

## ¿De qué hablamos cuando hablamos de justificación del castigo? Una propuesta dworkiniana

### What are we talking about when we talk about justification of punishment? A dworkinian proposal

*Tomás Fernández Fiks*<sup>\*</sup>

#### Resumen

Este artículo sostiene que gran parte de la ambigüedad persistente en torno a la discusión sobre la justificación del castigo se debe a una confusión inadvertida entre preguntas de naturaleza descriptiva, normativa y conceptual dentro de las denominadas teorías de la pena. Tal confusión, se sugiere, produce una perniciosa superposición de niveles de análisis, como se puede constatar, por ejemplo, en la teoría agnóstica defendida por Zaffaroni, Alagia y Slokar en su Manual de Derecho Penal. A partir de este diagnóstico, el trabajo propone una alternativa metodológica inspirada en la teoría de la interpretación de Ronald Dworkin. Se argumenta que el filósofo del derecho penal debería aproximarse a la práctica del castigo de un modo análogo al que el juez dworkiniano se aproxima al derecho. Este enfoque interpretativo, se postula, ofrece un marco atractivo para abordar algunas de las tensiones recurrentes en la literatura sobre la teoría de la pena.

**Palabras clave:** justificación del castigo – teorías de la pena – interpretación constructiva – derecho como integridad

#### Abstract

This article argues that much of the persistent ambiguity surrounding the discussion about the justification of punishment results from an unnoticed conflation of descriptive, normative, and conceptual questions within the so-called theories of

---

\* Abogado (UNMdP), LL.M. (Columbia Law School), candidato a Doctor (UBA). Quisiera agradecer los comentarios de Tomás Marino y de los miembros del grupo de investigación “Explicación y comprensión” de la Universidad Nacional de Mar del Plata a una versión anterior de este trabajo, como también las observaciones de un réferi anónimo. Asimismo, agradezco las intervenciones de los participantes de la conferencia de ICON-S realizada en dicha universidad durante los días 21 y 22 de agosto de 2025, donde expuse algunas ideas elaboradas en este artículo.

punishment. Such conflation, it is suggested, produces a confusion of analytical levels, as illustrated by the position advanced by Zaffaroni, Alagia, and Slokar in their Manual de Derecho Penal. Building on this diagnosis, the article proposes a methodological alternative inspired by Ronald Dworkin's theory of interpretation. It contends that the philosopher of criminal law should approach the practice of punishment in a manner analogous to the way Dworkin's judge approaches the law. This interpretive approach offers a coherent framework for addressing some of the recurrent tensions identified in the penal theory literature.

**Key words:** justification of punishment – punishment theory – constructive interpretation – law as integrity

## I. Introducción

La cuestión de la justificación del castigo constituye uno de los grandes temas de la filosofía moral, política y jurídica.<sup>1</sup> La ubicuidad del tema en estas disciplinas no debería causar sorpresa si se repara en que existen pocas prácticas tan universales, pero a la vez tan fácilmente cuestionables desde un punto de vista moral, como el castigo penal. En efecto, se trata esta de una práctica que, en un sentido evidente, necesita de una justificación.<sup>2</sup>

La razón no es misteriosa: normalmente, no estamos habilitados a tratarnos unos a otros de la manera en que lo hacemos cuando castigamos a alguien.<sup>3</sup> Por tanto, es de suponer que, si en determinadas ocasiones ese trato se vuelve aceptable, tiene que haber buenas razones para que así sea. Además, si es el Estado quien habrá de encargarse de imponer el castigo, es necesario explicar por qué deberían asignarse recursos públicos para esa tarea y por qué, en su caso, resultaría tolerable que la fuerza pública recaiga sobre un individuo, con la asimetría de poder que ello conlleva.

Además de su persistencia, una nota característica del debate acerca de la justificación del castigo es su relativa inmutabilidad. En efecto, si uno examina cualquier texto sobre la materia, se encontrará con que la discusión suele ser presentada siguiendo una estructura común, que no ha ido variando demasiado con el paso del tiempo.<sup>4</sup> En esta presentación estándar se contraponen distintas posiciones que ofrecen respuestas divergentes a la pregunta que indaga si, acaso, resulta posible justificar el castigo.

A grandes rasgos, podría decirse que de un lado, se encuentran quienes pretenden fundamentar la permisibilidad del castigo invocando razones relacionadas con el merecimiento negativo del sujeto que es castigado (los retribucionistas). Del otro, se hallan quienes apelan a los efectos beneficiosos que el castigo puede llegar a producir, ya

<sup>1</sup> Sancinetti destaca que ya en Platón puede encontrarse un diálogo entre Sócrates y Protágoras en el que se “leen pensamientos que más tarde serían motor de la discusión de siglos sobre la *esencia y fines* de la pena, una discusión que perdura hasta hoy.” Véase Marcelo A. Sancinetti., *Hacia un derecho penal más racional: contribuciones sobre castigo estatal e imputación*, 1ra ed. (Ciudad Autónoma de Buenos Aires: Editores del Sur, 2022), p. 87.

<sup>2</sup> Mitchell Berman, “Punishment and Justification”, *Ethics* 118, Nro. 2 (enero 2008), pp. 258-290.

<sup>3</sup> Piénsese, para ilustrar este punto, en lo que se le hace a una persona cuando se le impone la pena paradigmática en las sociedades contemporáneas: la pena de prisión.

<sup>4</sup> En un reciente artículo, Ferrante captura elocuentemente este rasgo con una metáfora musical, al equiparar el valor filosófico de la teorización contemporánea sobre la justificación del castigo al valor musical de una sesión de jazz, en que los músicos interpretan y reversionan estándares cuya melodía es conocida por todos. Véase Marcelo Ferrante, “¿Algo más sobre la justificación del castigo?”, *Lecciones y Ensayos*, Nro. 109 (2022), p. 18.

sea en el mismo sujeto que es castigado o en la sociedad en general (los consecuencialistas). Un tercer grupo pretende combinar los aspectos atractivos de las dos posiciones recién mencionadas (los partidarios de las teorías mixtas). Por último, también están aquellos que creen que la práctica del castigo no puede justificarse en absoluto, por lo que proponen eliminarla y reemplazarla por otra práctica distinta. Dentro de este último grupo se inscriben varias corrientes, pudiendo mencionarse, sin pretensión de exhaustividad, a los escépticos respecto del libre albedrío, algunos promotores de la justicia restaurativa y los abolicionistas.

Sin embargo, y a pesar de esta aparente uniformidad en el tratamiento de la cuestión, no siempre resulta claro qué es exactamente lo que está bajo discusión. Ello se debe a que, bajo el rótulo general del problema de la justificación del castigo, suelen agruparse diversas empresas teóricas que no siempre se distinguen con claridad. Según argumentaré, la falta de delimitación entre distintas preguntas que, —aunque no del todo desconectadas— son conceptualmente independientes, tiende a oscurecer el debate. Partiendo de este diagnóstico, uno de los objetivos del presente trabajo es distinguir esas preguntas para poder pasar en limpio de qué hablamos cuando hablamos de la justificación del castigo.

Creo que esta tarea es valiosa en sí misma, en la medida que puede aportar un poco de orden a una discusión que, en ocasiones, parece no avanzar porque sus participantes están discutiendo cosas distintas. El objeto de este trabajo no se agota, sin embargo, en esa labor taxonómica o conceptual. Por el contrario, una vez identificados algunos de los distintos caminos que suelen hallarse en la literatura a la hora de abordar el problema del castigo, propondré una ruta alternativa que considero prometedora y que, hasta ahora, no ha suscitado demasiada atención. Esta propuesta se basa en la obra de Ronald Dworkin.

Mi interés no radica, por tanto, en defender una tesis sustantiva en el marco de la discusión que es objeto de análisis, sino en realizar una especie de “reconstrucción racional” de esa discusión.<sup>5</sup> Podría entonces decirse que el presente artículo tiene una orientación más bien metateorética o metodológica.

<sup>5</sup> Tomo la expresión de Alchourrón y Bulygin, quienes entienden que “[l]a reconstrucción racional de un concepto es el método por medio del cual un concepto inexacto y vago —que puede pertenecer al lenguaje ordinario o a una etapa preliminar en el desarrollo de un lenguaje científico— es transformado en un concepto exacto o, por lo menos, más exacto que el primitivo.”, con la salvedad de que lo que aquí se examina es una discusión teórica y no un concepto. Véase Carlos E. Alchourrón y Eugenio Bulygin, *Sistemas normativos*, 2da ed. (Buenos Aires: Astrea, 2012), p. 11.

A continuación, brindo un panorama general de la estructura del trabajo. Comenzaré por establecer una primera distinción referida a la naturaleza misma del problema bajo análisis. En concreto, sostendré que quienes teorizan sobre la justificación del castigo discrepan acerca de si se trata, esencialmente, de un problema de filosofía moral o política. Quienes se inclinan por la primera alternativa suelen, asimismo, enfocarse en el castigo como un fenómeno amplio (que incluye, pero también excede, al castigo propiamente estatal), mientras que los que optan por la segunda suelen centrarse en la práctica específica del castigo penal.

En la sección siguiente, distinguiré la pregunta sobre el fin de la pena de aquella sobre su justificación moral. Mientras que la primera suele formularse con el enunciado “¿cuál es la razón por la que castigamos?”, la segunda se suele plantear en los siguientes términos: “¿por qué es moralmente admisible, si es que lo es, que se castigue a alguien?”. En el apartado que sigue, señalaré que estas dos preguntas deben diferenciarse, a su vez, de aquella de naturaleza conceptual que indaga acerca de qué tipo de fenómeno es la pena.

Finalmente, delinearé un posible abordaje al problema de la justificación del castigo tomando como base la teoría interpretativa elaborada por Ronald Dworkin. Según argumentaré, esta propuesta tiene potencial para lograr armonizar algunas de las tensiones identificadas en los puntos anteriores.

## II. El problema de la justificación del castigo: entre la filosofía moral y la filosofía política

¿Es la justificación del castigo un problema de filosofía moral o política? La respuesta que se brinde a esta pregunta no es trivial, en la medida que el arsenal teórico —las principales ideas y argumentos, conceptos básicos, recursos retóricos, etc.— que posee cada una de esas ramas de la filosofía es distinto al que posee la otra. Quienes hacen hincapié en el aspecto moral ven en la justificación del castigo un campo de batalla entre la ética deontológica y la ética utilitarista.<sup>6</sup> Los autores que privilegian el aspecto político, en cambio, encuadran su teoría acerca de la justificación del castigo dentro de la discusión más amplia existente entre las distintas teorías políticas.<sup>7</sup> En el primer caso, la discusión

<sup>6</sup> Guyora Binder, “Punishment Theory: Political or Moral?”, *Buffalo Criminal Law Review* 321 (2022), p. 1.

<sup>7</sup> Un ejemplo de esta corriente lo proveen Pettit y Braithwaite, quienes formulan una “teoría republicana” del castigo. Ver John Braithwaite y Philip Pettit, *No sólo su merecido. Por una justicia penal que vaya más allá del castigo* (Buenos Aires: Siglo Veintiuno Editores, 2015).

sobre el castigo resulta una derivación de aquella más general acerca de cómo deberíamos tratarnos unos a otros. En el segundo, la pregunta subyacente se centra en la relación existente entre el Estado y los ciudadanos, y en los deberes que se derivan del concepto de ciudadanía.

Ahora bien, es indudable que el castigo supone un problema de índole moral, en la medida que castigar a alguien implica ocasionarle un sufrimiento o dolor de manera intencional, y esto atenta contra normas morales que solemos aceptar sin demasiados reparos.<sup>8</sup> Algunos autores privilegian este problema por sobre otros al momento de analizar la posibilidad de justificar el castigo, decisión que tiñe todo su abordaje de una tonalidad eminentemente moral. Por ejemplo, Mitchell Berman sostiene que, tradicionalmente, el retribucionismo ha sido ofrecido como una respuesta hecha a la medida (*tailored*) de una objeción específica, a saber, aquella que cuestiona la permisibilidad moral del castigo con el argumento de que su imposición implica hacer sufrir intencionalmente a quien es castigado.<sup>9</sup> Esta es una objeción estrictamente moral, y por ende también lo es la respuesta fundada en el merecimiento que frente a ella se ensaya desde una perspectiva retribucionista.

Los autores que adoptan este enfoque no desconocen que la práctica del castigo penal presenta una variedad de problemas, incluyendo algunos de índole política. No obstante, deciden enfocarse, por razones metodológicas, en aquel que consideran central: a saber, el problema moral que se suscita cuando intentamos justificar por qué resultaría admisible que se haga sufrir intencionalmente a otra persona. Leo Zaibert expresa este punto de vista con claridad cuando dice:

deseo prestar mayor atención a la manera en que la tensión entre el retribucionismo y el consecuencialismo se manifiesta en el plano de “personas castigando a otras personas”, porque esto puede arrojar luz sobre el intrincado problema de la justificación del castigo antes de que nos enredemos en las formas en que las consideraciones políticas entran a jugar en la discusión una vez que consideramos la pregunta acerca de “ómo y cuándo debería el Estado castigar.”<sup>10</sup>

Ahora bien, como se infiere del pasaje recién citado, la objeción que denuncia la irrogación de sufrimiento que caracteriza al castigo resulta pertinente no solo contra el castigo penal, sino contra cualquier tipo de castigo. Pues todo castigo conlleva, por

<sup>8</sup> Berman, “Punishment and Justification”, p. 266.

<sup>9</sup> Mitchell Berman, “Two Kinds of Retributivism”, en R. A. Duff y Stuart Green, eds., *Philosophical Foundations of Criminal Law* (Oxford: Oxford University Press, 2011), p. 444.

<sup>10</sup> Leo Zaibert, *Punishment and Retribution* (New York: Ashgate, 2006), p. 21. La traducción del original en inglés al castellano me pertenece, como todas las siguientes salvo indicación en contrario.

definición, la intención de causarle sufrimiento a quien lo recibe.<sup>11</sup> Por este motivo, quienes dedican sus esfuerzos a intentar responder esta objeción suelen enfocarse en el castigo como un fenómeno más amplio o, incluso centrándose en el castigo penal, conciben a este último como una especie, o un caso paradigmático, del género castigo.<sup>12</sup>

Por otra parte, vinculados a la justificación del castigo se encuentran una serie de problemas que han sido extensamente desarrollados en la filosofía moral: particularmente, aquellos referidos a las condiciones de la responsabilidad moral y el reproche. En estos desarrollos se examinan cuestiones como, por ejemplo, si los seres humanos son moralmente responsables en el sentido básico que sería requerido para ser considerados blancos adecuados de juicios de merecimiento.<sup>13</sup> Ahora bien, de que los humanos sean moralmente responsables en este sentido depende la viabilidad de una de las tesis centrales en la discusión sobre la justificación del castigo, como es el retribucionismo.<sup>14</sup> Por lo tanto, la relevancia de la filosofía moral para dicha discusión resulta insoslayable.

Desde la postura opuesta, una corriente de autores hace hincapié en el carácter institucional del castigo penal. Este rasgo, argumentan, es lo suficientemente distintivo como para requerir un tratamiento diferente al que resultaría pertinente para analizar la permisibilidad del castigo como fenómeno general. La idea central es que, al versar la discusión sobre una práctica institucional, es necesario recurrir a la filosofía política. Guyora Binder es contundente al expresar esta visión:

[e]l castigo penal no es un comportamiento, sino una institución. Es parte de un sistema que involucra normas de conducta, un proceso autoritativo para generar esas normas, un proceso autoritativo para tomar decisiones que imponen sanciones, y alguna medida de poder práctico sobre personas o recursos. Castigar a alguien no es simplemente causarle un daño; ni siquiera es causarle un daño por algo que ha hecho. Es una forma de reivindicar un cierto tipo de autoridad institucional, incluso cuando la institución es solo la familia. Castigar a alguien implica afirmar un derecho y aceptar la obligación de castigar a cualquiera que se encuentre en las mismas circunstancias (...). El castigo nunca es el acto aislado de un individuo: castigar implica actuar

<sup>11</sup> Si bien la definición de castigo es un tema controvertido, prácticamente todas las definiciones que pueden encontrarse en la literatura especializada mencionan la causación intencional de sufrimiento como un rasgo esencial o, en otros términos, una condición necesaria del concepto de castigo.

<sup>12</sup> H. L. A. Hart, *Punishment and Responsibility. Essays in the Philosophy of Law*, 2da ed. (Oxford: Oxford University Press, 2008), p. 5.

<sup>13</sup> Derk Pereboom, “Free Will Skepticism and Criminal Punishment”, en Thomas A. Nadelhoffer, ed., *The Future of Punishment* (Oxford: Oxford University Press, 2013), p. 61.

<sup>14</sup> En línea con esto, autores como Gregg Caruso fundamentan su oposición al retribucionismo a partir de la defensa de una posición escéptica respecto del libre albedrío. Ver Gregg Caruso, *Rejecting Retributivism* (Cambridge: Cambridge University Press, 2021).

como un oficial o agente que participa en un sistema aplicando una norma promulgada con autoridad.<sup>15</sup>

Con todo, resulta claro que un fenómeno complejo como el castigo penal presenta una diversidad de problemas,<sup>16</sup> algunos de los cuales requieren auxilio de la filosofía moral y otros de la filosofía política. La decisión metodológica de privilegiar un problema en particular por sobre otro responde a la necesidad teórica de acotar el objeto de estudio hasta llegar a un punto manejable, pero no implica negar que los problemas remanentes existan, ni que merezcan ser analizados. El abordaje que pone el énfasis en el carácter institucional del castigo penal y examina el problema político inherente a este, y aquel que se centra en el problema moral que supone el castigo en general no son, por ende, excluyentes, sino complementarios. Y si bien nada obsta que se examine cada uno de esos problemas de forma individual, tampoco existe un impedimento para brindarles un tratamiento conjunto. Es más: en ocasiones, es posible advertir que ambas cuestiones son tenidas en cuenta a la hora de identificar lo que, de manera global, es caracterizado como el problema del castigo. Por ejemplo, David Boonin, en su libro que lleva justamente ese título, comienza señalando:

¿Cómo puede ser que el hecho de que una persona infrinja una norma justa y razonable vuelva moralmente permisible que el Estado la trate de formas que, de lo contrario, serían impermisibles?<sup>17</sup>

Nótese que, con esta manera de plantear el problema del castigo, se hace alusión al problema moral que implica hacer sufrir intencionalmente a una persona y al problema político que conlleva que sea el Estado quien trata de esa forma a un individuo por haber infringido una ley.

A modo de conclusión de esta sección, podría decirse que la discusión acerca de la justificación del castigo penal contiene problemas vinculados con la filosofía moral y otros vinculados con la filosofía política.<sup>18</sup> Los filósofos discrepan a la hora de privilegiar una u otra serie de problemas o darles un tratamiento global. En cualquier caso, es aconsejable tener en cuenta la pluralidad de problemas y enfoques existentes para poder contextualizar los aportes de cada autor. De esta manera se hace posible, por ejemplo,

<sup>15</sup> Binder, “Punishment Theory: Political or Moral?”, p. 321.

<sup>16</sup> Rabossi, *La justificación moral del castigo*, p. 18.

<sup>17</sup> David Boonin, *The Problem of Punishment* (Cambridge: Cambridge University Press, 2008), p. 1.

<sup>18</sup> En este sentido, Fletcher sostiene que “el Derecho penal es una especie de filosofía política y moral”. George Fletcher, *Rethinking Criminal Law* (Oxford: Oxford University Press, 2000), p. xix.

advertir que la defensa del retribucionismo axiológico efectuada por Zaibert<sup>19</sup> resulta perfectamente consistente con una teoría de corte contractualista como son las de Morris<sup>20</sup> y Pawlik.<sup>21</sup> Enfrentar estas teorías, o agruparlas a todas bajo el mismo rótulo, sería un error, dado que se enfocan en aspectos distintos.

### III. El fin de la pena y su justificación moral

Es común en la literatura —sobre todo, en la filosofía continental del derecho penal— que la pregunta acerca de la justificación del castigo se equipare con aquella sobre el fin de la pena. De acuerdo con esta manera de presentar las cosas, la cuestión relevante, al momento de analizar la permisibilidad del castigo, pasa por establecer cuál es la finalidad que persigue esa práctica. Las distintas teorías de la pena resultarían así teorías legitimantes<sup>22</sup> cuya diferencia radicaría en cuáles son los fines que atribuyen al castigo. En consonancia con esto, la pregunta acerca de los fines de la pena suele plantearse de la siguiente forma: ¿por qué castigamos a quienes cometen delitos? O bien, ¿cuál es la finalidad del castigo?

Bajo estas formulaciones, la pregunta en cuestión parecería apuntar a una cuestión susceptible de verificación empírica. Contestarla requeriría, por tanto, dilucidar cuáles son los motivos por los que de hecho castigamos. Se trataría, así, de una indagación

<sup>19</sup> Este autor concibe al retribucionismo como la tesis que sostiene que el sufrimiento merecido es intrínsecamente valioso. Su defensa del retribucionismo, por tanto, trasciende el fenómeno del castigo estatal.

<sup>20</sup> Morris es el principal exponente de la variante del retribucionismo que se conoce como “*fair play theory*”. Según esta, la pena se justifica como un mecanismo que remueve la ventaja injusta que el infractor ha obtenido, como resultado de no haber ajustado su conducta a lo que prescribe el derecho, respecto del resto de los ciudadanos. Este planteo parte de la premisa de que en una comunidad rigen ciertas normas que garantizan que toda persona posea un espacio libre de interferencias de terceros. Este “beneficio” que cada persona obtiene de la no injerencia de otras personas en su ámbito protegido conlleva la “carga” de no interferir en la esfera de los demás. Cuando alguien comete un delito, incumple su carga y obtiene una “ventaja injusta” por sobre sus conciudadanos, alterando la distribución equitativa de cargas y beneficios que rige en la sociedad. Esa ventaja radica en que el infractor continúa gozando de su propia esfera de inmunidad gracias al autocontrol ejercido por sus conciudadanos, mientras que él mismo no restringe su conducta de manera recíproca. Se convierte, de esta manera, en un *free rider* —alguien que goza de los beneficios del sistema sin pagar su parte—. En este esquema, el castigo del infractor implica la remoción de esa ventaja injusta y el restablecimiento de la adecuada distribución de cargas y beneficios en la sociedad. Véase Herbert Morris “Persons and Punishment”, *The Monist*, Vol. 52, Nro. 4, (Oct. 1968) pp 475-501

<sup>21</sup> De manera similar a Morris, Pawlik pretende justificar la pena a partir de la idea de que esta contribuye al mantenimiento de un estado de libertades. En esta teoría, el castigo resulta permisible en la medida que los infractores incurren en un deber secundario de tolerar la pena al violar, mediante la realización de un ilícito penal, su deber primario de colaborar con el mantenimiento del estado de libertades. Véase Michael Pawlik, *El injusto del ciudadano* (Bogotá, Barcelona: Universidad Externado de Colombia, Atelier, 2023).

<sup>22</sup> Raúl E. Zaffaroni, “Presentación del traductor”, en Anton Bauer, *La teoría de la advertencia* (Ciudad Autónoma de Buenos Aires: Ediar, 2019), p. 16.

sociológica.<sup>23</sup> Pero es claro que, sin perjuicio de la dificultad inherente a determinar cuáles son las razones que se persiguen a través de una práctica colectiva, tales razones no necesariamente coinciden con aquellas que son relevantes para determinar si la práctica en cuestión está justificada moralmente o no.

Este desfasaje entre ambas clases de razones puede ser ilustrado con el siguiente ejemplo. Imagínese una práctica distinta al castigo penal como, por ejemplo, un esquema progresivo de impuestos.<sup>24</sup> Podríamos preguntarnos cuál es el fin que tal esquema persigue y cómo pretende realizarlo. Una respuesta plausible podría ser que los impuestos progresivos buscan aumentar las arcas públicas y promover una distribución equitativa de la riqueza. Sin embargo, incluso si se aceptara que ese es su fin, quedaría abierta la pregunta de si un esquema de esa naturaleza puede justificarse moralmente. Un libertario à la Nozick, por ejemplo, podría admitir que se trata de un sistema efectivo para la consecución de los fines que pretende lograr, y aun así considerar que se trata de un esquema injusto porque implica, esencialmente, una apropiación indebida por parte del Estado de bienes individuales.

En la literatura se pueden encontrar distintas formas de distinguir entre estos dos tipos de razones: esto es, entre aquellas que permiten explicar por qué realizamos una determinada acción o establecemos una determinada práctica, y aquellas que resultan relevantes para evaluar si esa acción o práctica resulta moralmente admisible. Por ejemplo, en el ámbito anglosajón se ha trazado una distinción entre la justificación *racional* del castigo y su justificación moral.<sup>25</sup> Hart tiene en mente una distinción similar al separar el “propósito general justificante” (*general justifying aim*) de la práctica del castigo, y los principios que deben regir la distribución de las penas al interior de esa práctica, una vez instituida.<sup>26</sup> En el marco de la discusión penal continental se distingue, a su vez, entre el “fundamento jurídico” y el “fin real” de la pena estatal.<sup>27</sup>

<sup>23</sup> Luis E. Chiesa, “Taking Victims Seriously: A Dworkinian Theory of Punishment”, *Pace Law Faculty Publications* 561 (2007), disponible en: <https://digitalcommons.pace.edu/cgi/viewcontent.cgi?article=1560&context=lawfaculty> (enlace verificado por última vez el 17 de noviembre de 2025).

<sup>24</sup> David Dolinko, “Some Thoughts About Retributivism”, *Ethics* 101, Nro. 3 (abril 1991), p. 540.

<sup>25</sup> Dolinko, “Some Thoughts About Retributivism”, p. 539.

<sup>26</sup> Hart, *Punishment and Responsibility. Essays in the Philosophy of Law*, p.4.

<sup>27</sup> En este sentido: “Los representantes de las escuelas clásica y moderna no pudieron evitar hablarse sin entenderse porque los ‘clásicos’ se preguntaban siempre tan solo por el fundamento jurídico de la pena estatal e identificaban esa legitimación sin más con el fin efectivamente perseguido por la pena. En cambio, los ‘modernos’ nunca se preocuparon seriamente por la legitimación jurídica de la pena, sino que solo se centraron en sus fines reales, que, empero, en cuanto tales aún no justifican injerencias en la esfera del

Más allá de los matices existentes entre estas distintas clasificaciones, todas ellas apuntan a distinguir la pregunta acerca del fin o propósito de una cierta acción o práctica, de aquella que indaga acerca de su justificación moral. Y es precisamente esta distinción la que está ausente en los enfoques que tratan la cuestión de la justificación del castigo y la cuestión del fin de la pena de manera indistinta, como si se tratara de una misma pregunta. El problema de esta amalgamación es que de la respuesta que brindemos a la pregunta de por qué castigamos no se sigue que el castigo esté justificado o no. Podríamos intentar determinar, mediante complejos experimentos, si las razones por las que castigamos a otros son esencialmente retribucionistas o consecuencialistas,<sup>28</sup> pero el resultado de estos experimentos nada diría acerca de si el castigo que imponemos por esas razones es moralmente admisible o no. Por caso, podríamos descubrir a través de uno de estos experimentos que castigamos motivados por razones retribucionistas. Pero de ello no se seguiría que el retribucionismo constituye una justificación del castigo normativamente aceptable, ni que el castigo impuesto por razones retribucionistas es moralmente permisible. Sostener lo contrario implicaría confundir el plano descriptivo con el normativo, lo cual supone derivar una conclusión acerca de lo que *debería ser* de una observación acerca de lo que *es*.

Ahora bien, lo dicho hasta aquí no implica que no exista ninguna relación entre la finalidad que persigue el castigo y su justificación. En ocasiones, la admisibilidad moral de una acción depende precisamente de la finalidad por la cual es llevada a cabo. Esto puede ilustrarse con un ejemplo: que un padre castigue a su hijo prohibiéndole asistir a una reunión con sus amigos para enseñarle una lección, por haberse portado mal, no parece, en principio, objetable. Sin embargo, que lo haga porque disfruta ver sufrir a su hijo, sí lo es. Existe, por ende, un vínculo relevante entre las razones por las que llevamos adelante determinadas acciones y la permisibilidad de tales acciones.

Dicho esto, si se quiere desentrañar la conexión entre los fines que persigue el castigo y su justificación moral, la pregunta pertinente ya no deberá dirigirse a identificar las razones por las que de hecho castigamos, sino aquellas por las que deberíamos castigar para que el castigo resulte moralmente admisible. Bajo esta nueva formulación, la pregunta acerca de los fines de la pena sí resultará relevante para determinar si la pena

---

ciudadano.” Geerds Schild, *Einzelner und Staatsgewalt im geltenden Strafrecht* 1969), p. 16.; citado en Harro Otto, *Manual de Derecho Penal*, 7ma ed. (Barcelona: Atelier, 2017), p. 40.

<sup>28</sup> Fangfang Tan y Erte Xiao, “Third-Party Punishment: Retribution or Deterrence?”, *Journal of Economic Psychology*, Vol.67, (agosto 2018), pp. 34–46.

está justificada o no. Pero está claro que la respuesta a dicha pregunta no habrá de encontrarse en constataciones empíricas que permitan establecer por qué, en los hechos, castigamos, sino a partir de un argumento normativo tendiente a demostrar la permisibilidad de castigar por ciertas razones.

#### IV. La definición y la justificación: dos asuntos separados

Otra de las preguntas que suele estar entremezclada con las analizadas hasta aquí es aquella relacionada con el concepto de castigo, es decir, la pregunta que indaga acerca de qué es el castigo.

En un sentido importante, la cuestión de la definición del castigo es relevante para el análisis de su justificación. En efecto, en la literatura especializada es común que las reflexiones acerca de la permisibilidad de la pena estén precedidas por la formulación de una definición. Quienes siguen este camino parten de la premisa según la cual, para investigar si el castigo puede justificarse o no, primero es necesario saber qué es el castigo.<sup>29</sup> De acuerdo con esta mirada, la cuestión de la definición posee prioridad lógica respecto del análisis relativo a la justificación. Aunque mayoritaria, esta manera de proceder no es unánime: algunos autores, escépticos respecto de la posibilidad de definir de manera precisa un concepto de contornos difusos como es el castigo, prefieren evitar la empresa definicional. En vez de formular una definición se limitan, a la hora de analizar la permisibilidad de la pena, a identificar los aspectos de este fenómeno que resultan moralmente problemáticos y que, por tanto, requieren una justificación.<sup>30</sup>

No obstante, parecería claro que, sin perjuicio de la conexión que podría haber entre ellas, se trata de dos preguntas distintas, que requieren respuestas independientes.<sup>31</sup> La dilucidación del concepto de castigo supone identificar las condiciones necesarias y suficientes para que algo sea un castigo<sup>32</sup> o, de manera más modesta, identificar las

<sup>29</sup> Boonin, *The Problem of Punishment*, p. 4.

<sup>30</sup> Mitchell Berman, “The Justification of Punishment”, en Andrei Marmor (ed.), *The Routledge Companion to the Philosophy of Law*, (New York: Routledge, 2014), p. 143

<sup>31</sup> Entre los autores hispanoparlantes que han señalado esta distinción entre la definición y la justificación del castigo se destaca Pérez Barberá, con especial énfasis en el rol que ocupa la noción de retribución en cada una de esas dimensiones. Véase, al respecto, Gabriel Pérez Barberá, “Problemas y perspectivas de las teorías expresivas de la pena: Una justificación deontológica de la pena como institución”, *Indret* Nro. 4 (2014), disponible en: <https://indret.com/problemas-y-perspectivas-de-las-teorias-expresivas-de-la-pena/> (enlace verificado por última vez el 17 de noviembre de 2025).

<sup>32</sup> Zaibert, *Punishment and Retribution*, p. 7.

características centrales del caso paradigmático de castigo.<sup>33</sup> Se trata, por ende, de una indagación de índole conceptual. En cambio, la pregunta sobre la justificación del castigo busca averiguar cuándo y por qué razones es moral o políticamente legítimo imponer un castigo.<sup>34</sup> Es, por tanto, una indagación de índole normativa.

Sin embargo, también en este punto son frecuentes las confusiones. El ejemplo más famoso lo brindó Hart, quien introdujo el término “barrera definicional” (“*definitional stop*”),<sup>35</sup> para referirse a la maniobra a través de la cual se pretende clausurar un debate normativo mediante la introducción de una definición estipulativa que excluye la aplicabilidad de una de las posiciones en pugna. Para ilustrar esto, Hart tiene en mente la objeción tradicional que los retribucionistas formulan a los utilitaristas,<sup>36</sup> a saber: si se asume que el castigo se justifica por las consecuencias beneficiosas que produce, ¿por qué deberíamos limitar su aplicación a sujetos culpables de haber cometido un delito? ¿Por qué no castigar a personas elegidas al azar, o incluso a los familiares del delincuente, si ello produjera los mismos beneficios? Frente a esta objeción, una posible respuesta consiste en sostener que el castigo, por definición, solo puede ser aplicado a alguien culpable. Y esta es justamente la respuesta que Hart considera inaceptable, por entender que supone un uso abusivo de una definición. A esa maniobra denomina “barrera definicional”.

El problema de utilizar este recurso es que, en definitiva, lo que se pretende busca a través de él es dar por cerrado un debate normativo mediante una definición conceptual, sin aportar un argumento en apoyo de la tesis sustantiva que se quiere defender en el plano de la justificación. En el ejemplo citado, el utilitarista se defiende de la objeción del retribucionista apelando a una definición —por demás controvertible— del fenómeno cuya justificación es objeto de discusión.

<sup>33</sup> Un ejemplo de esta metodología, pero aplicada al concepto de reproche, puede encontrarse en Miranda Fricker, ¿“What’s the Point of Blame? A Paradigm Based Explanation”, *Nous* 50, Nro.1 (marzo 2016), pp. 165-183.

<sup>34</sup> Zaibert, *Punishment and Retribution*, p. 7.

<sup>35</sup> Hart, *Punishment and Responsibility: Essays in the Philosophy of Law*, p. 5. Pérez Barberá –en la misma línea que Zaibert– nos recuerda, sin embargo, que “...el temor a incurrir en barreras definicionales no debe tampoco conducirnos a renunciar lisa y llanamente a las definiciones. Estas son imprescindibles para poder saber con precisión de qué se está hablando. Trabajar sin definiciones sería, incluso, argumentativamente perezoso (...) Lo que no se puede hacer es blindar la introducción subrepticia de una cuestión normativa relevante mediante el recurso de bloquear la discusión sobre ella a través de una definición restrictiva” Ver Pérez Barberá, “Problemas y perspectivas de las teorías expresivas de la pena...”, p. 4.

<sup>36</sup> Siguiendo la terminología más moderna en la discusión sobre la justificación del castigo, sería preferible utilizar el término “consecuencialistas”. Se respeta aquí el término “utilitaristas” para resguardar la fidelidad con el texto original.

En ocasiones, esta estrategia consistente en recurrir a consideraciones conceptuales para resolver cuestiones normativas se manifiesta de manera más sutil, mediante la pretensión de descubrir cuál es la función de la pena. En estos casos, se argumenta que la pena cumple una determinada función y de ello se derivan consecuencias relativas a su justificación. Un ejemplo de esto lo provee Feinberg. Según este autor, una característica fundamental de la pena es que ella cumple una función expresiva de ciertas actitudes reactivas y juicios de desaprobación de parte de la autoridad que la impone o de la comunidad que esa autoridad representa.<sup>37</sup> De esto Feinberg deriva que una teoría acerca de la justificación del castigo debe dar cuenta de este carácter expresivo de manera satisfactoria. Pero esta afirmación puede ser entendida de manera tal que excluya sin más algunas de las posibles justificaciones del castigo alternativas.

Así, por ejemplo, Ferrante parece interpretar de esta manera a Feinberg cuando sostiene que

si el castigo es una práctica que tiene ese rasgo esencial [refiriéndose a su carácter expresivo], una teoría de la justificación del castigo tiene que hallar razones para involucrarse en una práctica que implique esencialmente esa expresión de condenación. Como las razones pertinentes para justificar la adopción de una actitud de condenación parecen fundamentalmente no consecuencialistas, el hallazgo de la dimensión expresiva del castigo parece poner en jaque a la pretensión de justificar la adopción de la práctica de castigo sobre fundamentos consecuencialistas como los que articula la teoría benthamiana.<sup>38</sup>

De esta manera, la cuestión de la justificación del castigo —ámbito en el que tradicionalmente se han enfrentado el consecuencialismo y el retribucionismo— parecería ya estar saldada en el momento en que se descubre cuál es la función del castigo.

Otro ejemplo de esta estrategia argumentativa puede hallarse en la obra de Michael Moore. Este autor comienza señalando que, ontológicamente, es posible distinguir tres tipos de clases (*kinds*): las clases naturales, las nominales y las funcionales. Las clases naturales están compuestas por elementos que comparten una misma esencia; las clases nominales están compuestas por elementos que simplemente comparten un mismo nombre; las clases funcionales están compuestas por elementos que comparten una función. Así, por ejemplo, los estómagos conforman una clase de este último tipo, cuya función común es digerir la comida, independientemente de cuál sea su forma o

<sup>37</sup> Joel Feinberg, “The Expressive Function of Punishment”, *The Monist* 20, nro. 43 (julio 1965), p. 400.

<sup>38</sup> Marcelo Ferrante, “Filosofía del Derecho penal”, en Jorge Luis Fabra Zamora y Ezequiel Spector (eds.), *Enciclopedia de Filosofía y Teoría del Derecho* (México: Universidad Nacional Autónoma de México, 2015), pp. 2090–2091.

estructura.<sup>39</sup> Moore agrega que identificar la función de una cosa requiere responder dos preguntas, a saber: qué es lo que esa cosa puede generar, y cuál de los efectos que puede producir es intrínsecamente valioso.<sup>40</sup> Siguiendo este método, concluye que el bien intrínsecamente valioso que el derecho penal puede generar y que, por tanto, constituye su función, es la retribución.<sup>41</sup>

Si bien, posteriormente, el autor brinda un argumento normativo independiente a favor del valor intrínseco de la retribución,<sup>42</sup> la discusión relativa a la justificación del castigo parece ya estar resuelta desde el momento en que asigna al derecho penal la función exclusiva de alcanzar la justicia retributiva. El problema común de todos estos enfoques es que, al vincular la pregunta sobre la permisibilidad de la pena con su definición o función, se corre el riesgo de cometer una petición de principio. Pues aquello que debería ser demostrado —i.e., por qué razón puede considerarse que el castigo está justificado— ya está presupuesto desde el inicio. Por ejemplo, si se asumiese que el castigo es un mal que sirve para prevenir futuros males, dicha definición inclinaría la balanza a favor de una justificación consecuencialista y excluiría la plausibilidad de la tesis retribucionista. A la inversa, si se asumiese que la función exclusiva del castigo es imponer retribución, la balanza quedaría inclinada de manera decisiva a favor del retribucionismo.

Para evitar la contaminación del debate normativo con consideraciones conceptuales, es preferible adoptar una definición de castigo que sea justificativamente neutral,<sup>43</sup> es decir, que no presuponga la corrección de ninguna de las tesis en disputa en el ámbito de la justificación.

## V. A modo de ejemplo: la teoría agnóstica de la pena de Zaffaroni, Alagia y Slokar

Trazadas las distinciones anteriores, a continuación, procuraré ilustrar, a partir de un ejemplo, su relevancia para la temática abordada. En su *Manual de Derecho Penal: Parte General*, Zaffaroni, Alagia y Slokar comienzan criticando las teorías positivas de la pena, entre las cuales incluyen la prevención general negativa y positiva, la prevención especial

<sup>39</sup> Michael Moore, *Placing Blame* (Oxford: Oxford University Press, 1997), p. 20.

<sup>40</sup> Moore, *Placing Blame*, p. 26.

<sup>41</sup> Moore, *Placing Blame*, p. 27.

<sup>42</sup> Moore, *Placing Blame*, ca p. 3 (“The Moral Worth of Retribution”).

<sup>43</sup> Boonin, *The Problem of Punishment*, p. 5.

positiva y negativa, y la retribución. Los defensores de estas teorías “creen que el castigo es un bien para la sociedad o para quien sufre la pena.”<sup>44</sup> Y a ello añaden: “[s]e trata de posiciones que hacen acto de fe de alguna función manifiesta y en ella fundan la legitimidad de la pena y de ella deducen una teoría del derecho penal.”<sup>45</sup> Es decir que, según Zaffaroni *et al.*, se trata de discursos que pretenden legitimar la pena a partir de la asignación a esta de alguna función en particular.

Aunque los autores mencionados dedican un tratamiento diferenciado a cada una de las teorías mencionadas, el argumento central que recorre su presentación es que todas ellas son falsas,<sup>46</sup> dado que, en la realidad, no se verifica ninguna de las distintas funciones que les asignan a la pena.<sup>47</sup> A partir de este aparente fracaso de las teorías positivas, proceden entonces a desarrollar un concepto negativo de pena, según el cual la pena es “(a) una coerción, (b) que impone una privación de derechos o un dolor, (c) que no repara ni restituye (d) ni tampoco detiene las lesiones en curso ni neutraliza los peligros inminentes”.<sup>48</sup>

Asimismo, agregan:

El concepto así enunciado se obtiene por exclusión: la pena es un ejercicio de poder que no tiene función reparadora o restitutiva ni es coacción administrativa directa (...) Este concepto de pena es negativo, tanto porque no le asigna ninguna función positiva a la pena como por ser obtenido por exclusión (es la coacción estatal que no entra en el modelo reparador ni en el administrativo directo) (...) [E]s agnóstico en cuanto a su función, pues parte de su desconocimiento.<sup>49</sup>

En consonancia con este concepto de pena, Zaffaroni, Alagia y Slokar proponen una concepción del derecho penal como dique de contención del poder punitivo. Según sugiere la metáfora

El dique debe filtrar ciertas aguas y contener otras, es decir, debe seleccionar y dejar pasar una cierta cantidad de agua, la indispensable para mantener su estabilidad y para evitar que el estado de derecho sea cubierto.<sup>50</sup>

<sup>44</sup> Raúl Eugenio Zaffaroni; Alejandro Alagia; Alejandro Slokar, *Manual de Derecho Penal: Parte General* (Buenos Aires: Ediar, 2009), p. 33.

<sup>45</sup> Zaffaroni; Alagia; Slokar, *Manual de Derecho Penal: Parte General*, p. 33.

<sup>46</sup> Zaffaroni; Alagia; Slokar, *Manual de Derecho Penal: Parte General*, p. 37.

<sup>47</sup> Zaffaroni; Alagia; Slokar, *Manual de Derecho Penal: Parte General*, p. 54.

<sup>48</sup> Zaffaroni; Alagia; Slokar, *Manual de Derecho Penal: Parte General*, p. 56.

<sup>49</sup> Zaffaroni; Alagia; Slokar, *Manual de Derecho Penal: Parte General*, p. 56.

<sup>50</sup> Zaffaroni; Alagia; Slokar, *Manual de Derecho Penal: Parte General*, p. 65.

Al derecho penal le correspondería, por tanto, limitar el poder punitivo lo máximo que sea posible. Pero de esto no se sigue que la porción de poder punitivo que inevitablemente habrá de filtrarse tenga valor positivo alguno.

Pues bien, creo que esta teoría incurre en una cierta confusión de planos que puede ser disipada si se repara en las distinciones elaboradas en los acápite anteriores. En efecto, los autores citados conciben a las teorías positivas de la pena como discursos que pretenden legitimar el castigo apelando a ciertas funciones que este vendría a cumplir. A esto agregan que, como en la realidad no se verifican esas funciones, tales teorías son falsas. Este último paso, sin embargo, supone un salto de niveles. Pues de la corroboración empírica de que el castigo, en los hechos, no cumple una determinada función —o, de manera similar, de la imposibilidad de corroborar empíricamente que cumpla una cierta tarea—, no se sigue la “falsedad” de una teoría normativa que pretenda fundar la permisibilidad moral de la pena apelando a esa función o tarea. Por ejemplo, que no se pueda verificar empíricamente que la pena ayude a prevenir delitos a través de la disuasión no implica que la prevención general negativa resulte falsa. Lo que podría ser falso, en todo caso, es una premisa empírica de la cual la plausibilidad de dicha teoría normativa en última instancia depende. Pero quedaría abierta la posibilidad de que, en otras condiciones —o incluso en bajo condiciones ideales—, la teoría justificatoria en cuestión resulte aceptable.

Al fundar su rechazo de las teorías positivas de la pena en observaciones fácticas (i.e., la pena no repara ni restituye), Zaffaroni, Alagia y Slokar desconocen el carácter normativo de aquellas, y esto hace que, en definitiva, su crítica resulte infundada. Pues poco sentido tiene decir que una teoría de la pena es falsa, en la medida que el objeto de una elaboración de ese tipo no es describir qué es lo que la pena hace o no hace en la realidad, sino ofrecer un argumento para demostrar en qué casos y por qué razones el castigo podría estar justificado. Por supuesto, una teoría que pretenda justificar—o demostrar por qué no puede justificarse— una cierta práctica, debe partir de una reconstrucción adecuada de esta. De lo contrario, la teoría se estaría refiriendo a un fenómeno distinto del que se supone es objeto de examinación. No obstante, un análisis de esta clase debe, además, asumir un punto de vista crítico y no meramente descriptivo. Como sostiene Duff,

nos debe ofrecer no solo una racionalización o justificación del *status quo*, sino un estándar crítico: una concepción ideal de lo que el castigo

debería ser, bajo la cual podamos evaluar (y seguramente encontrar deficiente) nuestras prácticas existentes.<sup>51</sup>

Por otra parte, los defensores de la teoría agnóstica de la pena formulaan, a partir de su oposición a las teorías positivas de la pena, un concepto negativo de pena. Pero aquí se constata un nuevo salto de niveles, dado que la plausibilidad (o implausibilidad) de una determinada posición relativa a la justificación del castigo no debería tener injerencia en la averiguación de aquello que la pena es. Sostener lo contrario, como parecen hacer Zaffaroni, Alagia y Slokar, implica confundir el plano conceptual y el normativo; es decir, mezclar la definición del castigo con su justificación. El resultado de esta confusión es doblemente pernicioso. Por un lado, la definición así alcanzada resulta ad-hoc, es decir, parece estar formulada a los efectos de reflejar el fracaso de las teorías positivas en el plano justificatorio, más que con el fin de esclarecer qué tipo de fenómeno la pena realmente es. Por el otro, la corrección de la tesis ofrecida en el plano de la justificación (en el caso de los autores citados, la tesis agnóstica) se termina derivando necesariamente del concepto negativo de pena formulado.

En efecto, esto es precisamente lo que ocurre cuando se sostiene que no es posible justificar la pena bajo ninguna de las teorías positivas (afirmación normativa) porque la pena no repara, restituye ni neutraliza los peligros inminentes (afirmación conceptual). Y el mismo error se comete cuando, a la inversa, se sostiene que la pena debe ser definida de modo negativo (como un ejercicio de poder que no tiene función reparadora ni restitutiva, ni es coacción administrativa directa) en virtud del fracaso de las teorías positivas en el plano de la justificación.

## VI. Una propuesta dworkiniana

Quisiera ahora explorar, aunque sea de manera superficial, un posible abordaje al problema de la justificación del castigo que no ha sido debidamente examinado en la literatura especializada. Considero que esta alternativa tiene potencial para resolver algunas de las tensiones identificadas en lo que va del trabajo.

El enfoque que propongo pretende trasladar al plano de la justificación del castigo la metodología que Ronald Dworkin desarrolló en sus escritos acerca de la interpretación y la naturaleza del derecho. Como señala con acierto Luis Chiesa, es llamativo que los

<sup>51</sup> R. A. Duff, *Punishment, Communication and Community* (Oxford: Oxford University Press, 2001), p. xv.

filósofos penales no se hayan interesado más en la obra de Dworkin, sobre todo si se tiene en cuenta el estrecho vínculo existente entre el derecho penal, la filosofía moral y la filosofía política, y el rol prominente ocupado por el filósofo estadounidense en estas dos últimas áreas.<sup>52</sup> En su artículo, Chiesa contribuye a subsanar esa omisión, pero lo hace ofreciendo una teoría sustantiva sobre la justificación del castigo inspirada en la obra de Dworkin, la cual pone el énfasis en el rol de las víctimas. Por contraste, en las secciones siguientes pretendo explorar algunas repercusiones metodológicas que podrían derivarse de la aplicación de las contribuciones del filósofo estadounidense a la teorización acerca del castigo. Para ello comenzaré brindando un panorama de la teoría interpretativa desarrollada por este autor, y luego intentaré demostrar la relevancia de dicha teoría para la discusión sobre la justificación del castigo. Por último, me dedicaré a contestar dos posibles objeciones.

### A. La teoría interpretativa de Dworkin

Dworkin concibe al derecho como una práctica social esencialmente interpretativa, en el sentido de que

[c]ada actor en la práctica comprende que aquello que esta permite o requiere depende de la verdad de ciertas proposiciones que tienen sentido sólo dentro de la práctica y por ella; la práctica consiste en gran parte en argumentar sobre estas proposiciones.<sup>53</sup>

Esta concepción del derecho va acompañada de una particular reconstrucción de los desacuerdos que se originan al interior de esa práctica, entre quienes participan en ella. En este sentido, el autor distingue entre desacuerdos empíricos, es decir, aquellos que versan sobre ciertos hechos —por ejemplo, sobre si un código contiene tal o cual norma, o sobre cuál es el significado convencional de un cierto término—, y desacuerdos teóricos, los cuales recaen sobre los fundamentos del derecho —por ejemplo, sobre si las normas contenidas en un determinado código agotan los fundamentos del derecho para resolver un cierto caso—.<sup>54</sup> Pues bien, según esta visión, cuando los participantes en la práctica del derecho (jueces, abogados, etc.) mantienen desacuerdos de este último tipo, lo que hacen es intercambiar argumentos que son, en última instancia, argumentos morales, y que sirven de apoyo a una determinada concepción que permita reflejar por

<sup>52</sup> Chiesa, “Taking Victims Seriously: A Dworkinian Theory of Punishment”, p. 117.

<sup>53</sup> Ronald Dworkin, *El imperio del derecho* (trads. Juan Iosa y Tomás Céspedes) (Barcelona: Gedisa, 2022), p. 16.

<sup>54</sup> Dworkin, *El imperio del derecho*, p. 9.

qué la práctica, así concebida, resulta valiosa; es decir, por qué vale la pena tener una práctica de esa naturaleza. Así,

cuando los participantes argumentan a favor de determinada comprensión de las exigencias de la práctica, lo que están haciendo es interpretarla, i.e., construir hipótesis sobre por qué tenemos una actividad así, hipótesis que la justifiquen, que nos la muestren como valiosa. Y esta tarea no es, no puede ser, sólo descriptiva. Requiere inevitablemente que recurramos a consideraciones morales.<sup>55</sup>

Para Dworkin, quienes realizan este tipo de ejercicio interpretativo no aspiran meramente a identificar cuál fue la intención de una persona al utilizar ciertas palabras, como ocurre en el contexto de una conversación, o cuál es la forma correcta de procesar ciertos datos de la realidad, como ocurre en una interpretación científica. Antes bien, en el contexto de los desacuerdos teóricos acerca de los fundamentos del derecho, los intérpretes realizan un tipo de interpretación creativa; es decir, una interpretación que tiene como finalidad dotar de sentido al objeto que es interpretado.<sup>56</sup>

La interpretación creativa supone la adopción de una “actitud interpretativa” por parte del intérprete. Esta actitud implica, por un lado, la asunción de que la práctica objeto de interpretación es en algún sentido valiosa y, por el otro, que las reglas o exigencias de la práctica no son inmutables y estáticas, sino más bien dinámicas y sensibles al valor previamente identificado.<sup>57</sup> De acuerdo con ello, el intérprete puede reinterpretar esas reglas o exigencias a la luz de lo que el valor del objeto requiere. Esto es justamente lo que implica interpretar la práctica bajo su mejor luz.

Señala Dworkin que, a la hora de hacer esto último, el intérprete procede en tres pasos interpretativos. La primera es la etapa preinterpretativa y consiste en la identificación de los insumos normativos relevantes; es decir, en la selección de “las reglas y estándares que proporcionan el contenido provvisorio de la práctica”.<sup>58</sup> La segunda es la etapa interpretativa propiamente dicha y consiste en el establecimiento “de una justificación general para los principales elementos de la práctica identificada en la etapa preinterpretativa”,<sup>59</sup> lo cual requiere la formulación de un argumento sobre por qué es valiosa una práctica con esa estructura general. La tercera es la etapa postinterpretativa o de reforma, en la cual el intérprete “ajusta su sentido sobre qué requiere ‘en realidad’ la

<sup>55</sup> Juan Iosa; Pablo A. Rapetti, “Estudio introductorio”, en Ronald Dworkin, *El imperio del derecho* (Barcelona: Gedisa, 2022), p. xxxiii.

<sup>56</sup> Dworkin, *El imperio del derecho*, pp. 51–52.

<sup>57</sup> Dworkin, *El imperio del derecho*, p. 49.

<sup>58</sup> Dworkin, *El imperio del derecho*, p. 68.

<sup>59</sup> Dworkin, *El imperio del derecho*, p. 68.

práctica a los fines de servir mejor a la justificación aceptada en la etapa interpretativa”.<sup>60</sup> Dworkin resume estas tres etapas señalando que “la interpretación busca establecer un equilibrio entre la descripción preinterpretativa de una práctica social y una justificación adecuada de dicha práctica.”<sup>61</sup>

Esta mirada general se corresponde con una concepción de la interpretación creativa como interpretación constructiva, la cual consiste en asignar un propósito a un objeto o práctica para hacer del mismo la mejor instancia posible de la forma o género al cual pertenece.<sup>62</sup> El criterio de corrección de una interpretación de este tipo se vincula con cuán exitosa sea ella en satisfacer los requisitos que se derivan de dos dimensiones evaluativas que Dworkin llama ajuste y justificación, respectivamente. De acuerdo con ellas, la interpretación en cuestión debe, por un lado, resultar congruente con los materiales jurídicos relevantes y poder dar cuenta de estos. Por el otro, los debe presentar bajo su mejor luz. La dimensión de ajuste opera entonces como un primer gran filtro que permite descartar aquellas interpretaciones que se contrapongan o que resulten inconsistentes con los insumos normativos relevantes, los cuales dependerán del caso concreto, pero pueden tratarse de leyes, precedentes jurisprudenciales, principios constitucionales, etc. Así,

las convicciones sobre ajuste proporcionarán un umbral grueso que cualquier interpretación de alguna parte del derecho deberá cumplir para ser siquiera elegible.<sup>63</sup>

Pero es posible que el umbral de ajuste resulte insuficiente por sí solo para determinar cuál es la interpretación correcta entre varias que resultan igualmente plausibles de acuerdo con ese parámetro. En tales casos, el intérprete deberá escoger aquella que ofrezca la mejor justificación moral de la práctica jurídica:

[d]eberá entonces decidir entre las interpretaciones elegibles preguntándose cuál muestra la estructura comunitaria de instituciones y decisiones (sus estándares públicos como un todo) bajo una mejor luz desde la perspectiva de la moralidad política.<sup>64</sup>

Como advierten Juan Iosa y Pablo A. Rapetti en el estudio preliminar que acompaña la reciente traducción, a manos del primero y de Tomás Céspedes, de *El imperio del*

<sup>60</sup> Dworkin, *El imperio del derecho*, p. 69.

<sup>61</sup> Dworkin, *El imperio del derecho*, p. 69.

<sup>62</sup> Dworkin, *El imperio del derecho*, p. 54.

<sup>63</sup> Dworkin, *El imperio del derecho*, p. 253.

<sup>64</sup> Dworkin, *El imperio del derecho*, p. 253.

*derecho* —la obra en la que Dworkin desarrolla con más extensión su teoría interpretativa—, la dimensión de justificación supone que

de entre distintos esquemas de principios (esto es, distintos principios y conjuntos de principios articulados) que pueden proporcionar una justificación global de una práctica jurídica, el intérprete debe elegir el que la muestre ‘en su mejor luz’, el que haga del derecho el mejor ejemplar del tipo más amplio de práctica al que genéricamente pertenece (e.g., la práctica de ordenación y guía de la conducta de las personas). En este sentido, se trata de dar con la mejor justificación. Y esta es una cuestión de convicciones sustantivas, de justificación moral.<sup>65</sup>

La distinción entre ambas dimensiones, sin embargo, no es tajante, dado que las consideraciones formales y estructurales que cobran protagonismo en la primera dimensión también son relevantes en la segunda. No obstante, el aspecto crucial es que la dimensión de ajuste restringe de manera significativa la dimensión de justificación. Esto es así porque el intérprete no puede hacer prevalecer sus propias convicciones morales por sobre lo que la historia de la práctica institucional demanda. Ninguna interpretación que no se adecúe a los antecedentes relevantes resultará admisible por más que, en términos morales sustantivos, sea la que el juez considera deseable. Conforme lo expuesto, y en relación con el sistema jurídico estadounidense, Dworkin ejemplifica:

incluso un juez que cree que la justicia en abstracto demanda igualdad económica no puede interpretar que la cláusula de protección igualitaria [*equal protection clause*] establezca que la igualdad de riqueza o la propiedad colectiva de los recursos productivos son requerimientos constitucionales, porque esa interpretación simplemente no se adecúa a la historia ni la práctica estadounidense, ni al resto de la Constitución.<sup>66</sup>

Para ilustrar la forma en que las dimensiones de ajuste y justificación se entrelazan en la interpretación constructiva, el autor recurre a la famosa analogía de la novela en cadena. De acuerdo con ella, la tarea de los jueces al aplicar el derecho —y, en especial, al tener que resolver los casos difíciles, es decir, aquellos en los que el derecho no prevé una solución evidente— es asimilable a la que realizarían los sucesivos autores de una novela que se escribe de manera ininterrumpida. Lo que ilustra con elocuencia la analogía es que, en tal supuesto hipotético, cada autor que escribe un nuevo capítulo debería asumir el compromiso de hacer avanzar la novela del mejor modo posible, lo cual exige tener en cuenta todo lo que se escribió con anterioridad. El intérprete debería entonces, en primer lugar, identificar a qué género pertenece la novela que le toca continuar, para luego poder

<sup>65</sup> Iosa, “Estudio introductorio”, p. xxxviii.

<sup>66</sup> Ronald Dworkin, *Freedom’s Law: The Moral Reading of the Constitution* (Oxford: Oxford University Press, 1996), p. 11.

plasmar su aporte con miras a convertirla en el mejor exponente de ese género. Idealmente, de acuerdo con estos parámetros, la novela debería lucir como la obra de un solo autor.<sup>67</sup>

La teoría interpretativa de Dworkin, que en líneas muy generales se acaba de esbozar, es parte de su ataque más general contra el positivismo jurídico. Este tuvo su primer embate en *The model of rules*,<sup>68</sup> con el cargo de que la concepción positivista resultaba incapaz de dar cuenta de los principios que forman parte del derecho en virtud de su contenido moral y no de su *pedigree* institucional. Posteriormente, en *El imperio del derecho*, el autor argumentó que el positivismo, dado su énfasis en el carácter convencional de los criterios de identificación del derecho, no puede explicar satisfactoriamente la persistencia de los desacuerdos teóricos respecto de los fundamentos del derecho.<sup>69</sup>

La posición defendida por Dworkin en estas y otras de sus obras tiene importantes ramificaciones teóricas. Entre ellas se destacan la tesis de la única respuesta correcta, y la defensa de una particular posición en materia metaética que, sin abrazar el realismo moral, constituye una forma de objetivismo moderado, al rechazar el escepticismo externo. No llama la atención, por tanto, que las contribuciones de Dworkin hayan sido tan influyentes en materias tales como la teoría general del derecho, la interpretación judicial, la ética normativa, la metaética, y la filosofía moral y política.<sup>70</sup>

Por el contrario, la obra de este autor no alcanzó un nivel similar de popularidad en el marco de la discusión entre los filósofos penales relativa al problema de la justificación del castigo. Sin embargo, esto se puede explicar fácilmente por el hecho de que dicho problema no formó parte de sus intereses teóricos. Su teoría, por lo tanto, no buscó dar respuesta a los interrogantes que se plantean al interior de la discusión sobre el castigo. No obstante, y siendo consciente de que mi propuesta implica remover a la teoría interpretativa de Dworkin de su hábitat natural, creo que dicha teoría puede resultar iluminadora en el contexto de la discusión acerca de la justificación del castigo. Esto es lo que intentaré demostrar a continuación.

<sup>67</sup> Dworkin, *El imperio del derecho*, pp. 228–231.

<sup>68</sup> Ronald Dworkin, *Taking Rights Seriously* (Cambridge (Mass.): Harvard University Press, 1977)

<sup>69</sup> Scott Shapiro, “The ‘Hart-Dworkin Debate’: A Short Guide for the Perplexed”, *Public Law and Legal Theory Working Paper Series*, University of Michigan Law School, Working Paper Nro. 77 (marzo 2007),

<sup>70</sup> Muestra de esto es que Dworkin sea el tercer jurista estadounidense más citado de todos los tiempos. Véase Fred R. Shapiro, “The Most Cited Legal Scholars Revisited”, *The University of Chicago Law Review* 88, nro. 1 (2021), p. 1602.

**B. Un abordaje dworkiniano al problema de la justificación del castigo**

Supongamos que el filósofo interesado en el problema de la justificación del castigo decidiera abordar esa temática utilizando las herramientas halladas en la teoría interpretativa de Dworkin. ¿Cómo debería proceder?

Según entiendo, deberá aproximarse a su objeto de estudio —la institución del castigo penal— como el intérprete dworkiniano se aproxima al derecho. Por lo tanto, deberá adoptar una actitud interpretativa; es decir, deberá asumir que la práctica bajo examen resulta en algún sentido valiosa, y que la configuración concreta de sus reglas puede variar de acuerdo con lo requiera ese valor. Asimismo, en un primer momento deberá identificar cuáles son las reglas y estándares que, en términos generales, estructuran dicha práctica. Para hacer esto, deberá tener en cuenta ciertos principios elementales que rigen y orientan la aplicación de las penas en las sociedades contemporáneas. Principios básicos del derecho penal liberal como el principio de culpabilidad y el principio de proporcionalidad de las penas habrán de ser entonces considerados, junto a otros rasgos centrales de la práctica, como, por ejemplo, la modalidad de las penas y los criterios utilizados comúnmente para seleccionar la clase de conductas que pueden ser criminalizadas. Luego, deberá argumentar por qué una práctica estructurada sobre los principios previamente identificados es valiosa, si es que lo es. Posteriormente, procederá a reinterpretar esa práctica bajo su mejor luz. Dado que el criterio de corrección de su teoría lo proveen las dimensiones de ajuste y justificación, esta deberá resultar coherente con los materiales normativos relevantes. Por ende, deberá poder explicar por qué la práctica del castigo penal está estructurada de la forma en que lo está. Pero, además, deberá ofrecer la mejor justificación moral disponible de esa práctica.

Llegado este punto cabría preguntarse: ¿cuál serían las ventajas de este abordaje? A continuación, menciono algunas. En primer lugar, esta perspectiva se ubicaría en un punto intermedio entre las posiciones que, desde la filosofía política, se enfocan en el problema de la legitimidad del Estado para castigar y aquellas que, desde la filosofía moral, se enfocan en el problema que supone la irrogación de sufrimiento que caracteriza a todo castigo. Al concebir a su objeto de estudio como práctica social, el filósofo penal dworkiniano deberá ir más allá del nivel de personas castigando personas y enfocarse en el castigo como fenómeno institucional y autoritativo, gobernado por ciertas reglas y principios de aplicación general. Su interpretación de la práctica deberá explicar por qué

resultaría deseable tener una institución de ese tipo, estructurada bajo esos principios. A su vez, deberá identificar cuál es el valor que dicha práctica tiene y reinterpretar sus exigencias a la luz de ese valor. Esto requerirá, inevitablemente, apelar a argumentos morales. A la hora de hacer todo esto, el intérprete deberá recurrir a consideraciones de moralidad interpersonal y política.

Resulta interesante destacar que, si bien he enfatizado otros aspectos de la teoría de Dworkin, la perspectiva que aquí se propone es consistente con la tesis de la unidad del valor que este autor defiende en otros trabajos, de acuerdo con la cual la ética, la moralidad política y el derecho forman parte de un todo coherente.<sup>71</sup> Por lo tanto, una primera ventaja de este abordaje es que permite evitar la fragmentación un tanto artificial de algunas de las preguntas que se suscitan en rededor del problema de la justificación del castigo, al darles un tratamiento unificado como cuestiones de las que es necesario dar cuenta para ofrecer una interpretación de la práctica examinada en su mejor luz.

En segundo lugar, la aproximación que aquí se propone permite evitar la confusión entre la descripción de la práctica del castigo penal y su justificación. En este sentido, resulta una alternativa superadora de aquellas teorías que pretenden llegar a conclusiones normativas (como, por ejemplo, “no es posible justificar el castigo”) mediante constataciones empíricas (como, por ejemplo, “en los hechos, la pena no disuade ni repara). A su vez, resulta preferible a aquellas que, al discurrir plenamente en un nivel normativo, terminan justificando una práctica que no se parece en nada a aquellas que realmente existen (y que son las que exigen una justificación en primer lugar). Este delicado equilibrio entre la aspiración de ofrecer un argumento justificatorio que pueda ser utilizado como estándar crítico para evaluar las prácticas reales y la reconstrucción fiel de esas prácticas, de manera tal que el producto final que termina siendo justificado no se aleje demasiado de aquel que motiva la necesidad de justificación, es adecuadamente capturado por los estándares de ajuste y justificación. De acuerdo con ellos, la interpretación correcta deberá poder explicar los elementos centrales de la práctica bajo examen y, además, justificarlos moralmente. Una interpretación de la práctica que no refleje sus características básicas simplemente fallaría en la dimensión de ajuste. Pero, a la vez, la etapa postinterpretativa abre la posibilidad de reinterpretar las exigencias y reglas de la práctica analizada a la luz del valor identificado en la etapa

<sup>71</sup> Ronald Dworkin, *Justice for Hedgehogs* (Cambridge (Mass.): The Belknap Press of Harvard University Press, 2011).

interpretativa. La interpretación constructiva de la práctica del castigo consigue, por tanto, armonizar los aspectos descriptivos y normativos que suelen entremezclarse en la discusión acerca de la justificación del castigo.

En tercer lugar, el abordaje defendido evita recurrir a una barrera definicional para clausurar de manera prematura el debate normativo relevante, como ocurre en aquellas teorías que no distinguen adecuadamente la definición del castigo y su justificación. Esto es así porque, de acuerdo con el enfoque propuesto, la corrección de una determinada posición sustantiva sobre el problema de la justificación del castigo —como podría ser el retribucionismo, la prevención general positiva, etc.— no se evalúa en función de su compatibilidad con una definición puntual de castigo, ni con una concepción en particular respecto de cuál es la función de aquel. El problema de hacer esto último, habíamos señalado, es que la pregunta de la justificación parece ya estar resuelta al momento en que se ofrece una definición estipulativa de castigo que excluye de antemano la plausibilidad de alguna de las posiciones justificatorias en pugna. Y lo mismo sucede cuando se pretende “descubrir” cuál es la verdadera función del castigo. En el caso del planteo dworkiniano que aquí se defiende esto no ocurre porque el teórico no comienza su indagación formulando una definición restrictiva del término ‘castigo’ o pretendiendo esclarecer cuál es la función de ese fenómeno, sino realizando una interpretación constructiva de la práctica penal, es decir, asignándole a esa práctica un propósito con el fin de volverla el mejor exponente del género al que se considera que pertenece. Y esto, a su vez, requiere de un argumento moral y no depende de un postulado metafísico sobre la naturaleza del castigo ni de lo que pueda arrojar un análisis lingüístico de la palabra castigo. En el marco de este ejercicio, es posible que la interpretación de la práctica que realiza un autor excluya alguna de las posiciones tradicionales que se enfrentan en el ámbito de la justificación. Podría darse este supuesto si, por ejemplo, la interpretación en cuestión condujese a descartar la prevención especial positiva al postular que dicha teoría no se ajusta a los materiales normativos relevantes. Pero, incluso en este caso, la interpretación de la práctica en esos términos sería objeto de debate y requeriría un argumento moral sustantivo. De esta manera, se evita la “trampa” de esquivar la discusión normativa relevante mediante afirmaciones conceptuales.

En definitiva, considero que la metodología inspirada en la teoría interpretativa de Dworkin que he propuesto como modelo para teorizar acerca del problema de la justificación del castigo contiene herramientas para resolver de manera satisfactoria

algunas tensiones comúnmente halladas en la literatura, de cuya presentación me encargué en los puntos II, III y IV de este trabajo.

### C. Dos posibles objeciones

Quisiera, antes de finalizar, responder dos posibles objeciones a mi planteo. La primera sostiene que las elaboraciones de Dworkin sobre la interpretación no resultan aplicables al problema de la justificación del castigo. La teoría de Dworkin dice la objeción, ofrece una reconstrucción del tipo de razonamiento que efectúan los jueces al tener que identificar qué es lo que el derecho dispone en un caso determinado. Por ende, no resulta aplicable al tipo de tarea que realiza un filósofo cuando se pregunta acerca de los fundamentos morales de la práctica del castigo. En especial, la idea de la escritura de una novela en cadena —y el énfasis en el valor de la integridad que ella conlleva— resulta particularmente extraña a la empresa del filósofo penal.<sup>72</sup>

De acuerdo con esta objeción, la inadecuación de la teoría interpretativa de Dworkin en el marco de la discusión sobre la justificación del castigo se manifestaría en tres frentes: respecto del sujeto que realiza la interpretación, respecto de la práctica que es examinada y respecto del tipo de operación realizada. Creo, sin embargo, que la objeción fracasa.

En primer lugar, y en lo que respecta al sujeto, Dworkin menciona explícitamente que su argumento adopta la perspectiva de los jueces, por ser la que de manera más evidente ilustra el carácter argumentativo de la práctica jurídica, pero aclara que también podría haber adoptado el punto de vista de los ciudadanos, los políticos y los profesores de derecho.<sup>73</sup> Aun así, el objetor podría replicar que existe una diferencia fundamental entre la perspectiva del juez y la del filósofo, dado que la primera es interna, es decir, asume el punto de vista de un participante que acepta y se siente obligado por el sistema normativo que debe aplicar; mientras que la segunda es externa, en el sentido de que adopta el punto de vista de un observador, que no necesariamente asume una posición con respecto al valor o la obligatoriedad del fenómeno que examina.

El problema de esta respuesta es que presupone la relevancia de la distinción entre el punto de vista interno y el punto de vista externo —y la caracterización de la actividad del filósofo como una indagación que tiene lugar desde el punto de vista externo—. Pero

<sup>72</sup> Agradezco a Roberto Gargarella por plantearme esta objeción.

<sup>73</sup> Dworkin, *El imperio del derecho*, p. 18.

este es un aspecto que el mismo Dworkin cuestiona. En efecto, el autor sostiene que un filósofo que elabora una teoría general del derecho realiza, al igual que un juez que resuelve un caso difícil, una interpretación constructiva, toda vez que las teorías generales del derecho “...tratan de mostrar la práctica jurídica en su mejor luz, lograr un equilibrio entre la práctica jurídica tal como ellas la encuentran y la mejor justificación de dicha práctica”.<sup>74</sup> Y a partir de esto concluye que “...ninguna línea firme divide la teoría del derecho de la adjudicación o de cualquier otro aspecto de la práctica jurídica.”<sup>75</sup>

Esto demuestra que, para Dworkin, la distinción entre un punto de vista interno y otro externo, central en la obra de Hart,<sup>76</sup> y que serviría para distinguir la tarea del juez de la tarea del filósofo, pierde relevancia, dado que también la actividad del filósofo del derecho supone la adopción de una posición justificatoria respecto de la práctica que es examinada. De hecho, es el mismo Hart quien advierte esto, al señalar que su proyecto, que es descriptivo, general y supone la adopción del punto de vista externo, “...es una empresa radicalmente distinta de la concepción de la teoría del derecho de Dworkin (o ‘jurisprudencia’ como él a veces la llama), que es en parte evaluativa y justificatoria y ‘dirigida a una cultura jurídica en particular’”.<sup>77</sup> Ahora bien, si esto es así, poco sentido tiene criticar mi propuesta como una extensión indebida de la teoría de Dworkin —al querer trasladarla desde la perspectiva interna (propia del juez) a la perspectiva externa (propia del filósofo)—, dado que el mismo pensador concibe a la actividad del filósofo como una que asume el punto de vista interno.

No obstante, el objector podría insistir en que el problema de mi propuesta no radica en el sujeto que realiza la interpretación —o en el punto de vista desde el cual lo hace—, sino en el objeto que es interpretado. De esta manera, la objeción consistiría en que la teoría interpretativa de Dworkin es aplicable al derecho, pero no al castigo penal. Creo, sin embargo, que también aquí fracasa la objeción. Pues, según este autor, la interpretación constructiva resulta procedente no solo en el ámbito del derecho sino también cuando el objeto interpretado es una práctica social. Ello es así porque, en el caso de esta última, y a diferencia de lo que ocurre en el contexto de una conversación entre dos personas, la interpretación no consiste en descubrir la intención de nadie, sino más bien en atribuirle un propósito a la práctica, de manera de volverla el mejor ejemplo

<sup>74</sup> Dworkin, *El imperio del derecho*, p. 94.

<sup>75</sup> Dworkin, *El imperio del derecho*, p. 94.

<sup>76</sup> H. L. A. Hart, *The Concept of Law* (Second Edition) (Oxford: Oxford University Press, 1994)

<sup>77</sup> Hart, *The Concept of Law*, p. 240.

posible del género al que pertenece.<sup>78</sup> De hecho, el ejemplo con el que Dworkin comienza el desarrollo de la noción de interpretación creativa —que luego reformula en términos de interpretación constructiva— es una práctica social distinta del derecho, como es la cortesía. Dado que también el castigo penal es una práctica social susceptible de ser interpretada de esta forma, no resulta convincente la objeción que acusa a mi posición de trasladar a ámbitos extraños una teoría que está pensada exclusivamente para dar cuenta del derecho como fenómeno general.

Sin embargo, la objeción no se acaba aquí. El objetor podría aún señalar que, si bien no hay ningún inconveniente en que sea un filósofo (y no un juez) quien realiza una interpretación constructiva de la práctica del castigo penal (y no del derecho), el problema radica en que la empresa del teórico interesado en el problema de la justificación moral del castigo es radicalmente distinta a la que tiene en mente Dworkin cuando desarrolla su teoría de la interpretación. En particular, continua la objeción, es difícil concebir el intento de arribar a una justificación del castigo como la escritura de una novela en cadena. Las dificultades son varias: ¿cuáles serían los anteriores capítulos que el filósofo interesado en el castigo debería continuar? ¿Los trabajos de otros filósofos? ¿Por qué no podría el filósofo comenzar una nueva novela desde cero, ofreciendo una justificación novedosa de la práctica que sea superior a todas las que vinieron antes?

Creo que esta última variante de la objeción es la más fuerte y se centra en un aspecto importante. Parecería, en efecto, que la idea de la escritura de una novela en cadena no refleja de manera adecuada el tipo de tarea —ni los estándares de corrección que sobre ella rigen— que realizan los filósofos penales, ni ofrece un ideal particularmente atractivo de dicha tarea. Esta disimilitud puede explicarse mediante el papel que juega el valor de la integridad en uno y otro ámbito. Dworkin enfatiza el papel que cumple este valor en el marco de un sistema jurídico. La idea central es que

el gobierno hable con una sola voz, que actúe en base a los mismos principios y de una forma coherente con todos sus ciudadanos, que extienda a cada uno de ellos los estándares sustantivos de justicia o equidad que utiliza con algunos.<sup>79</sup>

Y luego argumenta que

<sup>78</sup> Iosa, “Estudio introductorio”, p. xxxv.

<sup>79</sup> Dworkin, *El imperio del derecho*, p. 166.

la integridad es la llave de la mejor interpretación constructiva de nuestras prácticas jurídicas y, en particular, de la forma en que nuestros jueces deciden los casos difíciles.<sup>80</sup>

Pero esto no parece reproducirse en el ámbito de la discusión acerca de la justificación del castigo. Si bien la integridad constituye un valor relevante en el marco de la aplicación del castigo, la pregunta que interesa al filósofo que indaga acerca de la permisibilidad moral de ese fenómeno tiene carácter previo, a saber: ¿puede la práctica del castigo penal justificarse moralmente? Responder esta pregunta requiere un argumento tendiente a demostrar que se trata de una práctica valiosa, pero el valor relevante no puede residir en la aplicación coherente de los criterios con los que se rige la aplicación de una pena porque la admisibilidad de la práctica misma en cuyo marco se aplica una pena es lo que está bajo discusión. Como señalan Iosa y Rapetti,

la integridad (...) es una virtud política que no funciona en el marco de una teoría política utópica o ideal, sino en el marco de un sistema jurídico ya establecido.<sup>81</sup>

Si esto es así, parecería que la integridad entonces no cumple un papel central en la discusión sobre la justificación de la práctica del castigo penal, puesto que al menos una de las preguntas relevantes en el marco de esta discusión es si deberíamos establecer una práctica de esa naturaleza o no. Dada esta asimetría entre el rol que ocupa la integridad en el derecho y en la práctica del castigo, es razonable concluir que la tarea del filósofo del derecho penal no se ve constreñida por los antecedentes relevantes —los anteriores capítulos de la novela— del mismo modo en que lo está la tarea del juez al resolver un caso difícil.

Estoy dispuesto a conceder parcialmente esta objeción y admitir que, por lo menos en el aspecto identificado, hay una importante diferencia entre el ámbito de aplicación que tiene en mente Dworkin al desarrollar su teoría interpretativa y el tipo de tarea que realiza el filósofo interesado en el problema de la justificación del castigo. No obstante, no creo que reconocer la existencia de estos matices invalide por completo el argumento que he presentado, pues las ventajas del abordaje propuesto —identificadas en el punto anterior— subsisten. En particular, la posibilidad de aproximarse a la discusión sobre la justificación del castigo como lo haría el intérprete dworkiniano, es decir, mediante una interpretación constructiva tendiente a presentar la práctica del castigo en su mejor luz —

<sup>80</sup> Dworkin, *El imperio del derecho*, p. 216.

<sup>81</sup> Juan Iosa; Pablo A. Rapett, “Estudio introductorio”, p. xliv.

lo cual supone que la interpretación propuesta satisface en cierta medida los requisitos de ajuste y justificación— sigue resultando atractivo, por las ventajas ya destacadas.

Ahora bien, si la influencia de Dworkin se limita a esto último, ¿es necesario recurrir a este autor? Este es, básicamente, el contenido de la segunda objeción que quisiera responder.<sup>82</sup> Según ella, una vez trazadas las diferencias relevantes entre el planteo de Dworkin y el que aquí propongo, las conexiones remanentes son tan modestas que no justifican la calificación del abordaje defendido como “dworkiniano”. Bastaría con remarcar que la tarea del filósofo del castigo tiene aspectos descriptivos y normativos (como, por ejemplo, sostiene Duff)<sup>83</sup> sin tener que recurrir a Dworkin.

Creo que esta objeción tampoco es convincente dado que, más allá de las disimilitudes identificadas entre el proyecto de Dworkin y el que aquí se defiende, la metodología que he propuesto para analizar el problema del castigo se nutre de conceptos que son lo suficientemente distintivos de la obra de Dworkin —como los de interpretación constructiva, ajuste y justificación, etc.— para justificar que se haga explícita la influencia de este autor. Así como podría calificarse como “rawlsiano” un abordaje que asigna un papel central a los conceptos de “equilibrio reflexivo” o “velo de ignorancia”, el enfoque defendido admite entonces ser definido como “dworkiniano”. Y si bien es cierto que se podría prescindir de la obra de este autor para elaborar los argumentos centrales que he presentado en este artículo, considero que aquellos se nutren de conceptos que han sido popularizados y desarrollados con mayor fineza en dicha obra. Esto solo, a mi juicio, provee una razón de peso para indagar en la obra del profesor estadounidense si lo que se busca es avanzar en la discusión teórica sobre la justificación del castigo.

## VII. Conclusión

Comencé señalando que, sin perjuicio de los ríos de tinta que han corrido sobre el tema, no siempre resulta claro qué es lo que se está discutiendo cuando se examina el problema de la justificación del castigo. Esta imprecisión, sostuve, se debe en parte a que es común que se mezclen preguntas de distinta índole (descriptivas, normativas, conceptuales), sin que los propios autores que las formulan lo adviertan. Luego de distinguir las distintas cuestiones que suelen abordarse bajo el rótulo de “teorías de la pena” y de señalar algunos de los problemas que trae aparejada su confusión, ofrecí un

<sup>82</sup> Agradezco a Claudina Orunesu por plantearme esta objeción

<sup>83</sup> Duff, *Punishment, Communication, and Community*, p. xv.

ejemplo de esos problemas refiriéndome a la posición defendida por Zaffaroni, Alagia y Slokar en su Manual de Derecho Penal.

En la segunda parte del artículo, presenté una alternativa metodológica para teorizar sobre el problema que plantea la justificación del castigo, inspirada en la teoría de la interpretación desarrollada por Ronald Dworkin. Según argumenté, el filósofo penal debería aproximarse a su objeto de estudio —la práctica del castigo penal— del mismo modo en que el juez dworkiniano concibe al derecho. Es decir, interpretando dicha práctica de manera tal que se la presente como la mejor exponente del género al que pertenece. Este ejercicio posee elementos tanto descriptivos como normativos y se evalúa en función de los parámetros de ajuste y justificación. Tales características, sostuve, permiten resolver de modo armónico algunas de las tensiones comúnmente encontradas en la literatura, de cuya identificación me ocupé en la primera parte del trabajo.

Por último, defendí mi propuesta frente a dos posibles objeciones, argumentando, por un lado, que la teoría de la interpretación de Dworkin resulta útil para abordar el problema de la justificación del castigo y, por otro, que el abordaje resultante merece ser calificado como “dworkiniano”.

Me he limitado, entonces, a señalar las ventajas que conlleva, desde un punto de vista metodológico, aproximarse al problema del castigo con algunas de las herramientas teóricas elaboradas por Dworkin. Los aspectos sustantivos, o de contenido, de una teoría dworkiniana de la pena deberán quedar para otra ocasión.

## Bibliografía

- Alchourrón, Carlos E.; Bulygin, Eugenio. *Sistemas normativos* (2<sup>a</sup> ed.). Buenos Aires: Astrea, 2012.
- Bauer, Anton. *La teoría de la advertencia*. Ciudad Autónoma de Buenos Aires: Ediar, 2019.
- Berman, Mitchell. “Punishment and Justification”. *Ethics*, 118, nro. 2 (enero 2008): pp. 258-290.
- Berman, Mitchell. “Two Kinds of Retributivism”. En Duff, R. A.; Green, Stuart (eds.), *Philosophical Foundations of Criminal Law*. Oxford: Oxford University Press, 2011, pp. 433-461.

- Berman, Mitchell. "The Justification of Punishment". En Marmor, Andrei (ed.), *The Routledge Companion to the Philosophy of Law*. New York: Routledge, 2014, pp. 141-157.
- Binder, Guyora. "Punishment Theory: Political or Moral?". Buffalo Criminal Law Review, nro. 5 (2022).
- Boonin, David. *The Problem of Punishment*. Cambridge: Cambridge University Press, 2008.
- Braithwaite, John, y Philip Pettit. *No sólo su merecido. Por una justicia penal que vaya más allá del castigo*. Buenos Aires: Siglo Veintiuno Editores, 2015.
- Caruso, Gregg. *Rejecting Retributivism*. Cambridge: Cambridge University Press, 2021.
- Chiesa, Luis E. "Taking Victims Seriously: A Dworkinian Theory of Punishment". *Pace Law Faculty Publications*, 561 (2007).
- Dolinko, David. "Some Thoughts About Retributivism". *Ethics* 101, nro 3 (abril 1991): pp. 537-559
- Duff, R.A. *Punishment, Communication and Community*. Oxford: Oxford University Press, 2001.
- Dworkin, Ronald, *Taking Rights Seriously*. Cambridge (Mass.): Harvard University Press, 1977.
- Dworkin, Ronald. *Freedom's Law: The Moral Reading of the Constitution*. Oxford: Oxford University Press, 1996.
- Dworkin, Ronald, *Justice for Hedgehogs*. Cambridge (Mass.): The Belknap Press of Harvard University Press, 2011.
- Dworkin, Ronald. *El imperio del derecho*. Traducido por Juan Iosa y Tomás Céspedes. Barcelona: Gedisa, 2022.
- Feinberg, Joel. "The Expressive Function of Punishment". *The Monist* 20, nro 43 (julio 1965): pp. 397-423.
- Ferrante, Marcelo. "¿Algo más sobre la justificación del castigo?". *Lecciones y Ensayos* Nro. 109, (2022): pp. 17-35.
- Ferrante, Marcelo. "Filosofía del Derecho penal". En Fabra Zamora, Jorge Luis y Spector, Ezequiel (eds.), *Enciclopedia de Filosofía y Teoría del Derecho*, México: Universidad Nacional Autónoma de México, 2015. pp. 2090–2091.

- Fletcher, George. *Rethinking Criminal Law*. Oxford: Oxford University Press, 2000.
- Fricker, Miranda. "What's the Point of Blame? A Paradigm Based Explanation". *Nous* 50, nro. 1 (marzo 2016): pp. 165-183.
- Hart, H.L.A. *Punishment and Responsibility. Essays in the Philosophy of Law* (2<sup>a</sup> ed.). Oxford: Oxford University Press, 2008.
- Morris, Herbert. "Persons and Punishment". *The Monist* 52, nro 4 (octubre 1968): pp. 475-501.
- Otto, Harro. *Manual de Derecho Penal* (7<sup>a</sup> ed.). Barcelona: Atelier, 2017.
- Pawlik, Michael. *El injusto del ciudadano*. Bogotá; Barcelona: Universidad Externado de Colombia; Atelier, 2023.
- Pereboom, Derk. "Free Will Skepticism and Criminal Punishment". En *The Future of Punishment*, editado por Thomas A. Nadelhoffer, Oxford: Oxford University Press, 2013, pp. 49-79.
- Pérez Barberá, Gabriel. "Problemas y perspectivas de las teorías expresivas de la pena: Una justificación deontológica de la pena como institución". *InDret* nro 4 (octubre de 2014).
- Rabossi, Eduardo. *La justificación moral del castigo*. Buenos Aires: Astrea, 1976.
- Rodríguez, Jorge L. *Teoría analítica del derecho*. Madrid: Marcial Pons, 2021.
- Sancinetti, Marcelo A. *Hacia un derecho penal más racional: contribuciones sobre castigo estatal e imputación* (1<sup>a</sup> ed.). Ciudad Autónoma de Buenos Aires: Editores del Sur, 2022.
- Shapiro, Scott. "The 'Hart-Dworkin Debate': A Short Guide for the Perplexed". *Public Law and Legal Theory Working Paper Series*, University of Michigan Law School, Working Paper nro. 77 (marzo 2007)
- Shapiro, Fred R. "The Most Cited Legal Scholars Revisited". *The University of Chicago Law Review* 88, nro. 1 (2021):
- Tan, Fangfang, y Erte Xiao. "Third-Party Punishment: Retribution or Deterrence?". *Journal of Economic Psychology*, nro. 67 (agosto 2018):
- Zaffaroni, Eugenio R., Alejandro Alagia, y Alejandro Slokar. *Manual de Derecho Penal: Parte General*. Buenos Aires: Ediar, 2009.
- Zaibert, Leo. *Punishment and Retribution*. New York: Ashgate, 2006.