminuyendo la voz de los otros en la elección, y eso es injusto. Entonces, la injusticia, en uno de sus sentidos más familiares, es usurpar el rol de otro en los procedimientos deliberativos que definen la acción colectiva. Es entrometerse en el trabajo de otro.

⁶ Sócrates no sólo reconoce abiertamente esta rareza más adelante, sino que de hecho sugiere que el principio de la especialización del trabajo es “beneficioso” porque es “una especie de imagen de la justicia” (443c).

Dije en un sentido, pues esto es en gran medida lo que a veces se denomina una concepción *procesal* de la justicia, en contraposición a una concepción *sustantiva*. Esta distinción representa una importante tensión en nuestro concepto de justicia y una causa constante de confusión sobre la fuente de su normatividad. Por un lado, la idea de justicia implica esencialmente la idea de seguir ciertos procedimientos. En el estado, como he venido diciendo, estos son los procedimientos que la constitución establece para la acción deliberativa colectiva: para hacer leyes, librar guerras, juzgar casos, recaudar impuestos, distribuir servicios, y todas las diversas cosas que hace un estado. De acuerdo con la concepción procesal de la justicia, una acción del estado es justa si y sólo si es el resultado de seguir efectivamente estos procedimientos. Ésta es una *ley* que ha sido aprobada en forma por un cuerpo legislativo debidamente constituido; esta ley es *constitucional* si (por ejemplo) la corte suprema dice que lo es; una persona es *inocente* de un cierto crimen cuando ha sido considerada así por un jurado; alguien es *el presidente* si cumple las cualificaciones legales y ha sido debidamente electo, etc. Todos estos son juicios normativos — los términos *ley*, *constitucional*, *inocente* y *presidente* implican todos la existencia de ciertas razones para la acción — y su normatividad *deriva de* la ejecución de los procedimientos que los han establecido.

Por otro lado, sin embargo, ciertamente hay casos en los que tenemos una idea independiente de qué resultado deben generar los procedimientos. Estas ideas independientes sirven como criterios para nuestros juicios más sustantivos — en algunos casos, de lo que es justo, en otros, simplemente de lo que es correcto o mejor. Y estos juicios sustantivos pueden entrar en conflicto con los resultados reales de la ejecución de los procedimientos. Tal vez la ley sea inconstitucional, aunque el cuerpo legislativo la haya aprobado; quizás el acusado es culpable, aunque el jurado lo haya dejado libre; quizás el candidato electo no sea la mejor persona para el trabajo, o ni siquiera la mejor de las que se presentaron, o quizás no represente realmente la voluntad de la mayoría debido a accidentes en la participación electoral. Como muestra este último ejemplo, la distinción entre lo procesalmente justo y lo sustantivamente justo, correcto o mejor, es una distinción rudimentaria y relativa al caso bajo consideración. ¿Quién debería ser electo? La mejor persona para el trabajo, la mejor de aquellas que efectivamente se presentan, la preferida por la mayoría de los ciudadanos, la preferida por la mayoría de los votantes registrados, la electa por la mayoría de aquellos que efectivamente concurren el día de la elección... A medida que descendemos en la lista, la respuesta a la pregunta se vuelve cada vez más procesal; la respuesta superior a ésta es, relativamente, más sustantiva. Podemos intentar diseñar nuestros procedimientos para garantizar el resultado sustancialmente correcto, mejor o justo. Pero —

y aquí está el punto importante — según la concepción procesal de justicia, la normatividad de estos procedimientos, no obstante, no surge de la eficiencia, la bondad, *o incluso la justicia sustantiva* de los resultados que producen. Es cierto lo contrario: son los procedimientos mismos — o más bien la ejecución efectiva de los procedimientos — los que confieren normatividad a aquellos resultados. La persona que resulta electa ocupa el cargo, sin importar cuán lejos esté de ser la mejor persona para este trabajo. La absolución del jurado se mantiene, aunque posteriormente descubramos nueva evidencia de que, después de todo, el acusado era culpable.

Ahora bien, si la normatividad de los resultados surge de la ejecución de los procedimientos, ¿de dónde, podemos preguntar, puede venir la normatividad de los procedimientos mismos? Y aquí nos topamos con la causa de confusión que mencioné al comienzo, pues hay una tentación constante de creer que los procedimientos mismos deben derivar su normatividad de la buena calidad de sus resultados. Eso no puede ser correcto, como he dicho, pues si la normatividad de nuestros procedimientos proviniera de la calidad sustantiva de sus resultados, estaríamos dispuestos a dejar de lado esos procedimientos cuando supiéramos que sus resultados habrían de ser pobres. Y como he dicho hace poco, no hacemos aquello. Ahí donde haya procedimientos constitucionales, la corrección, la bondad, la excelencia o incluso la justicia sustantivas no son necesarias ni suficientes para el estatus normativo de sus resultados.

Quizás ahora podemos atrevernos a decir que lo que hace normativos a los procedimientos es la calidad *usual* de sus resultados, el hecho de que acierten la mayoría del tiempo. Después de todo, incluso si respaldamos los resultados de nuestros procedimientos, aunque en tal o cual caso sean malos, ciertamente cambiaríamos esos procedimientos si los resultados fueran malos *muy frecuentemente*. Pero ésta no puede ser la respuesta completa, no sólo porque no siempre es verdadera — piensa en el sistema de jurado — sino también porque, como los utilitaristas de acto llevan años diciéndonos, es irracional seguir un procedimiento meramente porque éste usualmente obtiene un buen resultado, cuando sabes que esta vez tendrá uno malo. Entonces, quizás deberíamos decir que la normatividad de los procedimientos proviene de la calidad usual de sus resultados *combinada* con el hecho de que debemos tener algunos procedimientos tales y debemos respaldar sus resultados. Pero, ¿por qué debemos tener procedimientos tales? Porque sin ellos la acción colectiva es imposible. Y ahora nos hemos convencido del punto de vista de Platón. Para actuar juntos — para hacer leyes y políticas, aplicarlas, hacerlas cumplir — de una manera que represente no a algunos de nosotros tiranizando a

otros, sino a todos nosotros actuando como una unidad — debemos tener una constitución que defina los procedimientos de la acción deliberativa colectiva, y debemos respaldar sus resultados.⁷

Según Platón, la fuerza normativa de una constitución *consiste* en el hecho de que hace posible que la ciudad funcione como un único agente unificado. Pues una ciudad sin justicia, según Platón, sobre todo carece de unidad — no es una ciudad, dice, sino muchas (422d-423c; véase también 462a-e). Cuando la justicia colapsa, la ciudad cae en guerra civil, ya que los gobernantes, los soldados y la gente luchan todos por el control. Los procedimientos deliberativos que unifican la ciudad como un único agente colapsan y la ciudad *como tal* no puede actuar. Los ciudadanos individuales y las clases de ésta aún pueden realizar diversas acciones, pero la ciudad no puede actuar como una unidad.

Y esto también aplica para la justicia y la injusticia dentro de la persona individual. Sócrates dice:

Quien es justo no permite que ninguna parte de sí mismo haga el trabajo de otra parte ni permite que las diversas clases que hay en él se entrometan una en otra. Regula bien lo que es realmente propio y se gobierna a sí mismo. Se pone a sí mismo en orden, es su propio amigo y armoniza las tres partes de sí mismo como tres notas límite en una escala musical — alta, baja y media. Une aquellas partes y cualquier otra que pudiese haber entre ellas, y a partir de haber sido muchas cosas se convierte completamente en una sola, moderada y armoniosa. Sólo entonces actúa. (443d-e)

Pero si la justicia es lo que hace posible que una persona funcione como un único agente unificado, entonces la injusticia lo hace imposible. La guerra civil estalla entre el apetito, el espíritu y la razón, cada uno de los cuales intenta usurpar los roles y los cargos de los otros. Los procedimientos deliberativos que unifican el alma en un único agente colapsan y la persona *como tal* no puede actuar. Entonces, el argumento de Sócrates en el Libro I resulta ser cierto. Los deseos e impulsos pueden operar dentro de la persona injusta, como los ciudadanos individuales pueden operar dentro del estado injusto. Pero la *persona* injusta es “completamente incapaz de lograr nada” (352c) porque la *persona* injusta no puede actuar en absoluto.

⁷ También he discutido estos puntos en “Taking the Law in Our Own Hands: Kant on the Right to Revolution,” en Andrews Reath, Barbara Herman y Christine M. Korsgaard (eds.), *Reclaiming the History of Ethics: Essays for John Rawls* (Cambridge: Cambridge University Press, 1997), pp. 308-309. La discusión que aquí se hace está tomada en gran parte de esa discusión.

III. Kant

Pasemos ahora a Kant. El mejor modo de ver que Kant está pensando en términos del modelo constitucional es considerar el argumento que utiliza para establecer que el imperativo categórico es la ley de una voluntad libre (G 4:446-448).⁸ Kant sostiene que, en la medida en que eres un ser racional, debes actuar bajo la idea de la libertad. Y una voluntad libre es aquella que no está determinada por ninguna causa ajena — por ninguna ley fuera de sí misma. No es, en el lenguaje de Kant, “heterónoma.” Pero Kant afirma que la voluntad libre debe estar determinada por una ley u otra — retomaré el argumento para ello en la sección VII — y, por lo tanto, debe ser “autónoma.” Es decir, debe actuar según una ley que se dé a sí misma. Y Kant dice que esto significa que el imperativo categórico *es precisamente* la ley de una voluntad libre.

Para ver por qué, sólo tenemos que considerar cómo debe deliberar una voluntad libre. Entonces, he aquí la voluntad libre, completamente autogobernante, sin nada fuera de ella dándole ninguna ley. Y llega una inclinación, y le presenta a la voluntad libre una propuesta. Ahora bien, las inclinaciones, según Kant, están fundadas en lo que él llama “incentivos,” que son las características de los objetos de aquellas inclinaciones que los hacen parecer atractivos y elegibles.⁹ Supongamos que el incentivo es que el objeto sea placentero. Entonces, la inclinación dice: el fin-F sería algo muy placentero de realizar. Entonces, ¿qué tal el fin-F? ¿No parece un fin-a-producir? Ahora bien, lo que la voluntad elige son, estrictamente hablando, acciones, entonces, antes de que la propuesta esté completa, tenemos que hacerla una propuesta para la acción. El razonamiento instrumental determina que podrías producir el fin-F haciendo el acto-A. Entonces, la propuesta es: que deberías hacer el acto-A para producir este tan placentero fin-F.

Ahora bien, si tu voluntad fuera heterónoma, y el placer fuera una ley para ti, esto es todo lo que necesitarías saber, e inmediatamente harías el acto-A para producir ese placentero fin-F. Pero dado

⁸ Las referencias a las obras de Kant se insertan en el texto utilizando una abreviatura para el título seguida de los números del volumen y la página de la edición de la Academia Alemana, que se encuentran en los márgenes de la mayoría de las traducciones. Las abreviaturas y ediciones utilizadas son las siguientes:

C2 = *Critique of practical reason*, editada y traducida por Mary Gregor, con una introducción de Andrews Reath (Cambridge: Cambridge University Press, 1997).

G = *Groundwork of the Metaphysics of Morals*, editada y traducida por Mary Gregor, con una introducción de Christine M. Korsgaard (Cambridge: Cambridge University Press, 1998).

R = *Religion within the Limits of Reason Alone*, editada y traducida por Theodore M. Greene y Hoyt H. Hudson, con una introducción de John R. Silber (New York: Harper Torchbooks, 1960).

⁹ Para una exposición más completa de estas ideas y de la psicología moral de Kant en general, véase la primera sección de mi “Motivation, Metaphysics, and the Value of the Self: A reply to Ginsborg, Guyer, and Schneewind,” *Ethics*, 109 (1998), pp. 49-66.

que eres autónoma, el placer no es una ley para ti: nada es una ley para ti excepto aquello que tú haces una ley para ti misma. Por lo tanto, te planteas una pregunta diferente. La propuesta es que deberías hacer el acto-A para conseguir el placentero fin-F. Dado que nada es una ley para ti excepto lo que haces una ley para ti misma, te preguntas si puedes tomar *aquello* como tu ley. Tu pregunta es si puedes querer como una ley la máxima de hacer el acto-A para producir el fin-F. Tu pregunta, en otras palabras, es si tu máxima pasa el test del imperativo categórico. El imperativo categórico es, por lo tanto, la ley de una voluntad libre.

La inclinación presenta la propuesta; la razón decide si actúa o no según ella, y la decisión toma la forma de un *acto legislativo*. Éste es claramente el modelo constitucional.

IV. Estándares para la acción

El principal punto de semejanza entre las teorías de Platón y Kant aparece, sin embargo, en su tratamiento de la acción mala. En el modelo combativo, ¿qué ocurre cuando una persona actúa mal? La respuesta debe ser que la persona es vencida por la pasión. Pero en el modelo constitucional podríamos decir igual de bien que cuando una persona actúa bien es vencida por la razón, ya que ambas fuerzas parecen estar al mismo nivel. En cambio, según el modelo constitucional, una persona actúa bien cuando actúa de acuerdo con su constitución. Si la razón vence a la pasión, ella debería actuar de acuerdo con la razón, no porque se identifique con la razón, sino porque se identifica con su constitución, y ésta dice que la razón debería gobernar.¹⁰ Entonces, ¿qué ocurre cuando una persona actúa mal? Aquí nos encontramos con lo que parece, en principio, una dificultad para el modelo constitucional: ésta resulta ser la fuente de sus ideas más profundas.

La dificultad es, por supuesto, que, según la explicación sobre Platón que acabo de dar, una *persona* injusta no puede actuar en absoluto, porque una persona injusta no está unificada por una regla constitucional. Cuando una ciudad se encuentra en estado de guerra civil, no actúa, aunque las diversas facciones dentro de ella puedan hacer diversas cosas. La analogía sugiere que cuando un alma se encuentra en estado de guerra civil, y las diversas fuerzas dentro de ella están luchando por el control,

¹⁰ Julia Annas y otros me han señalado que hay cierta tensión entre esta idea y algunos pasajes de los últimos libros de la *República* que sugieren fuertemente que la opinión de Platón es que deberíamos identificarnos con la razón — de modo más notable el pasaje de 588b-e, en el que Platón compara las tres partes del alma con una bestia de muchas cabezas (apetito), un león (espíritu), y un ser humano (razón). Estoy de acuerdo, pero creo que la tensión está dentro del texto mismo de la *República*, que es parte de una tensión general entre las concepciones del alma en los primeros libros y en los últimos.

lo que para el mundo exterior parecen ser *las acciones de la persona* son realmente sólo las manifestaciones de las fuerzas que operan *dentro* de la persona. Así, al principio parece como si *nada contara exactamente como una acción mala*.

Y hay una analogía *exacta* de esta dificultad en la ética kantiana. Pues un problema bien conocido en la *Fundamentación* es que Kant parece decir que sólo la acción autónoma, es decir, la acción regida por el imperativo categórico, es realmente una acción libre, mientras que la “acción” mala o heterónoma es un comportamiento *causado* por la obra de los deseos e inclinaciones en nosotros (G 4:453 55). Pero si esto fuera así, sería difícil ver cómo podríamos ser responsables de la acción mala o heterónoma, o por qué deberíamos considerarla como algo que nosotros *hacemos*. Parece más bien algo que ocurre en nosotros o a nosotros. Este problema surge a causa del argumento por medio del cual Kant establece la autoridad del imperativo categórico, el argumento que acabamos de ver. Pues ese argumento parece mostrar que la acción es *esencialmente* autónoma. La acción debe tener lugar bajo la idea de libertad; y una voluntad libre debe ser autónoma. Así, al principio parece como si *nada contara exactamente como una acción mala*.

Es importante observar que la *estructura* del problema en estas dos teorías es exactamente la misma. Kant identifica primero la acción con la acción autónoma, afirmando que es esencial para la acción que ésta deba ser autónoma. Luego identifica la acción autónoma con la acción regida por el imperativo categórico, la acción universalizable. De exactamente la misma manera, Platón identifica primero la acción con la acción que emerge del procedimiento constitucional, afirmando que es esencial para la acción que ésta deba emerger del procedimiento constitucional. A continuación, identifica la acción que surge del procedimiento constitucional con la acción justa. En otras palabras, cada argumento identifica primero una propiedad metafísica esencial de la acción — la autonomía en el argumento de Kant y la constitucionalidad en el de Platón — y luego, a su vez, identifica esta propiedad metafísica con una propiedad normativa — la universalizabilidad en el argumento de Kant y la justicia en el de Platón. Y así es como se justifica el requerimiento normativo.

Además, en ambos argumentos la identificación de la propiedad metafísica es un intento de aprehender un rasgo específico de la acción, algo importante que distingue una acción de un mero evento, a saber, que una acción es *atribuible* a la persona que la hace. La propiedad metafísica que Platón y Kant buscan es la que hace verdadero que la acción no sea sólo algo que sucede en o a la persona, sino que sea más bien algo que ella como persona *hace*. Es la propiedad que hace que la *persona* sea la autora de la acción. El uso explícito por parte de Platón del modelo constitucional deja claro

que trata de identificar esta propiedad. Pues ciertamente distinguimos las acciones que atribuimos a una ciudad como tal de las acciones que atribuiríamos sólo a algunos de los individuos que la componen. Y la base de esta distinción es si la acción fue el resultado de seguir procedimientos constitucionales o no. Si un espartano ataca a un ateniense, por ejemplo, no concluiremos que *Esparta* está yendo a la guerra contra Atenas a menos que el ataque haya sido hecho por un soldado que actúe bajo la dirección de los gobernantes: es decir, a menos que surja de los procedimientos constitucionales de Esparta. Por analogía, sólo atribuiremos una acción a una persona, y no a algo en ella, si ésta fue dirigida por su razón, su parte rectora. De modo similar, Kant piensa que lo que hace que una acción sea atribuible a la persona es que surja de la autonomía o el autogobierno de la persona. El ejercicio de la autonomía de la persona es lo que hace que la acción sea *suya* y, por lo tanto, lo que la hace que sea una acción.

Y así llegamos al problema. La naturaleza esencial de la acción es que tiene una cierta propiedad metafísica. Pero para tener esa propiedad metafísica debe tener una cierta propiedad normativa. Esto explica por qué la acción debe cumplir con el estándar normativo: *simplemente no es acción* si no lo hace. Pero también parece que explica aquello demasiado bien, pues parece implicar que sólo la acción buena es realmente acción, y que no queda nada para que haya acción mala.

Ahora bien, en lugar de encontrar en esto una razón para rechazar estos argumentos, creo que deberíamos verlo como nuestra principal razón para aceptarlos. Pues lo que acabamos de observar es que, según Platón y Kant, los estándares morales que aplicamos a las acciones son lo que he llamado en otro lugar “estándares internos” — estándares que una cosa debe cumplir en virtud de lo que es.¹¹ Un estándar interno es aquel que surge de la naturaleza del objeto al que se aplica, de las normas funcionales o teleológicas que lo hacen el objeto que es. Digamos que una casa, por ejemplo, es un refugio habitable. Entonces, una casa buena es una casa que tiene las características que le permiten servir como refugio habitable — las esquinas están apropiadamente selladas, el techo es impermeable y hermético, las habitaciones son lo suficientemente altas como para estar de pie dentro, y cosas por el estilo. Estos estándares internos son los que hacen que algo sea *una casa buena*.

Tenemos que distinguir aquí entre que algo sea una *casa* buena o mala y que sea una casa que resulta ser una *casa* buena o mala debido a algún estándar externo. La gran mansión que bloquea la

¹¹ Discuto la concepción de una norma interna en “The Normativity of Instrumental Reason,” en Garrett Cullity y Berys Gaut (eds.), *Ethics and Practical Reason* (Oxford: Clarendon Press, 1997), pp. 215-254. Véase especialmente pp. 249-250. Allí sostengo que el imperativo hipotético es un estándar interno para los actos de la voluntad.

vista del lago para todo el vecindario puede ser una *cosa mala* para el vecindario, pero no es por ello una *casa mala*. Por otro lado, una casa que no ofrece refugio de modo satisfactorio es una casa mala. Permíteme dar a esta clase de maldad un nombre especial. Una entidad que no cumple sus estándares internos es *defectuosa*.

La distinción entre estándares internos y externos es importante, porque los estándares internos afrontan los desafíos a su normatividad con perfecta facilidad. Supongamos que vas a construir una casa. ¿Por qué no deberías construir una casa que bloquee la vista al lago de todo el vecindario? Tal vez porque disgustará a los vecinos. Ahora bien, *ahí* hay una consideración que simplemente puedes dejar de lado, si eres lo suficientemente egoísta o rudo como para afrontar el disgusto de los vecinos. Pero como no tiene sentido preguntar por qué una casa debería servir de refugio habitable, tampoco tiene sentido preguntar por qué las esquinas deberían estar selladas y el techo debería ser impermeable y hermético. Porque si te quedas demasiado lejos del estándar interno para las casas, lo que produces simplemente no será una casa. Y esto significa que hay un sentido en el que incluso el constructor más venal y torpe debe intentar construir una casa buena, por la sencilla razón de que no hay otra forma de intentar construir una casa. Construir una casa buena y construir una casa no son actividades diferentes: porque ambas son actividades en las que debemos guiarnos por las normas funcionales o teleológicas implícitas en la idea de casa. Obviamente, no se sigue que toda casa sea una casa buena. Sin embargo, sí se deduce que la construcción de casas malas no es una actividad diferente de la construcción de casas buenas. *Es la misma actividad, hecha mal.*

Las acciones justas en Platón, las acciones universalizables en Kant, son acciones que son buenas *como* acciones, del mismo modo en que una casa que ofrece refugio con éxito es buena como casa. Y si esto es correcto, deberíamos obtener las mismas conclusiones. Si la justicia y la universalizabilidad son estándares internos, entonces no son consideraciones ajenas de cuya normatividad se pueda dudar. Un agente no puede simplemente dejar de lado la cuestión de si su acción es universalizable o justa, ya que, si se aleja demasiado de los estándares internos de las acciones, lo que produce simplemente no será una acción. En efecto, esto significa que incluso el agente más venal y torpe debe intentar realizar una acción buena, por la sencilla razón de que no hay otra forma de intentar realizar una acción. Realizar una acción buena y realizar una acción no son actividades diferentes: pues ambas son actividades en las que debemos guiarnos por las normas funcionales o teleológicas implícitas en la idea de acción. Obviamente, no se deduce que toda acción sea una acción

buena. Sin embargo, sí se deduce que realizar acciones malas no es una actividad diferente de realizar acciones buenas. *Es la misma actividad, hecha mal.*

V. La acción defectuosa

Entonces, si pudiéramos hacer que estas afirmaciones fueran plausibles, o incluso inteligibles, tendríamos aquí un resultado importante: una respuesta a la pregunta sobre por qué nuestras acciones deben cumplir estándares morales. Las acciones injustas o no-universalizables serían *defectuosas*: serían malas *como acciones*. Pero, ¿cómo pueden las acciones ser defectuosas y aún *ser* acciones? El modelo constitucional nuevamente nos proporciona los recursos para una respuesta. Pues todos sabemos que la acción de una ciudad puede ser formal o procedimentalmente constitucional y, sin embargo, no ser sustantivamente justa. De hecho, nada es más conocido: una ley debidamente legislada por el congreso e incluso avalada por la corte suprema puede ser, a pesar de todo ello, injusta. Entonces, no es que no haya territorio alguno entre una ciudad perfectamente justa y la desintegración completa de una guerra civil. Una ciudad puede estar gobernada y, sin embargo, estar gobernada por una ley errada. Y así también puede estarlo un alma. Esto, según Platón y Kant, explica cómo es posible la acción mala.

En la obra de Kant esto surge con mayor claridad en la primera parte de *La religión dentro de los límites de la mera razón*. Allí aprendemos que una persona mala no es, después de todo, aquella que es impulsada, o provocada a actuar, por deseos e inclinaciones. En cambio, una persona mala es aquella que se rige por lo que Kant llama el principio del amor a sí mismo, por un principio que subordina las consideraciones morales a aquellas que surgen de la inclinación (R 6:36). La persona que actúa según el principio del amor a sí mismo *elige* actuar como la inclinación incita (R 6:32-39). Permíteme intentar aclarar por qué Kant piensa que una acción basada en el principio del amor a sí mismo es *defectuosa*, en lugar de meramente mala en sentido externo.

Imagina a una persona a la que llamaré Harriet, que es, en el sentido *formal* que quieras, una persona autónoma. Tiene una mente humana, es autoconsciente, con una asignación normal de facultades de reflexión. No es una esclava ni una sirvienta contratada, y la situaremos — a diferencia de la original según la cual la estoy modelando — en una democracia constitucional moderna avanzada, con los plenos derechos de la ciudadanía libre y todos sus derechos humanos legalmente garantizados. En cualquier sentido formal legal y psicológico, lo que haga Harriet *depende de ella*. Sin embargo, cada

vez que tiene que tomar alguna decisión o elección importante en su vida, el modo en que Harriet lo hace es preguntando a Emma qué debería hacer, y entonces eso es lo que ella hace.¹²

Ésta es una acción autónoma y, sin embargo, es *defectuosa* como acción autónoma. Harriet es autogobernada y, sin embargo, no lo es, porque se deja gobernar por Emma. Harriet es heterónoma, no en el sentido de que sus acciones sean causadas por Emma en lugar de ser elegidas por ella misma, sino en el sentido de que se deja gobernar en sus elecciones por una ley fuera de ella misma. Incluso ayuda a mi caso que Harriet haga esto porque tiene miedo de pensar por sí misma. Pues, como he argumentado en otro lugar, así es como Kant concibe el funcionamiento del principio del amor a sí mismo.¹³ Kant no concibe que la persona que actúa a partir del amor a sí mismo reflexione activamente sobre para hacer qué tiene una razón y llegue a la conclusión de que debe hacer lo que quiere. Por el contrario, Kant lo concibe como alguien que simplemente sigue la dirección del deseo, sin reflexión suficiente. Él es heterónimo, y obtiene su ley a partir de la naturaleza, no en el sentido de que ésta cause sus acciones, sino en el sentido de que se deja gobernar por sus sugerencias — al igual que Harriet se deja gobernar por las de Emma.

La doctrina análoga en Platón es mucho más elaborada, y por ello Platón merece ser destacado. Porque lo que Kant dice aquí es incompleto y confuso. Por lo menos, parece que Kant debería haber distinguido entre un principio intemperante del amor a sí mismo — el principio de actuar según el deseo del momento — y un principio prudente del amor a sí mismo — que busca, digamos, la mayor satisfacción de los deseos a lo largo del tiempo.¹⁴ Ambos caracteres se encuentran en Platón, y además otros. En los libros VIII y IX de la *República*, Platón distingue, de hecho, cinco maneras diferentes según las cuales el alma puede ser gobernada, comparándolas con cinco tipos diferentes de constituciones posibles para una ciudad: la manera buena, que es la monarquía o la aristocracia; y cuatro malas, cada una peor que la otra: la timocracia, la oligarquía, la democracia y la peor de todas,

¹² El modelo para mi Harriet es la influenciada Harriet Smith de la novela *Emma* de Jane Austen (Oxford: Oxford University Press, 1990).

¹³ Véase mi “From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action” en Stephen Engstrom y Jennifer Whiting (eds.), *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty* (New York: Cambridge University Press, 1996), pp. 203-236, especialmente pp. 208-212.

¹⁴ Si estoy en lo cierto al decir que Kant considera que el amor a sí mismo opera irreflexivamente, esto podría parecer que favorece un principio intemperante del amor a sí mismo. Sin embargo, a veces está claro que Kant tiene en mente un principio prudente del amor a sí mismo — véase, por ejemplo, C2 5:35-36. Aunque creo que el principio intemperante de hecho se ajusta mejor a los argumentos de Kant, también creo que debería ser posible hacer que los pasajes de la segunda *Crítica* sean consistentes con la opinión según la cual quienes actúan a partir del amor a sí mismo son irreflexivos. Sólo tenemos que argumentar que hay una diferencia entre ser *reflexivo* y ser *calculador*.

la tiranía. En los tres casos intermedios, lo que hace que la constitución sea mala es que la unidad de la persona que vive bajo ella depende de circunstancias contingentes.

Lo más cercano al alma aristocrática es la persona timocrática, que, como la ciudad que le da nombre, se rige por consideraciones de honor. Tal persona ama la forma exterior, la belleza, de la bondad, casi como si fuera la bondad misma. Esta persona se equivoca y se divide contra sí misma en un cierto tipo de casos — a saber, el tipo de caso en que lo correcto es algo que parece deshonroso. Supón, por ejemplo, que la persona timocrática está luchando por el bien de la ciudad, pero se llega a un punto en el que la rendición es realmente el mejor camino. La persona timocrática puede estar tan apegada a la honorabilidad, a la belleza, al glamour si se quiere, de este tipo de acción, de luchar-por-el-bien-de-la-ciudad, que puede ser incapaz de rendirse, aunque realmente sea por el bien de la ciudad que deba hacerlo.¹⁵

A continuación viene la persona oligárquica, que parece regirse por la prudencia: es cautelosa, no-lujosa y se preocupa por el enriquecimiento a largo plazo. Al describirla, Platón emplea una distinción entre deseos necesarios, cuya satisfacción es beneficiosa o esencial para la supervivencia, y deseos innecesarios o lujosos, que son perjudiciales y no deben ser satisfechos. La persona oligárquica está atenta a los deseos necesarios y al dinero, mientras que reprime sus deseos innecesarios. Pero los reprime porque no son rentables, no porque sea malo satisfacerlos. El resultado de esta potente represión, según Sócrates, es que “alguien así no estaría completamente libre de la guerra civil interna y no sería uno, sino en cierto modo dos.” Este tipo de prudencia gobierna despóticamente sobre la parte apetitiva, como los ricos que gobiernan sobre una clase trabajadora descontenta. Si alguna fuerza exterior — quizás simplemente una tentación suficiente — refuerza y aviva sus deseos innecesarios, la persona oligárquica puede literalmente perder el control de sí misma. Si en general la persona oligárquica consigue mantenerse unida, es porque tiene el tipo de virtud imitativa de la que se burla Sócrates en el *Fedón*, la virtud de aquellos que son capaces de dominar algunos de sus placeres y miedos

¹⁵ Aunque las limitaciones de espacio no me permiten explicar la idea con suficiente detalle aquí, me atrevería a decir que el problema de la persona timocrática es que es incapaz de hacer frente a las contingencias que exigen la aplicación de lo que he llamado en otro lugar, siguiendo a John Rawls, “teoría no-ideal” [Véase mi “The Right to Lie: Kant on Dealing with Evil,” en *Creating the Kingdom of Ends* (New York: Cambridge University Press, 1996), pp. 133-158, especialmente pp. 147-154]. Es decir, actúa bien, excepto en aquellos momentos en que la verdadera bondad exige una concesión, un compromiso, una regla menos estricta, o incluso — aunque esto es raro — acciones que son formalmente incorrectas [Véase mi “Taking the Law into Our Own Hands: Kant on the Right to Revolution,” en Andrews Reath, Barbara Herman y Christine M. Korsgaard (eds.), *Reclaiming the History of Ethics: Essays for John Rawls* (Cambridge: Cambridge University Press, 1997) para una discusión de este tipo de casos].

porque a su vez son dominados por otros.¹⁶ Sócrates tiene en mente aquí argumentos tales como que debes ser moderada porque así obtendrás más placer en general. En general, Platón parece pensar que el honor y la prudencia son principios de elección lo suficientemente parecidos a la virtud verdadera como para mantener íntegra un alma a lo largo de la mayoría de los tipos de presión, aunque en la persona oligárquica las grietas sean cada vez más visibles.¹⁷

Le sigue la persona democrática, en la que no se reprimen los deseos innecesarios, y que como resultado es un intemperante. Sócrates dice que la persona democrática:

pone sus placeres en igualdad de condiciones... cediendo siempre el gobierno sobre sí mismo a cualquier deseo que se presente, como si fuera elegido por sorteo. Y cuando éste se satisface, cede el gobierno a otro, sin desdeñar ninguno, sino satisfaciendo todos por igual. (R 561b)

La democracia es un caso degenerado de autogobierno, ya que tal persona se gobierna a sí misma sólo en un sentido mínimo o formal, al igual que elegir por sorteo es diferente sólo en un sentido mínimo o formal de no elegir en absoluto. La coherencia de la vida de la persona democrática depende completamente de la coherencia accidental de sus deseos. Para ver el problema, considera un relato:

Jeremy, un estudiante universitario, se sienta en su escritorio una tarde para estudiar para un examen. Dado que se siente demasiado inquieto como para concentrarse, decide dar un paseo al aire libre. Su paseo lo lleva a una librería cercana, donde la visión de un título tentador lo atrae a mirar el libro. Sin embargo, antes de conseguirlo, se encuentra con su amigo Neil, quien le invita a reunirse con otros chicos en el bar de al lado para tomar una cerveza. Jeremy decide que puede permitirse tomar sólo una y se va con Neil al bar. No obstante, mientras espera su cerveza descubre que el ruido le da dolor de cabeza y decide volver a casa sin tomar la cerveza. Sin embargo, ahora

¹⁶ Véase Platón, *Fedón* 68d-69c.

¹⁷ Varias personas han sostenido que el problema descrito aquí no se presentaría en el caso del egoísta racional en el sentido moderno más ordinario, la persona que busca maximizar la satisfacción de sus propios intereses. De hecho, esto es lo que sugieren mis propias observaciones sobre cómo la virtud imitativa puede ayudar al oligarca a mantenerse íntegro, pues el egoísmo moderno se parece mucho a la virtud imitativa de Platón. Si es correcta, esta objeción sugeriría que te puedes constituir a ti misma por medio del principio egoísta. Una respuesta completa a esta objeción requiere un tratamiento exhaustivo de la afirmación según la cual existe un principio coherentemente formulable del egoísmo racional. Véase mi "The Myth of Egoism" de próxima aparición.

tiene demasiado dolor como para estudiar. De este modo, Jeremy no estudia para su examen, apenas sale a pasear, no compra un libro y no toma una cerveza.¹⁸

Por supuesto, la vida democrática no tiene por qué ser así; sólo es un accidente que cada uno de los impulsos de Jeremy lo lleve a una acción que socave por completo la satisfacción de la anterior. Pero ese es justamente el problema, pues también es sólo un accidente si esto *no* ocurre. La persona democrática no tiene recursos para dar forma a sus deseos para evitar esto, y por lo tanto está a merced del accidente. Al igual que Jeremy, puede ser casi completamente *incapaz de acción eficaz*.

Es a partir del caos resultante de este tipo de vida que surge el alma tiránica. Este tipo de alma nuevamente está unificada, pero no bajo el gobierno de la razón que busca el bien del conjunto. Según Platón, el alma tiránica está gobernada por algún deseo erótico (R 572d-573a), que subordina toda el alma a sus propósitos, dejando a la persona como esclava absoluta de una única obsesión dominante (R 571a-580a).¹⁹

En el argumento de Platón, como en el de Kant, la acción mala es la acción regida por un principio de elección que no es el propio de la razón: un principio de honor (timocracia), de prudencia (oligarquía), de desenfreno (democracia) o de obsesión (tiranía). Es acción, porque se elige de acuerdo con el ejercicio de un principio por el que el agente se gobierna a sí mismo y bajo cuyo gobierno está — en cierto sentido — unificado. Pero es defectuosa, porque no es el principio propio de la razón, y la unidad que produce es, al menos en los tres casos intermedios, contingente e inestable. Y Platón puede decir con Kant que la persona que se gobierna a sí misma de una de estas maneras no es, después de todo, completamente autogobernada. Porque está apoyada, por así decirlo, por el hecho de que las circunstancias que crearían una guerra civil en su alma no ocurren.

VI. La acción buena y la unidad del alma platónica

¹⁸ He tomado este ejemplo de una nota al pie de mi “The Normativity of Instrumental Reason”, Garrett Cullity y Berys Gaut (eds.), *Ethics and Practical Reason* (Oxford: Clarendon Press, 1997), pp. 215-254, p. 247, n. 64.

¹⁹ El problema con la tiranía no es el mismo que con la timocracia, la oligarquía y la democracia — no es que la unidad que produce en el alma sea contingente. Platón concibe la tiranía como una especie de locura (véase R 573c y ss.). Según como imagino al tirano, su relación con su obsesión es como la relación de un psicótico con su delirio: es capaz de y está preparado para organizar todo lo demás en torno a éste, pero a costa de una pérdida de su control sobre la realidad, sobre el mundo. Pero esto es sólo un esbozo, y un tratamiento más completo de este principio, y de la cuestión sobre por qué una persona no puede integrarse a sí misma con éxito bajo su gobierno, es necesario para la completud de la discusión de este artículo.

Ahora estamos casi listos para hablar de lo que hace que la acción sea buena. Pero primero quiero abordar una posible objeción. Recién dije que en las condiciones de la timocracia, la oligarquía y la democracia, tu unidad y, por lo tanto, tu autogobierno, están sostenidos por circunstancias externas, por la ausencia de condiciones bajo las cuales te desmoronarías. Pero podrías preguntar: ¿qué hay de malo en eso? El defecto de estos caracteres es como una grieta, un potencial de desintegración que no necesariamente se manifiesta, y mientras no lo haga, estas personas tienen procedimientos constitucionales y así pueden actuar. Así que, ¿por qué no seguir adelante y ser, digamos, oligárquica? Te mantienes íntegra la mayor parte del tiempo, podrás realizar acciones, y además ahorrarás mucho dinero.

Hay otra forma de plantear esta misma cuestión, que consiste en preguntarse si el desafío de Glaucón no es demasiado extremo. Glaucón quiere que Sócrates le diga lo que la justicia y la injusticia hacen al alma. Así que formula el siguiente desafío: toma, por un lado, a una persona que tiene un alma completamente injusta, y dale todos los beneficios externos de la justicia, es decir, todos los beneficios que provienen de que la gente crea que eres justa. Y toma, por otro lado, una persona que tiene un alma completamente justa, y dale todas las desventajas externas de la injusticia, todas las desventajas que vienen de que la gente crea que eres injusta (R 360d-361e). En particular, la persona justa que es considerada injusta será — y cito ahora — “azotada, extendida sobre un potro de tortura, encadenada, [y] cegada con fuego” (R 361e). Se supone que Sócrates tiene que demostrar que *incluso entonces* es mejor ser justo que injusto. ¿Pero no es eso pedir demasiado?

En el contexto de la discusión de la *República*, no lo es. Porque la pregunta de la *República* se plantea como una pregunta *práctica*: es la pregunta sobre si la vida justa es más digna de *elección* que la vida injusta. Y si eliges ser una persona justa, y vivir una vida justa, estás eligiendo por lo tanto hacer lo justo incluso si eso significa que serás azotada, extendida sobre el potro de tortura, encadenada y cegada con fuego. No puedes hacer un compromiso condicional con la justicia, un compromiso de ser justa a menos que la situación se torne difícil. Tu justicia se basa en la naturaleza de tus compromisos, y un compromiso como ese no *sería* un compromiso con la justicia. Por eso, a la hora de decidir si ser una persona justa, tienes que estar convencida de antemano de que valdrá la pena, aunque la situación se torne de este modo.

Supongamos —pues es suficientemente plausible— que hay una persona que vive una vida justa, es decente y honrada, siempre cumple con su parte, nunca se aprovecha de una ventaja injusta, mantiene su palabra —todo eso—, pero un día es puesta en el potro de tortura y, bajo presión de

tortura, hace algo injusto. Digamos que divulga un secreto militar o el paradero de un fugitivo injustamente perseguido. ¿Estoy diciendo que esto demuestra que nunca estuvo realmente comprometida con la justicia, porque su compromiso debe haber sido condicional? *Por supuesto que no.* Lo que el caso muestra es que el rango de cosas que las personas pueden *ser* es más amplio que el rango de cosas que pueden elegir, por así decirlo, *de antemano ser*. Esta persona se comprometió a mantener sus secretos sobre el potro de tortura, pero fracasó, eso es todo —y también de modo bastante comprensible. Pero el hecho de que puedas ser una persona justa que en estas circunstancias fracasará no muestra que puedas decidir de antemano ser una persona justa que en estas circunstancias fracasará: es decir, no muestra que puedas hacer un compromiso condicional con la justicia. Porque supongamos que te sorprendes a ti misma y sí aguantas y mantienes el secreto incluso cuando te ponen sobre el potro de tortura. ¿Fallaste entonces en *mantener* tu compromiso condicional?

Entonces, el desafío de Glaucón es uno justo. Pero Platón lo supera holgadamente. Porque no prueba meramente que la vida justa es la más digna de ser elegida. Demuestra que la vida justa es la única que puedes elegir. Permíteme que intente explicar por qué.

Considera la explicación de Platón sobre el principio de la acción justa o aristocrática. Platón dice del alma aristocrática que:

cuando él hace cualquier cosa, ya sea adquirir riqueza, cuidar su cuerpo, participar en política, o celebrar contratos privados — en todas estas cosas, cree que la acción que preserva esta armonía interna y ayuda a lograrla es justa y buena, y la llama así, y considera como sabiduría el conocimiento que supervisa tales acciones. Y cree que la acción que destruye esta armonía es injusta, y la llama así, y considera como ignorancia la creencia que la supervisa. (R 443e-444a)

En otras palabras, el principio de justicia nos dirige a realizar aquellas acciones que establecen y mantienen nuestra unidad volitiva. Ahora bien, ya hemos visto que, según Platón, la unidad volitiva es esencial si vas a actuar como persona, como una única agente unificada. Entonces, el principio propio de la razón *justamente es* el principio de actuar de una manera que te constituya en una única agente unificada. La acción deliberativa es autoconstitución.

De hecho, la acción deliberativa, por su propia naturaleza, impone un orden constitucional al alma. Cuando deliberas sobre qué hacer y luego lo haces, lo que estás haciendo es organizar tu apetito, razón y espíritu en el sistema unificado que genera una acción que se te puede atribuir como persona. La acción deliberativa junta las partes del alma en un sistema unificado. Lo que sea que estés haciendo cuando eliges una acción deliberativa, también te estás unificando en una persona. Esto significa que el principio de justicia de Platón, el principio propio de la razón, es el principio *formal* de la acción deliberativa. Es como si Glaucón preguntara: ¿qué condición puede ser ésta, que permite a la persona justa atenerse a sus principios incluso sobre el potro de tortura? Y es como si Platón respondiera: no busques ninguna condición *ulterior* que tenga eso como *efecto*. La justicia no es ninguna otra condición o una condición ulterior que nos permita mantener nuestra unidad como agentes. Es justamente esa condición misma — la condición de ser capaces de mantener nuestra unidad como agentes.

Para ver que esto es formal, considera la siguiente comparación. Se podría preguntar a Kant: ¿qué principio podría ser éste, que permite a la persona libre ser autónoma, gobernarse a sí misma? Y Kant respondería: no busques algún principio *ulterior* que tenga eso como *efecto*. El imperativo categórico no es algún otro principio o un principio ulterior que nos permita gobernarnos a nosotros mismos. Es justamente ese mismo principio, el principio de darnos leyes a nosotros mismos.

Por un lado, esta explicación del alma aristocrática nos muestra por qué las exigencias de la justicia platónica son tan altas. En ciertas ocasiones, las personas con las otras constituciones se desmoronan. Para la persona verdaderamente justa, el alma aristocrática, no hay tales ocasiones. Es completamente autogobernada, de modo que todas sus acciones, en todas las circunstancias de su vida, son real y plenamente suyas: nunca meramente las manifestaciones de fuerzas que operan en ella o sobre ella, sino siempre la expresión de su propia elección. Es completamente autogobernada: no necesariamente feliz sobre el potro de tortura — pero *ella misma* sobre el potro de tortura, ella misma incluso allí.

Y, sin embargo, el argumento de Platón muestra al mismo tiempo que esta constitución aristocrática es la única que puedes elegir. Pues no puedes, en el momento de la acción deliberativa, elegir ser algo menos que una única agente unificada. Y eso significa que no puedes propiamente elegir actuar según otro principio sino el principio de justicia. Las almas timocráticas, oligárquicas y democráticas se desintegran bajo ciertas condiciones, por lo que decidir ser una de ellas sería como hacer un compromiso condicional con tu propia unidad, con tu propia personalidad. Y eso no es posible. En efecto, considera lo que ocurre cuando las condiciones que causan la desintegración en

estas constituciones realmente ocurren. Si no te desmoronas, ¿has fallado en mantener tu compromiso, como la persona condicionalmente justa que después de todo resiste sobre el potro de tortura? Pero si te desmoronas, ¿quién es la que ha mantenido el compromiso? Si te desmoronas, no queda ninguna persona. Se puede ser una persona timocrática, oligárquica o democrática, de la misma manera que se puede ser una persona justa que fracasa sobre el potro de tortura. Pero no puedes decidir de antemano que eso es lo que vas a ser.

Por supuesto, esto no significa que todos elijan la vida justa. Lo que significa es que la elección de una vida injusta no es una actividad diferente de elegir una vida justa. Es la misma actividad — la actividad de autoconstitución — hecha mal.

VII. La acción buena y la unidad de la voluntad kantiana

Queda por mostrar que ésta es también la opinión de Kant; y para ello tenemos que revisar el argumento por medio del cual Kant establece que la acción debe estar de acuerdo con el imperativo categórico y completar el paso que le falta. Kant sostiene que, en la medida en que eres un ser racional, debes actuar bajo la idea de libertad — y esto significa que no te consideras a ti misma, o te experimentas a ti misma, como impelida a la acción, sino que decides qué hacer. Te consideras *a ti misma*, en lugar del incentivo según el cual decides actuar, como la *causa* de tu acción.²⁰ Y Kant piensa que para que esto sea así debes actuar según una ley universal. No puedes considerarte a ti misma como la *causa* de tu acción — no puedes considerar la acción como el producto de tu voluntad — a menos que quieras universalmente.

Para ver por qué, consideremos lo que ocurre si intentamos negarlo.²¹ Si nuestras razones no tuvieran que ser universales, entonces podrían ser completamente particulares — sería posible tener una razón que aplicara sólo al caso que tienes ante ti, y que no tuviera implicaciones para ningún otro caso. Querer actuar según una razón de este tipo sería lo que llamaré “querer particularista.” Si el

²⁰ Para decirlo de forma un poco más estricta, te tomas a ti misma como la causa de tus movimientos inteligibles, ya que es realmente una *acción* sólo si tú eres, o hasta el punto en que eres, la causa. Creo que hay preguntas filosóficas importantes, aún por resolver, sobre cómo debería formularse exactamente este punto, pero por ahora dejo en el texto la formulación más familiar. Estoy en deuda con Sophia Reibetanz y Tamar Schapiro por las discusiones sobre estos puntos.

²¹ El argumento que sigue apareció por primera vez en la sección 1 de mi “Reply” en *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), pp. 225-233. Espero que la presente versión sea más clara.

querer particularista es imposible, entonces se sigue que el querer debe ser universal — es decir, una máxima, para ser querida del todo, debe ser querida como una ley universal.

Ahora bien, hay que tener en cuenta dos cosas aquí. En primer lugar, la cuestión no es si podemos querer una nueva máxima para cada nueva ocasión. Perfectamente podemos hacer eso, ya que cada ocasión puede tener diferencias relevantes respecto a la última que encontramos. Cualquier diferencia en la situación que sea realmente relevante para la decisión pertenece propiamente a nuestra máxima, y esto significa que nuestra máxima puede ser bastante específica para la situación en cuestión. Este argumento no pretende mostrar que las razones son generales. Pretende mostrar que las razones son universales, y la universalidad es bastante compatible con — de hecho, requiere — un alto grado de especificidad.

El segundo punto es que será suficiente para el argumento si el principio que se quiere se quiere, como lo llamaré, como provisionalmente universal. Para explicar lo que quiero decir con esto utilizaré un par de contrastes. Hay tres formas diferentes en las que podemos considerar que nuestros principios abarcan una variedad de casos, y es importante mantenerlas diferenciados. Tratamos un principio como *general* cuando pensamos que aplica a un amplio rango de casos similares. Tratamos un principio como universal, o, como diré a veces, *absolutamente universal*, cuando pensamos que aplica absolutamente a todos los casos de un determinado tipo, pero todos los casos deben ser exactamente de ese tipo. Tratamos un principio como *provisionalmente universal* cuando pensamos que aplica a todos los casos de un cierto tipo, a menos que haya alguna buena razón para no hacerlo. La diferencia entre considerar un principio como universal y considerarlo como provisionalmente universal es marginal. Tratar un principio como provisionalmente universal equivale a hacer un reconocimiento mental según el cual es posible que no hayas pensado en todo lo necesario para que el principio sea universal y, por lo tanto, que no lo hayas especificado completamente. Tratar los principios como generales y tratarlos como provisionalmente universales son superficialmente similares, porque en ambos casos admitimos que puede haber excepciones. Pero en realidad son profunda y esencialmente diferentes, y esto se manifiesta en lo que ocurre cuando nos encontramos con excepciones. Si pensamos que un principio es meramente general y nos encontramos con una excepción, no pasa nada. El principio era sólo general, y esperábamos que hubiera algunas excepciones. Pero si un principio era provisionalmente universal, y nos encontramos con un caso excepcional, entonces debemos volver atrás y revisarlo, acercándolo un poco más a la universalidad absoluta a la que la universalidad provisional esencialmente aspira.

Los principios causales aproximados con los que operamos en la vida cotidiana (no estoy hablando ahora de física cuántica) son provisionalmente universales, y lo señalamos a veces utilizando la frase “todo lo demás igual.” El principio de que al encender un fósforo se produce una llama se cumple todo lo demás igual, donde las cosas que tienen que ser iguales son que no haya una ráfaga de viento o un chorro de agua o una rareza en la composición química de la atmósfera que interfiera con la conexión habitual. Hay condiciones contextuales para la operación de estas leyes, y sin enumerarlas y posiblemente sin conocerlas todas, mencionamos que deben cumplirse cuando decimos “todo lo demás igual.” Aunque ciertamente hay excepciones, una ley natural no es meramente general, pues siempre que ocurre una excepción, buscamos una explicación. Algo debe haber hecho que este caso sea diferente: una de sus condiciones contextuales no se cumplió.

Para ver cómo funciona esto en el caso práctico, consideremos una aporía estándar del criterio de universalizabilidad de Kant. Puede parecer que querer ser médico es una razón adecuada para convertirse en médico, ya que no hay nada incorrecto en ser médico — de hecho, es bastante admirable — y si te preguntas si podría ser una ley que todo el que quiera ser médico se convierta en uno, parece, superficialmente, bien. Pero entonces llega el objetor y dice, pero mira, supón que realmente *todos* quisieran ser médicos y nadie quisiera ser otra cosa. Todo el sistema económico colapsaría, y entonces no podrías ser médico, ¡así que tu máxima se habría contradicho a sí misma! Entonces, ¿demuestra esto que es incorrecto ser médico simplemente porque lo quieres?

Lo que demuestra es que el mero deseo de ejercer una determinada profesión es sólo una razón provisionalmente universal para hacerlo. Hay una condición contextual para que sea correcto ser médico porque lo quieres, a saber, que la sociedad tenga alguna necesidad de que la gente ejerza esta profesión. En efecto, el caso sí muestra que es incorrecto ser médico simplemente porque lo quieres — la máxima necesita revisión, ya que no es absolutamente universal a menos que mencione como parte de tu razón para convertirte en médico que hay una necesidad social. Alguien que decide convertirse en médico bajo la plena luz de la reflexión también lo tiene en cuenta.

Ese caso es fácil, pero no hay ninguna razón general para suponer que podamos pensar en todo de antemano. Cuando adoptamos una máxima como ley universal, sabemos que podría haber casos, casos en los que no hemos pensado, que nos mostrarían que, después de todo, no es universal. Podemos admitir excepciones en ese sentido. Pero mientras se mantenga el compromiso de revisión ante excepciones, la máxima no es meramente general. Es provisionalmente universal.

Entonces, la voluntad particularista no es una cuestión de querer una nueva máxima para cada ocasión, ni una cuestión de querer una máxima que podrías tener que cambiar en otra ocasión. Ambas posibilidades son compatibles con considerar las razones como universales. En cambio, la voluntad particularista sería una cuestión de querer una máxima exactamente para esta ocasión sin considerar que tenga otras implicancias de cualquier tipo para cualquier otra ocasión. Quieres una máxima pensando que puedes usarla sólo esta vez y luego, por así decirlo, desecharla; ni siquiera necesitas una razón para cambiar de opinión.

Ahora voy a argumentar que ese tipo de querer es imposible. El primer paso es éste: Concebirte a ti misma como la causa de tus acciones es identificarte con el principio de elección según el cual actúas. Una voluntad racional es una causalidad autoconsciente, y una causalidad autoconsciente es consciente de sí misma como causa. Ser consciente de ti misma como causa es identificarte con algo en el escenario que da lugar a la acción, y esto debe ser el principio de elección. Por ejemplo, supón que experimentas un conflicto de deseos: tienes el deseo de hacer tanto *A* como *B*, y son incompatibles. Tienes algún principio que favorece a *A* por sobre *B*, así que ejecutas este principio, y eliges hacer *A*. En este tipo de caso, no te consideras una mera espectadora pasiva de la batalla entre *A* y *B*. Consideras la elección como tuya, como el producto de tu propia actividad, porque consideras el principio de elección como expresivo, o representativo, de ti misma. Debes hacerlo así, pues la única alternativa a la identificación con el principio de elección es considerar el principio de elección como una tercera cosa en ti, otra fuerza a la par con los incentivos para hacer *A* y para hacer *B*, que resulta que ha puesto su peso en favor de *A*, en una batalla en la que serías, después de todo, una mera espectadora pasiva. Pero entonces no eres la causa de la acción. La agencia autoconsciente o racional, entonces, requiere la identificación con el principio de elección según el cual actúas.

El segundo paso es ver que el querer particularista hace imposible que te distingas a ti misma, tu principio de elección, de los diversos incentivos según los cuales actúas. Según Kant, siempre tienes que actuar según un incentivo u otro, porque toda acción, incluso la acción por deber, implica una decisión sobre una propuesta: algo te tiene que sugerir la acción. Y para querer de forma particularista, tienes que identificarte en cada caso completamente con el incentivo de tu acción. Ese incentivo sería, por el momento, tu ley, la ley que definió tu agencia o tu voluntad.

Es importante ver que, si tuvieras una voluntad particularista, no te identificarías con el incentivo como representativo de ninguna clase de tipo, ya que si lo tomaras como representante de un tipo lo estarías tomando como universal. Por ejemplo, no podrías decir que decidiste actuar según

la inclinación del momento, *porque estabas inclinada de ese modo*. Alguien que toma como máxima “haré las cosas a las que me inclino, sean las que sean” ha adoptado un principio universal, no uno particular: tiene el principio de tratar como razones sus inclinaciones *como tales*. Una voluntad verdaderamente particularista debe comprender el incentivo en toda su particularidad: éste, de ninguna manera que sea más descriptible, es la ley de tal voluntad. Por lo tanto, quien entra en un querer particularista ni siquiera tiene un alma democrática. Sólo existe la tiranía del momento: la completa dominación del agente por algo que está dentro de él.

El querer particularista erradica la distinción entre persona e incentivos según los cuales actúa. Pero entonces no queda nada que sea la *persona*, el agente, que sea su voluntad como distinta del juego de incentivos en su interior. No es una persona, sino una serie, un mero conglomerado, de impulsos no-relacionados. No hay diferencia entre alguien que tiene una voluntad particularista y alguien que no tiene voluntad en absoluto. La voluntad particularista carece de un sujeto, de una persona que sea la causa de estas acciones. Por lo tanto, el querer particularista no es querer en absoluto.

Si una voluntad particularista es imposible, entonces, cuando quieres una máxima, tienes que tomarla como universal. Si no lo haces, no estás operando como una causa autoconsciente, y entonces no estás queriendo. Para poner el punto en términos kantianos familiares, sólo podemos añadir el “yo quiero” a nuestras elecciones si queremos nuestras máximas como leyes universales.²² El imperativo categórico es un estándar interno para las acciones, porque la conformidad a éste es constitutiva de un ejercicio de la voluntad, de una acción de una persona como opuesta a una acción de algo dentro de ella.

Y este argumento también muestra que la concepción de Kant es la misma que la de Platón. Pues si el querer particularista es lo que nos hace colapsar, el querer universal es lo que nos mantiene unidos. Para Kant, como para Platón, la acción deliberativa, por su propia naturaleza, impone unidad sobre la voluntad. Sólo cuando te preguntas si tu máxima puede ser una ley universal, ejerces la causalidad autoconsciente, la autonomía, que genera una acción que se te puede atribuir como una persona completa. Entonces, al margen de cualquier otra cosa que estés haciendo cuando eliges una acción deliberativa, también te estás unificando a ti misma en una persona. Para Kant, como para Platón, la acción es autoconstitución.

²² Debo esta formulación de mi punto a Govert den Hartogh.

VIII. Conclusión

Concluiré revisando el curso de la discusión y diciendo lo que considero que ha establecido. Comencé criticando el modelo combativo por fallar en identificar la persona que es autora de sus acciones. Espero que a estas alturas sea claro por qué falla de tal manera. El modelo combativo no es un retrato del alma humana. Es un retrato del alma humana en ruinas, desgarrada por la guerra civil y, por lo tanto, incapaz de actuar. Según el modelo constitucional, una acción es tuya cuando se elige de acuerdo con tu constitución. Tu constitución es lo que te da el tipo de unidad volitiva que necesitas para ser la autora de tus acciones. Y la persona que actúa de acuerdo con la mejor constitución, la constitución más unificada, es quien más verdaderamente es la autora de sus acciones. Para Kant, como para Platón, la integridad es la esencia metafísica de la moralidad.

La discusión de este artículo no nos lleva, por sí misma, a la necesidad de actuar moralmente. El objetivo de la discusión ha sido establecer que el principio platónico de justicia y el imperativo categórico de Kant son estándares formales de la acción deliberativa. Tanto Kant como Platón creían que un cierto contenido, el contenido de la moralidad ordinaria, podía derivarse a partir de estos principios formales. La convicción de Platón aparece en uno de los momentos más tristemente célebres de la *República*, cuando Sócrates propone a Glaucón que pueden eliminar cualquier duda que puedan tener sobre su definición de justicia “apelando a casos comunes” (442d-e). Consecuentemente, le pregunta a Glaucón si la persona justa, como la ha descrito Sócrates, malversaría depósitos, asaltaría templos, robaría, traicionaría a sus amigos o a su ciudad, violaría sus juramentos o sus otros acuerdos, cometería adulterio, sería irrespetuoso con sus padres o ignoraría a los dioses, a todo lo cual Glaucón responde, con una complacencia alarmante para quien lee, que no, que la persona justa que Sócrates ha descrito no haría esta clase de cosas. No se nos dice exactamente por qué está tan seguro. Kant, por supuesto, sí intenta mostrarnos cómo se puede derivar contenido a partir de su principio formal y, en ese sentido, su versión del argumento es superior a la de Platón, aunque el éxito de sus esfuerzos es objeto de un viejo y famoso debate. Creo que la postura de Kant se puede defender, pero no he tratado de hacerlo aquí.²³ Tanto los argumentos de Platón como los de Kant se mueven (1) a partir de la propiedad metafísica de la acción que la hace poseer un autor y, por lo tanto, la hace acción — la

²³ En la Lección 3 de *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996) doy un argumento que pretende moverse desde la versión formal del imperativo categórico hacia las exigencias morales a través de la Fórmula de la Humanidad de Kant. Véanse especialmente las secciones 3.3.7-3.4.10.

autonomía en el caso de Kant, la constitucionalidad en el de Platón — hacia (2) una exigencia normativa formal que las acciones deben cumplir para tener esa propiedad — la universalidad en el caso de Kant, la justicia en el de Platón — y luego a través de (3) la derivación de contenido a partir de la exigencia formal para llegar a las exigencias morales ordinarias. Los dos primeros pasos son los que aquí han sido mi objeto.

Al menos en el sentido formal, entonces, la justicia platónica y el imperativo categórico de Kant son estándares internos para las acciones, porque es sólo en la medida en que tus acciones surgen de toda tu persona, en lugar de algo en tí, que pueden ser acciones. No se sigue exactamente que debamos elegir las acciones de modo justo y de acuerdo con el imperativo categórico, pues en cierto sentido no podemos elegir de ninguna otra manera. Elegir acciones malas no es una actividad distinta de elegir acciones buenas. Es la misma actividad — la actividad de autoconstitución — hecha mal.

Department of Philosophy

Harvard University

208 Emerson Hall

Cambridge, MA 02138

USA