

SOBRE LA JUSTIFICACIÓN MORAL DE LAS ACCIONES

El Tema del Castigo

EDUARDO A. RABOSI
Consejo Nacional de la Investigación
Científica y Técnica, Argentina

I

1. El tema del castigo ha preocupado a muchos filósofos clásicos con una persistencia que suele pasar inadvertida. Platón, Aristóteles, Tomás de Aquino, Hobbes, Spinoza, Rousseau, Kant, Schopenhauer, Hegel, Bentham y Mill son sólo algunos de los componentes de la larga nómina que se puede confeccionar.¹ Es verdad que las razones teóricas que han llevado a analizar la cuestión a pensadores de tan diversa extracción son muchas veces diferentes, como también lo son los contextos en los que sus análisis tienen lugar. Pese a ello es posible encontrar en sus discusiones una evidente unidad temática. *El tema del castigo* posee, pues, una importante tradición filosófica.

Los filósofos contemporáneos no han sido ajenos a esta preocupación. En lo que va del siglo han contribuido de manera importante a desarrollar y a superar los planteos tradicionales acerca del castigo. Y si bien la cuestión ha alcanzado en sus manos un notable grado de dificultad teórica —signo quizá de progreso filosófico— no cabe duda de que han logrado enriquecer de un modo notable el análisis del tema. Es este enfoque reciente del tema del castigo el que mueve nuestro interés. En él centramos nuestra atención, sin perjuicio de algunas referencias necesarias a los clásicos.

¹ En H. B. Acton (ed.) *The Philosophy of Punishment* (Londres, 1969) —una excelente compilación de trabajos sobre el tema del castigo— puede encontrarse una bibliografía muy completa, ordenada cronológicamente, que incluye referencias que abarcan desde Platón en adelante.

2. Pueden distinguirse dos momentos —sería excesivo hablar de etapas— en el tratamiento moderno del tema. El primero, que se extiende hasta los comienzos de la Segunda Guerra Mundial y comprende, entre otros, aportes de A. C. Ewing y W. D. Ross, culmina con el excelente trabajo de J. D. Mabbott, "Punishment".² El segundo momento se inicia luego de concluida la guerra y reconoce como uno de sus puntos de partida la discusión del trabajo de Mabbott. La década del cincuenta es particularmente fructífera. En 1953 dicho artículo es discutido por A. Flew y C. W. K. Mundle en el "Scots Philosophical Club". Sus trabajos se publican en 1954. Ese mismo año aparece un interesante trabajo de A. Quinton y en los años subsiguientes K. Baier, J. D. Mabbott, J. Rawls, S. I. Benn y H. L. A. Hart —entre otros— realizan importantes contribuciones, que junto con las anteriores condicionan en gran medida el curso futuro de la discusión.³ De ahí en más la bibliografía se multiplica de manera notable.⁴ Tan numerosos son los trabajos dedicados a la cuestión que no es exagerado afirmar que el *tema del castigo* es uno de los capítulos de la filosofía moral que han atraído con mayor persistencia el interés de muchos filósofos contemporáneos.

3. La expresión 'el tema del castigo' es muy general y vaga. Quizá el lector haya reparado ya en ambos caracteres. Pero es precisamente en ellos en donde reside su valor —sobre

² A. C. Ewing, *The Morality of Punishment* (Londres, 1929); W. D. Ross, *The Right and the Good* (Oxford, 1930), esp. capítulo 2, apéndice 2; J. D. Mabbott, "Punishment", *Mind*, 1939.

³ A. Flew, "The Justification of Punishment", *Philosophy*, 1954; C. W. K. Mundle, "Punishment and Desert", *Philosophical Quarterly*, 1954; A. Quinton, "On Punishment", *Analysis*, 1954; K. Baier, "Is Punishment Retributive?", *Analysis*, 1955; J. D. Mabbott, "Professor Flew on Punishment", *Philosophy*, 1955; J. Rawls, "Two Concepts of Rules", *Philosophical Review*, 1955; S. I. Benn, "An Approach to the Principles of Punishment", *Philosophy*, 1958; H. L. A. Hart, "Prolegomenon to the Principles of Punishment", *Proceedings of the Aristotelian Society*, 1959-60. (Incluido, junto con otros importantes trabajos de Hart sobre el tema, en H. L. A. Hart, *Punishment and Responsibility*, Oxford, 1968.)

⁴ En la ya citada obra de H. B. Acton la bibliografía incluye referencias a los más importantes trabajos publicados en los últimos años.

todo cuando se tratan de dar los primeros pasos en la discusión que nos preocupa. Por cierto que puede hablarse —tal como se ha hecho con frecuencia— del *problema del castigo* y, consecuentemente, de las *teorías* que se han propuesto como solución. Es más, esta “ecuación” *problema-teoría* representa el planteo efectivo de la mayoría de los autores clásicos y de algunos pensadores modernos. Pero es evidente que a esta altura de la discusión desarrollada en torno al tema del castigo, la citada “ecuación” resulta impropia, aún como mera terminología, pues lleva a crear desde el comienzo una “imagen” teórica inadecuada.

Lo que queremos señalar es lo siguiente. Uno de los resultados provechosos de la discusión reciente consiste en poner de manifiesto las limitaciones y la artificialidad de todo enfoque que al discutir el castigo pretenda estar lidiando con *un único* problema al que se debe dar solución proponiendo *un único* criterio (o quizá un conjunto de criterios alternativos). La tendencia actual consiste, más bien, en puntualizar los variados aspectos que presenta la institución del castigo y, en consecuencia, en tener presente la necesidad de apelar a criterios y respuestas variadas, cuando se trata de analizarlos. En otras palabras, si se quiere conservar la terminología tradicional conviene, en todo caso, hablar de *los* problemas que plantea el castigo, eliminando, por lo demás, el empleo de la dudosa expresión ‘teoría’. Al margen de no saber muy bien qué significa ‘teoría’ cuando se la emplea en contextos filosóficos, su uso en relación con la discusión en torno al castigo resulta singularmente equívoco y teóricamente malsano. Que esto es así, resultará evidente —esperamos— de la discusión que nos proponemos realizar.

4. Llamar la atención sobre los eventuales inconvenientes de cierta manera de encarar la discusión —es decir, la idea de que tenemos que habérsela, en tanto filósofos, con *un* problema al que de un modo u otro debemos dar *una* solución— no implica negar la posible existencia de una *moti-*

vación común que ha llevado y lleva a ver en la práctica del castigo una fuente de dificultades. No es fácil decir en unas pocas palabras en qué consiste esa motivación, pero no creemos descaminado describirla de la siguiente manera.

Matar intencionalmente a alguien, privarlo de su libertad, quitarle algo que le espreciado, hacerlo sufrir física o psíquicamente son acciones que merecen, en situaciones standard, *reprobación moral*. La pérdida de la vida, de la libertad, de bienes materiales y el sufrimiento son *males* y por lo tanto no puede ni debe admitirse que puedan ser sufridos por alguna persona como consecuencia de un acto consciente de alguna otra persona. Ahora bien, cuando describimos la acción de una persona —digamos, A— sobre otra persona —digamos, B— como un castigo, parte de lo que queremos significar es que A inflige a B alguno de esos “males”. Es decir, que A ha ordenado privar de la vida a B, o que le ha privado de su libertad, o de bienes materiales, o que lo ha hecho sufrir —dándole una paliza por ejemplo. Pero si esto es así, ¿cómo justificar *moralmente* castigar a alguien? ¿Por qué un acto que parece ser el mismo que otro acto —digamos, propinar a alguien una paliza— puede y debe merecer en un caso reprobación moral y en otro caso —en tanto describimos dicho acto como un castigo— puede merecer quizá asentimiento y aprobación o, al menos, abstención de juicio moral?

En este punto puede sugerirse una respuesta simple y *prima facie* admisible: “La justificación se encuentra, sencillamente, en que cuando A castiga a B, B ha hecho algo que *no debía* hacer; ha cometido una *ofensa*. En esto reside la diferencia entre el caso en que, por ejemplo, A da una paliza a B “por gusto” y aquel en que A castiga a B dándole una paliza. Y la diferencia es, obviamente, crucial. Aún en una de las aplicaciones más terribles de la *lex talionis* —privar de la vista a quien ha privado de la vista a otro— ambos actos (privar de la vista alguien) no pueden ser considerados como *el mismo acto*, en el sentido de ser ambos igualmente malos

o de ser productivos de males que acaecen igualmente a dos individuos distintos. En el primer caso el acto consiste en el castigo a un ofensor, en el segundo, el acto consiste en una ofensa. Esto es lo que hace a la diferencia entre ellos. Y es en esta relación castigo-ofensa en donde se diluye la necesidad de buscar alguna justificación, moral o no.”

Pero esta respuesta —que contiene en germen una de las posiciones tradicionales sostenidas en relación con el castigo— no es admisible. En todo caso mueve a plantear la cuestión en términos algo distintos. Si B ha llevado a cabo una ofensa —es decir, si ha hecho algo malo— y si castigar a B implica infligirle un “mal”, parece irrazonable provocar un mal simplemente porque otro mal ha tenido efecto. La dificultad sigue en pie. O si se prefiere plantearla en términos más dramáticos: ¿por qué es moralmente preferible un mundo en el que se dé el complejo *ofensa-castigo* a un mundo en el que sólo se dé la ofensa? Intuitivamente nos inclinamos a preferir el primer tipo de mundo, pero ¿por qué? ¿Por qué pensamos que es *moralmente* mejor? En otras palabras, ¿qué justificación *moral* podemos dar al castigo?

5. Estas preguntas acerca del castigo expresan en parte la motivación principal que ha llevado a los filósofos a meditar sobre la cuestión. No pretendemos sugerir, por cierto, que sean las únicas ni que hayan sido propuestas necesariamente de esta manera. Lo que sí creemos es que de uno u otro modo constituyen el meollo de la problemática que ha servido de base a la prolongada discusión sobre el tema. Además, la insistencia en buscar una *justificación moral* al castigo, como tema filosófico central, posee un interés que excede su valor teórico intrínseco. En realidad, el problema de la justificación moral de las acciones (en general) ocupa un lugar de preeminencia dentro de la filosofía moral contemporánea y, como tendremos oportunidad de ver más adelante, gran parte de la reciente discusión sobre el tema del castigo puede considerarse como un sucesivo refinamiento de la cuestión rela-

tiva a su justificación moral de las acciones (en general). Este es un punto importante que merece ser tenido en cuenta; aunque dada la índole de este trabajo sólo nos limitaremos a señalarlo aquí, sin pretender internarnos en el marco más general que presupone.

6. Estas consideraciones preliminares nos ponen en condiciones de indicar el plan que nos proponemos desarrollar. En primer lugar (sección II) expondremos las dos concepciones tradicionales acerca del castigo —la *retribucionista* y la *utilitarista*— tratando de mostrar sus presupuestos, contenido y limitaciones. La opción *retribucionismo o utilitarismo* parece llevar a un punto muerto teórico que debe ser superado. En consecuencia, analizaremos luego (sección III) tres enfoques alterantivos que pretenden superar tal opción. Este análisis nos brindará elementos suficientes para encarar, por último (sección IV), la consideración de una interesante familia de temas que deben discernirse cuidadosamente en un enfoque realmente fructífero del tema del castigo.

Como puede apercibirse, nuestro trabajo contiene además de una parte expositiva, el desarrollo de cierto aparato crítico. Y si bien no creemos que nada espectacular se siga de él, mostrará —esperamos— cierta virtud en la presentación del tema y en algunas de las sugerencias finales.

II

1. Comencemos con la exposición —‘reconstrucción’ es posiblemente el término más adecuado— de las dos concepciones o “teorías” tradicionales acerca del castigo: la *retribucionista* y la *utilitarista*. Ambas concepciones pueden interpretarse, en principio, como intentos de dar respuesta a un mismo problema. Y hay dos cosas curiosas acerca de ellas. Primera, que dichas concepciones o “teorías” parecen agotar las posibles respuestas a la cuestión que pretenden resolver; segunda, que las dos se adecuan al sentido común: ninguna de ellas está rodeada del aire esotérico que tan comúnmente caracte-

riza las respuestas de los filósofos; son, en cambio, respuestas sensatas que cualquier persona inteligente podría dar o aceptar a poco que se pusiera a meditar sobre el problema.

2. Según los *retribucionistas*, el castigo que se inflige a un individuo se encuentra *moralmente* justificado por el hecho de que dicho individuo sea *culpable* de haber cometido una *ofensa*. En otras palabras, que B —por ejemplo— haya cometido intencionalmente una ofensa constituye la razón para que se le administre el castigo que le corresponde, pues es *moralmente correcto* que el ofensor sufra por la ofensa que ha cometido. “El castigo es castigo —afirma concluyentemente Bradley— sólo cuando es merecido. Pagamos la ofensa porque la debemos y por ninguna otra razón.”⁵

Esta es, en términos generales, la tesis básica del retribucionismo. Aunque resulta difícil ampliarla o precisarla de manera satisfactoria pues aparece expresada e interpretada de maneras muy distintas. La siguiente miscelánea de opiniones ayudará a ejemplificar este punto.

a) “El castigo judicial —dice Kant— ...nunca puede ser infligido como un medio de lograr algún bien para el criminal o para la sociedad, sino *únicamente* porque (el criminal) ha cometido un crimen... un ser humano nunca puede ser usado como un medio para los fines de otras personas.”⁶

b) “Al discutir este problema —sostiene Hegel— las únicas cosas importantes son: primero, que el crimen debe ser anulado no porque consista en la producción de un mal, sino porque es una infracción al derecho como derecho; segundo, el problema lo plantea esa exigencia positiva que posee el crimen y que debe ser anulada; es esta existencia lo que es realmente malo, lo que debe eliminarse, y el punto crucial es determinar en dónde reside (en qué consiste).” Hegel termina el párrafo con una frase digna de ser puesta en la-

⁵ F. H. Bradley, *Ethical Studies* (Oxford, 1927, 2da. ed.), págs. 26-7.

⁶ E. Kant, *Metaphysische Anfangsgründe der Rechtslehre*, pág. 49.

bios de G. E. Moore: “En tanto no se aprehendan con claridad los conceptos en cuestión, la confusión seguirá reinando en la teoría del castigo.”⁷

c) “La única justificación para castigar a un hombre —afirma Mabbott— es que haya violado la ley.”⁸

d) “El hecho de que una persona haya cometido una ofensa moral —dice Mundle— es una razón suficiente para que se la haga sufrir; entendiéndose por tal a una razón que es independiente del valor de los resultados previsibles del acto particular de castigo o de la imposición de las leyes en cuestión.”⁹

No nos detendremos a considerar las diferencias, en parte obvias, que existen entre estas versiones de la concepción retribucionista. Nos interesa señalar en cambio una importante consecuencia que parece seguirse de todas ellas. Si se considera que el castigo queda justificado *moralmente* por la culpa del ofensor, resulta claro que el monto del castigo que se le proporciona debe adecuarse con exactitud a la magnitud del agravio que ha cometido. La severidad del castigo debe corresponder exactamente a la gravedad de la ofensa. De no ser así, parecería que el equilibrio que logra reestablecer la producción del castigo —según la tesis retribucionista— quedaría de alguna manera alterado. Y esto plantea de modo inmediato el problema de ofrecer criterios de adecuación entre la gravedad de la ofensa y la gravedad de la pena. Las dificultades que surgen en conexión con este problema son grandes y motivan, como veremos, fundadas críticas de los utilitaristas.

3. Para los *utilitaristas* el castigo que se inflige a un indivi-

⁷ G. Hegel, *The Philosophy of Right* (tr. T. M. Knox, Oxford, 1942), págs. 68-70.

⁸ J. D. Mabbott, “Punishment”, pág. 45 (A). (El signo ‘(A)’ indica que el número de página es el correspondiente a la compilación de artículos efectuada por H. B. Acton.)

⁹ C. W. K. Mundle resume de esta manera la concepción retribucionista que defendió en “Punishment and Desert” y que afirma haber abandonado. Cf. en H. B. Acton (Postscript, 1968), págs. 81-2 (A).

duo no puede justificarse *moralmente* basándose en la culpa que supone el haber cometido una ofensa. “Todo castigo —sostiene Bentham— es algo dañino; todo castigo es en sí mismo malo.”¹⁰ Si este es el caso, ¿cómo justificar *moralmente* algo que es malo en sí mismo relacionándolo con un acto pasado, también malo e indeseable? Los retribucionistas están errados. Si se quiere ofrecer una justificación *moral* adecuada del castigo no debe mirarse al pasado sino al futuro: el castigo sólo puede justificarse *moralmente* cuando se toman en cuenta las *consecuencias valiosas* que su aplicación puede llegar a producir. Si no se consideran tales consecuencias no se justifica *moralmente* nada.

Cuáles son esas consecuencias valiosas puede depender, en parte, de las circunstancias de cada caso. Aunque moralmente se menciona a alguna o algunas consecuencias que se considera que representan paradigmáticamente el fin valioso que debe perseguir la aplicación del castigo. El castigo puede perseguir la *reforma* del ofensor, o puede tratar de *desalentarlo* o de *disuadirlo* de realizar en el futuro ofensas similares (o *desalentar* o *disuadir* con el ejemplo a otros miembros del grupo social a que realicen ofensas parecidas). El castigo también puede justificarse como un medio para *prevenir* que el ofensor pueda repetir en el futuro la ofensa cometida.

También puede plantearse al utilitarista el problema de la adecuación entre la magnitud de la ofensa y la severidad del castigo. Pero resulta claro que su posición no lo obliga a ofrecer criterios exactos de adecuación. En este punto pareciera encontrarse en mejor posición que el retribucionista. El utilitarista puede responder que la severidad del castigo debe ser tal que no resulte conveniente al ofensor cometer la ofensa y recibir el castigo, o bien que la severidad de la ofensa debe ser tal que asegure el logro del fin valioso que persigue, etc.

¹⁰ J. Bentham, *Introduction to the Principles of Morals and Legislation*, capítulo 13.

4. Pasemos revista a algunas de las *objeciones* a que están sujetas ambas concepciones o “teorías” clásicas del castigo.

La principal crítica que los retribucionistas formulan contra los utilitaristas es que de su concepción del castigo se sigue que en determinadas circunstancias resultaría justificado castigar a personas inocentes. El argumento es el siguiente. Si el castigo se considera justificado *moralmente* por las consecuencias valiosas que su aplicación puede producir, es posible —y hasta deseable— que en ciertos casos se llegue a castigar a una persona inocente si se tiene la certidumbre de que tales consecuencias se obtendrán. Y esto es, sin duda, algo moralmente inadmisibile.

Veamos algunos ejemplos.

Supongamos que B no ha cometido cierta ofensa —es decir, que es inocente— pero que C y D creen que B la ha cometido. Supongamos, además, que tenemos motivos para pensar que C y D tienen propensión a cometer la ofensa en cuestión, o que están dispuestos a cometerla. Si pensamos, por ejemplo, que la justificación moral del castigo se funda en la *prevención* de ofensas o en *desalentar* o *disuadir* a quienes estén dispuestos a cometerlas parece seguirse que en la hipótesis se debe castigar a B —pese a su inocencia y sabiéndolo inocente— como un medio útil para alcanzar los fines valiosos que justifican el castigo; es decir, prevenir en este caso que C y D pueden llegar a cometer la ofensa o disuadirlos a que no la cometan.

Supongamos ahora que B es un individuo de mal comportamiento que no ha cometido cierta ofensa —nuevamente, supongamos que B es inocente. Admitamos que tenemos motivos fundados para pensar que B podría cometer en el futuro dicha ofensa. Si justificamos el castigo de un individuo por las consecuencias valiosas implícitas en su *reforma* parece seguirse que también en este caso debemos castigar a B —pese a su inocencia y sabiéndolo inocente— en tanto alcancemos de tal manera el fin valioso deseado.

Un último caso. Consideremos nuevamente la concepción

utilitarista en su versión preventiva. “Si algún tipo de crimen sumamente cruel —observa Carritt— ha proliferado en exceso y ninguno de los criminales ha podido ser descubierto, podría resultar sumamente eficaz como medio ejemplificador colgar a un hombre inocente, si la acusación en su contra pudiera ser organizada de tal manera que todos lo creyeran culpable.”¹¹ De tal manera se lograría, quizá, prevenir la realización de nuevos crímenes.

Obsérvese que en ningún momento el argumento atribuye a la concepción utilitarista sostener la tesis de que *debe* castigarse a personas inocentes. A lo que la crítica apunta es al hecho de que si el castigo se considera *moralmente* justificado por las consecuencias valiosas que permite o puede permitir alcanzar, existen o pueden existir situaciones en las que sea deseable alcanzar tales consecuencias *con prescindencia* de la culpa de la persona castigada. Y nadie —se supone— está dispuesto a admitir que se castigue a un inocente, cualquiera que sea el valor de las consecuencias que se puedan seguir de tal hecho. El castigo de personas inocentes es algo moralmente inadmisibile.

Los retribucionistas también esgrimen una segunda crítica poderosa. En un sentido estricto, sostienen no es el castigo en sí mismo el que previene, reforma o desalienta. Estos fines se alcanzan muchas veces mediante la *amenaza* del castigo, o la *publicidad* que se da a la aplicación del castigo, o los *consejos*, el *tratamiento psiquiátrico*, la *educación*, etc., que se da a los ofensores. Pero es evidente que todos estos medios son “externos” al castigo mismo. Como ha señalado con agudeza Mabbott refiriéndose a la publicidad que normalmente se da a los juicios penales y a las sentencias que los jueces dictan en ellos: “. . . aunque a veces se habla de la publicidad como si fuera ‘parte del castigo de un hombre’ no es parte de su castigo; como no lo es su arresto o detención. . . aunque ambos pueden resultarles desagradables y acarrearle mala reputación. El juez condena a un

¹¹ E. F. Carritt, *Ethical and Political Thinking* (Oxford, 1947), pág. 65.

hombre a tres años de prisión y no a tres años *más* tres columnas en los diarios. . . La visita del capellán de la prisión no forma parte del castigo de un hombre, como tampoco la visita de (un asistente social) o del Ratón Mickey.”¹²

En síntesis, la pretensión de justificar *moralmente* al castigo apelando a las consecuencias valiosas que puede llegar a producir su aplicación resulta —según los retribucionistas— errada. Y lo es, entre otros argumentos, porque: a) puede llegar a justificar en ciertos casos el castigo de inocentes, y b) no distingue entre el castigo propiamente dicho y un cúmulo de factores “externos” a él.

5. Los utilitaristas también tienen fuertes objeciones que formular a la concepción retribucionista.

En primer lugar, en ésta concepción parece estar implícita una especie de injustificada valoración del castigo por el castigo mismo. Esto no es nada más —se alega— que la manifestación refinada o la racionalización de quien ve en el castigo una exteriorización del deseo de *venganza* que naturalmente experimenta el ofendido ante el ofensor.¹³

Consideremos un caso extremo. Supongamos que en una escuela el director ha dispuesto que los alumnos de los años superiores —que ocupan aulas ubicadas en el primer piso— no deben descender ni intentar descender a la planta baja

¹² J. D. Mabbott, *Op. cit.*, pág. 40 (A).

¹³ Debe distinguirse entre la concepción retribucionista del castigo y la posible justificación del castigo en base al mero deseo de venganza. Algunos autores hablan de una concepción del castigo fundada en el deseo de venganza, que colocan en un pie de igualdad con las dos concepciones que venimos analizando. (Cf. J. Feinberg (ed.), *Reason and Responsibility*, California, 1965, págs. 296-8). Se menciona también a O. W. Holmes, Jr., como un defensor de tal concepción. (Ver su libro, *The Common Law*, Boston, 1881, págs. 40-51). Consideramos que dicha nivelación teórica es inaceptable. Si se trata de justificar *moralmente* al castigo no se ve cómo puede apelarse para ello a una reacción o a un sentimiento *instintivo* que puede acompañar al conocimiento de la ofensa por parte del ofendido, para producir sobre tal base la elucidación que se pretende. Sea como fuere, lo importante es insistir en que el retribucionismo no implica una justificación del castigo basada en el mero deseo de venganza. Este punto no siempre se ve con claridad. Expositores cuidadosos como J. Hospers parecen incurrir en ciertas confusiones al respecto. (Cf. *Human Conduct*, Londres, 1963, págs. 453-4).

durante los recreos. El castigo a quien viole esta regla consiste en la privación de los recreos durante un mes. B, un alumno travieso de último año, viola abiertamente lo dispuesto por el director y al intentar descender por la escalera, cae fracturándose las piernas. B vuelve al colegio luego de algún tiempo, enyesado e imposibilitado de caminar. Según el dictamen médico, no puede volver a hacerlo hasta dentro de dos meses: un mes después de que hayan terminado las clases. Está claro que en la hipótesis B no puede volver a repetir la ofensa y que la privación de los recreos puede afectar su estado de ánimo. Como es natural, sus compañeros experimentan hacia él una natural simpatía; un propósito de ayudarlo a salir del mal trance. El director lo castiga de conformidad con lo estipulado con la norma: lo priva de los recreos durante un mes de clases. Y al ser criticado por la decisión que ha tomado responde que el castigo aplicado a B está plenamente justificado por ser B culpable: ha intentado descender a la planta baja del colegio durante un recreo. Es posible que el director agregue que no sólo el castigo se encuentra justificado sino que él, en tanto autoridad, ha tenido el *deber* de aplicarlo. El planteo del director es de un neto corte retribucionista.

Pero, ¿es este planteo admisible? Dados los *hechos relevantes*, ¿no muestra la respuesta del director la peligrosa estrictez que caracteriza a la concepción retribucionista? ¿Es que sólo la culpa puede ser lo único *moralmente relevante* cuando se trata de castigar a alguien? ¿No proporciona este tipo de ejemplos un argumento parecido —aunque no, por cierto, similar— al que los retribucionistas esgrimen contra los utilitaristas? Si bien no parece que pueda llegar a justificarse *moralmente* el castigo de inocentes, ¿es siempre *moralmente* justificable el castigo de culpables?

Por cierto que el director del colegio podría ofrecer una respuesta distinta a la anterior. Podría decir que el castigo de B está justificado porque B cometió la ofensa y porque la aplicación del castigo resulta necesaria en este caso para

mantener el “principio de autoridad”. Pero esta respuesta abriría al utilitarista un camino fácil para la réplica, porque justificar el castigo de B en base al “principio de autoridad” es colocar la culpa de B en un plano secundario. Lo importante ahora serían las consecuencias que trae aparejado el castigo de B, es decir, evitar que se resienta la disciplina, impedir que otros alumnos crean que “siempre puede haber excepciones”, etc. En otras palabras, sería apelar a una justificación de tipo utilitarista.

Otro aspecto de la concepción retribucionista sujeto a críticas es el relativo a la adecuación entre el castigo y la ofensa. Como ya hemos señalado, si el castigo de una persona se justifica *moralmente* por su culpa parece seguirse que el monto de castigo que se le administra debe corresponder al monto de la gravedad de la ofensa que ha realizado. Si por ejemplo, el castigo que se inflige a B equivale a diez unidades de disvalor y la ofensa equivale a cinco unidades o quizá a veinte, el castigo sería *injusto*; no alcanzaría a restablecer la “balanza moral” que contrapesa castigos con ofensas. Pero, ¿cómo fijar tal relación de equilibrio? ¿Qué factores tener en cuenta? ¿Qué criterios seguir? El argumento utilitarista puede plantearse en dos etapas: 1) o bien es imposible —como parece serlo— determinar con *exactitud* dicha relación, o 2) si es posible hacerlo es porque entran en juego, explícita o implícitamente, consideraciones de tipo utilitarista, es decir, consideraciones que toman en cuenta las consecuencias valiosas que se siguen de la aplicación del castigo.

En suma, los utilitaristas argumentan —entre otras cosas— que el retribucionismo ofrece una respuesta estéril y artificial cuando intenta justificar *moralmente* el castigo en base a la culpa del ofensor. Estos caracteres se ponen en evidencia cuando se trata de aplicar la tesis a casos concretos. En ellos, parece inevitable tomar en cuenta, tarde o temprano, las consecuencias del castigo; es decir, aceptar básicamente la relevancia de una interpretación de tipo utilitarista.

III

1. Estas son, en términos muy generales, las tesis básicas que caracterizan a las dos concepciones o “teorías” clásicas del castigo y las críticas canónicas a que están sujetas. ¿Cuáles son las ventajas teóricas que pueden llevar a preferir a una de estas concepciones en particular? La pregunta parece inevitable. Pero es en este punto cuando comienza a tomar cuerpo la sensación de haber arribado a una especie de punto muerto. Por un lado, no existen —en este nivel del análisis— criterios adecuados que permitan ofrecer razones concluyentes en favor de una u otra concepción. Esto parece claro. Por otro lado (y este punto está implícito, en realidad, en lo anterior), no se ve muy bien cuáles son *exactamente* los términos del problema a resolver. Hemos hablado hasta aquí de ‘justificar *moralmente* el castigo’, pero la expresión, así usada, cubre demasiado campo como para admitir que sirva a los fines de delimitar con precisión el marco específico de la discusión.

Una manera de superar la *impasse* es proponerse responder varias preguntas que surgen naturalmente: (a) ¿se plantean los retribucionistas y los utilitaristas el mismo *tipo de pregunta* respecto del castigo?; (b) ¿dan solución al mismo *tipo de problema* —en el sentido de que sus respectivas tesis pueden ser ubicadas dentro del mismo compartimento filosófico?; (c) ¿toman en cuenta ambas concepciones el mismo *tipo de fenómeno* o proceso?

En lo que sigue prestaremos atención a cada una de estas tres opciones.

2. Comencemos con la primera. ¿Se plantean los retribucionistas y los utilitaristas el mismo *tipo de pregunta* respecto del castigo?

Una respuesta posible es que no; que las diferencias de fondo entre ambas concepciones se deben al hecho —trivial quizá— de que formulan ante un mismo problema un tipo distinto de pregunta. Dado un caso concreto de castigo los

retribucionistas parecen preguntar '¿Por qué se castiga a B?'. Los utilitaristas, en cambio, parecen preguntar '¿Para qué se castiga a B?'. Y es natural que si se formulan preguntas de distinto tipo, se obtengan respuestas diferentes. La diferencia de fondo entre ambas preguntas es, por lo demás, obvia. Si preguntamos a alguien '¿Para qué hace *p*?' (en donde '*p*' es una frase que hace referencia a una acción cualquiera), la respuesta *tiene* que hacer mención —entre otras cosas— a un hecho futuro relacionado con *p*, en términos que incluyan la referencia a ciertas consecuencias que el agente persigue al llevar a cabo *p*. A su vez, si preguntamos a alguien '¿Por qué hace *p*?' —en un sentido relevante de '¿Por qué...?'— la respuesta *tiene* que hacer referencia a un hecho o a una acción pasados, conectados de alguna manera con *p*. La diferencia es clara y sus consecuencias teóricas son evidentes.

Si aceptamos *prima facie* este planteo, es posible formular varias observaciones de interés. Entre otras cosas, parecería seguirse de él (a) que no existiría entre la concepción retribucionista y la utilitarista una oposición teórica real; (b) que dado el grado de aceptabilidad que poseen ambas concepciones, podría admitirse en principio que sus respectivas respuestas son correctas simultáneamente; (c) que el error básico de los sostenedores del retribucionismo y del utilitarismo consiste en creer que *su* concepción ofrece la justificación *moral* del castigo, y (d) que quizá fuera posible llegar a integrar a ambas concepciones en lo que podría considerarse una concepción *general* de la justificación *moral* del castigo. No puede negarse el interés que poseen estas consecuencias, pero lamentablemente el planteo del cual se desprenden no es aceptable. Y ello es así por varias razones.

Primera y principal, porque no hay motivos para pensar que tanto a los retribucionistas como a los utilitaristas se les haya podido pasar por alto un hecho tan trivial como el que pone de manifiesto la distinción entre los dos tipos de pre-

gunta. Ello es así por un importante motivo que no siempre se suele tener en cuenta. La diferencia entre la respuesta retribucionista y la utilitarista frente a la justificación *moral* del castigo se debe en gran medida a la adopción de una posición teórica diferente frente a lo que podríamos denominar 'el problema de justificar *moralmente* una acción'. Es seguro que si tomamos a un defensor ortodoxo de cada posición llegaremos tarde o temprano a descubrir en este nivel último una diferencia crucial. Para un utilitarista, el castigo como cualquiera otra acción humana con dimensión moral sólo puede justificarse por sus consecuencias. Para un retribucionista, tal justificación de las acciones humanas con dimensión moral sólo puede hacerse mediante una referencia a hechos pasados que incluyan la intención del agente, o determinadas características volitivas, o cosas parecidas. No se trata, pues, del hecho simple de que al formular preguntas distintas arriben, como es de suponer, a respuestas diferentes, sino que los retribucionistas y los utilitaristas formulan las preguntas que formulan porque consideran que de un modo u otro solamente es *ese* el tipo de pregunta que corresponde hacer ante este tipo de problema.

En segundo lugar, la nítida diferencia que hemos trazado entre '¿Por qué se castiga a B?' y '¿Para qué se castiga B?' está sujeta a una importante excepción. En general, '¿Por qué hace p?' puede responderse con sentido tanto haciendo referencia a un acto pasado ('Porque ocurrió m') como a un acto futuro ('Porque así ocurrirá n'). '¿Por qué hace p?' es una pregunta ambigua: en uno de sus sentidos equivale a '¿Para qué hace p?'. Y si bien esta circunstancia no prueba que no exista la diferencia conceptual que hemos tratado de poner de manifiesto al hacer referencia al funcionamiento de '¿Por qué...?' y '¿Para qué...?', pone en tela de juicio cuáles pueden ser los criterios a emplear para llegar a determinar el sentido relevante de '¿Por qué...?' que interesa poner en claro para salvar la distinción. La discu-

sión de tales criterios se volvería a plantear de reabrir válidamente la disputa entre retribucionistas y utilitaristas.

Por último, ¿qué se lograría llevando a la práctica el plan de “integrar” ambas respuestas? Dadas las observaciones precedentes difícilmente podría lograrse tal objetivo. Y aunque se la lograra, ¿qué otra cosa se obtendría como no fuera una respuesta ecléctica expuesta, en tanto tal, a las críticas de retribucionistas y utilitaristas? En la bibliografía sobre el tema existe algún ejemplo notable de este resultado indeseable.¹⁴

¹⁴ Quizá el esfuerzo más valioso realizado en esta dirección se deba a A. C. Ewing (*The Morality of Punishment*). Ewing critica las teorías tradicionales del castigo —es decir, el retribucionismo y el utilitarismo (en su versión reformadora y preventiva)— y basa su posición en la *función educativa* del castigo. El castigo es educativo en tanto y en cuanto hace que la gente se abstenga de realizar malas acciones porque las considera tales —no porque meramente teme al castigo que puede recibir por llevarlas a cabo. En realidad, “el objeto moral del castigo como tal es hacer que la gente piense que determinado tipo de acción es muy malo”. (pág. 104). De ahí, por ejemplo, la importante función educativa que cabe atribuir a la ley penal. Ewing piensa que esta función de castigo debe ser complementada con la que le atribuyen las concepciones tradicionales —aunque con expresas reservas. En consecuencia, modifica de la siguiente manera al principio de retribución, es decir, la tesis de que es moralmente correcto castigar al culpable de una ofensa: (i) es moralmente correcto desaprobador una mala acción; (ii) infligir sufrimiento es un método adecuado para expresar nuestra desaprobación ante una mala acción. Frente a la eventual pregunta, ¿por qué no infligir castigo a una persona inocente, si es que el castigo es educativo? Ewing responde: “...porque el castigo sólo tiene valor educativo para el culpable.” (pág. 91).

Como se ve, Ewing intenta desarrollar una posición básicamente utilitarista aunque con fuertes tintes retribucionistas. Y no hay duda de que su análisis ofrece algunas sugerencias valiosas. Pero, ¿qué opinión le puede merecer a un retribucionista o a un utilitarista más o menos ortodoxo? Un buen ejemplo de cómo puede reaccionar un retribucionista lo ofrece C. W. K. Mundle (“Punishment and Desert”). Mundle observa: “...desde un punto de vista filosófico, la teorías reformadora, educativa y preventiva son meras variantes del tema utilitarista... La controversia básica se da entre quienes sostienen que el castigo debe ser justificado por el valor de sus efectos y los defensores de la teoría retributiva que niegan tal cosa. Ewing intenta llegar a un acuerdo diciendo que el castigo está parcialmente justificado porque expresa una actitud correcta de desaprobación y porque, en gran medida, tiene buenas consecuencias... (Pero lo que está en discusión es si el hecho de que una persona haya cometido una ofensa moral es una razón suficiente para infligirle un sufrimiento. Aunque Ewing no plantea la cuestión en esta forma, parece claro —por lo que dice— que se inclinaría por responder negativamente.” (*Op. cit.*, pág. 69 (A).) E de suponer que a su vez, un utilitarista podría criticar a Ewing por hacer un uso innecesario del principio de retribución. Los elementos utilitaristas implícitos en su teoría justificarían tal observación.

3. Demos pues por fracasado este primer intento de superar el antagonismo entre las dos "teorías" clásicas del castigo y pasemos a considerar la segunda de las preguntas: ¿dan solución los retribucionistas y los utilitaristas al mismo *tipo de problema* —en el sentido de que sus respectivas tesis pueden ubicarse dentro del mismo compartimento filosófico?

Una respuesta negativa a esta pregunta exige formular un planteo mucho más elaborado que el anterior, aunque también mucho más interesante. La idea —en términos generales— es la siguiente: la crítica más fuerte que los retribucionistas esgrimen contra los utilitaristas es —como hemos señalado— que su concepción del castigo puede llevar a justificar el castigo de personas inocentes. Esta consecuencia es inaceptable. En esto pareciera haber consenso general. Pero, ¿por qué asentimos tan fácilmente a la afirmación de que es moralmente malo castigar a inocentes? Pues porque entre otras cosas parece *necesario* que para que podamos decir con sentido que A castiga a B por haber hecho *m*, B *tiene* que haber cometido la ofensa que trae aparejado el castigo que recibe; B tiene que ser culpable de haber hecho *m*.

¿Y qué tipo de necesidad es ésta? ¿Lógica? ¿Moral (si es que existe tal cosa)?

Una respuesta sumamente plausible es que se trata de una necesidad de carácter *lógico* —en el siguiente sentido. No es —como piensan los retribucionistas y quizá la mayoría de la gente— que no *debamos* (moralmente) castigar a los inocentes; lo que pasa es que no *podemos* (lógicamente) hacerlo. En términos de Quinton: "... el infligir un sufrimiento a una persona solamente puede ser descrito de una manera correcta como castigo si esa persona es culpable... En consecuencia, la tesis retribucionista no es una doctrina moral sino una doctrina que da cuenta del significado de la palabra 'castigo'... El castigo *no puede ser in-*

Creemos que lo expuesto ejemplifica adecuadamente la objeción que formulamos en el texto.

fligido al inocente. Tampoco *debe* infligírsele el sufrimiento asociado al castigo. En primer lugar, porque es algo brutal. Y además, porque si se lo presenta como un castigo supone una mentira (un fraude)."¹⁵

De lo anterior se sigue que expresiones del tipo 'Te castigo pese a que eres inocente' o 'Te castigo por algo que no has hecho' son lógicamente absurdas. Si *ser culpable* es una condición necesaria de *ser castigado* —y esta parece ser la tesis mínima que defienden los retribucionistas— resulta claro que si no se es culpable no puede haber castigo. Decir, pues, con pretensión de seriedad, 'Te castigo pese a que eres inocente' resulta tan absurdo como decir 'Te acuso de haber hecho *m*, aunque no hayas hecho *m*' —o quizá en otro orden de expresiones lógicamente absurdas 'Sé que la Tierra es redonda pero no es verdad que la Tierra sea redonda'.

Muchas formulaciones de la concepción utilitarista encajan dentro de este planteo crítico. Recordemos a modo de ejemplo la frase de Bradley ya citada. Dice Bradley: "El castigo sólo es castigo cuando es merecido. Pagamos la pena porque la debemos y no por otra razón; y si el castigo es infligido por cualquier otra razón que no sea la de que se lo merece cuando se es culpable, entonces es una inmoralidad mayúscula, una injusticia evidente, un crimen abominable y nunca lo que pretende ser." Pero, ¿qué función lógica tienen las frases 'El castigo sólo es castigo cuando es merecido', o 'Pagamos la pena porque la debemos y no por otra razón', o 'El castigo... se lo merece cuando se es culpable'? ¿Explicitan acaso el significado de la palabra 'castigo'? ¿Expresan condiciones de uso de expresiones del tipo 'A castiga a B'? Lo cierto es que cualquiera que sea la respuesta que se dé, las dramáticas afirmaciones finales de Bradley resultan irrelevantes. Como ha señalado Flew: "... por mayúscula que sea la inmoralidad, evidente la injusticia y abominable el crimen, *no puede* haber castigo si

¹⁵ A. Quinton, *Op. cit.*, pág. 59 (A).

no es a la vez merecido y se lo paga porque se lo debe.”¹⁶ En otras palabras, lo que viene a afirmar Bradley es algo evidente: ‘El castigo sólo es castigo cuando es merecido’, ‘el castigo se lo merece cuando se es culpable’; pero tal evidencia deriva nada más —ni nada menos— del hecho de que en esas frases se explicitan en realidad el significado de la palabra ‘castigo’ o las condiciones de uso de expresiones del tipo ‘A castiga a B’.

Las consecuencias que se siguen de este intento de superar el antagonismo entre la concepción utilitarista y la retribucionista son claras.

En primer lugar, se establece entre ellas una diferencia en cuanto al nivel disciplinario en el que operan. El retribucionismo ofrece una respuesta de carácter *lógico*: elucida el significado de la palabra ‘castigo’ o las condiciones de aplicación de expresiones del tipo ‘A castiga a B’. El utilitarismo ofrece, a su vez, una respuesta de carácter *ético*: intenta justificar moralmente al castigo. No hay pues oposición entre ambas concepciones; cada una ofrece su solución dentro de un área disciplinaria determinada.

En segundo lugar, el planteo propuesto lleva a distinguir —y este es, sin duda, su mérito teórico mayor— dos aspectos peligrosamente encapsulados en la frase ‘la justificación moral del castigo’: un aspecto *conceptual* y un aspecto *ético*. Ambos aspectos son obviamente diferentes y deben ser separados con cuidado. Una cosa es la explicitación del significado o de las condiciones de uso de ‘castigo’, ‘castigar’, etc., y otra muy distinta es el problema ético de *justificar moralmente* el castigo o la acción de castigar. Esta distinción es teóricamente importante en cuanto a la discusión del problema.

En tercer lugar, resulta del planteo que ciertas frases que pueden llegar a usarse en contextos usuales resultan lógicamente absurdas. ‘Te castigo pese a que eres inocente’ y ‘Te castigo por algo que no has hecho’ son dos ejemplos de tal

¹⁶ A. Flew, *Op. cit.*, pág. 92-3 (A).

tipo de frases. El por qué de su absurdo lógico surge de las razones dadas más arriba.

Por último —y este es el resultado que motiva y da fuerza a la propuesta que estamos analizando— la conocida crítica al utilitarismo que consiste en achacarle la posibilidad de poder llegar a justificar el castigo de inocentes pierde toda base de sustentación. No puede haber castigo de inocentes. Castigar a un inocente es una imposibilidad lógica. Puede hacerse sufrir a alguien bajo el rótulo de castigo, pero ello supone mera maldad, o fraude, o mentira —como sugiere Quinton.

El planteo es, como se ve, sumamente atractivo, aunque no exento de dificultades. Quizá las objeciones más terminantes que pueden formularse en su contra tienen que ver con la tercera de las consecuencias que acabamos de enunciar.

Por de pronto cabe señalar una *cuestión de hecho* innegable. Existen, o pueden existir, seres suficientemente sádicos dispuestos a castigar a alguien advirtiéndole al mismo tiempo 'Te castigo por algo que no has hecho'. Existen, o pueden existir, autoridades o jueces corruptos que cumpliendo con las formalidades estatutarias o legales dispongan el castigo de una persona sabiendo perfectamente de su inocencia. Existe, también, la posibilidad de error en cuanto a la posibilidad real del supuesto ofensor, castigado —en consecuencia— por algo que no hizo. En todos estos casos tiene sentido decir, además, 'A fue castigado pese a que era inocente', o frases similares. Y no sólo tiene sentido decirlo en un plano de estricta corrección lingüística sino que en la mayoría de los casos en que empleamos u oímos expresar tal frase (o frases similares) transmitimos o recibimos una *información* que puede ser o que podemos considerar importante. Parece excesivo e incorrecto, pues, pretender que en estos casos está presupuesto algo así como un absurdo lógico.

Esta objeción puede responderse diciendo que en ninguno de estos casos 'castigo' ha sido empleada correctamente. Y

aún podría intentarse explicar la eventual “corrupción” de estos usos lingüísticos en base a la confusión que parece existir en el uso común entre (a) llamar ‘castigo’ al infligir un sufrimiento al individuo *culpable* y (b) llamar también ‘castigo’ al *mero* infligir un sufrimiento a alguien. En el caso (a), decir ‘A fue castigado pese a que era inocente’ es afirmar algo lógicamente absurdo, tal como lo son los ejemplos dados al presentar esta posición y por razones similares. En el caso (b), decir ‘A fue castigado pese a que era inocente’ es una manera de decir, en realidad, ‘Le dieron una paliza a A’ (o lo privaron de sus bienes, o lo expulsaron del colegio, o lo privaron de su libertad, etc.). Puede resultar interesante e importante llegar a conocer estos hechos, pero la expresión correspondiente no equivale en significado a ‘A fue castigado pese a que era inocente’, en el sentido (a).¹⁷

¹⁷ Quinton ensaya una respuesta distinta, al considerar la objeción de que el castigo de inocentes es *de hecho* posible. Sugiere que ‘castigar’ forma parte de la clase de verbos —a la que pertenece, por ejemplo, ‘prometer’— cuyo uso en primera persona del presente en el modo indicativo posee características peculiares. A su criterio, ‘(Yo) Te castigo por algo que no has hecho’ y ‘(Yo) Te prometo hacer algo que no puedo cumplir’ son expresiones absurdas y por razones análogas. Es impropio decir ‘(Yo) Te voy a castigar’, a menos que aquel a quien dirijo la expresión sea culpable. De la misma manera, es impropio decir ‘(Yo) Te prometo hacer *x*’, a menos que *x* esté dentro de mis posibilidades. Caso contrario, no me encuentro en una *posición* adecuada para enunciar una u otra expresión. Y es *moralmente* impropio hacerlo si sé que aquel a quien dirijo la expresión es inocente o que puedo llevar a cabo el acto prometido. Todo esto no vale para expresiones en tercera persona: ...“así como es perfectamente adecuado decir de otro ‘Él prometió hacer *x*’, sea que él haya pensado o no que podría hacer lo que prometió, resulta perfectamente adecuado decir ‘Ellos lo castigaron’, sea que ellos lo hayan considerado (a él) culpable o no —siempre y cuando ellos hayan dicho ‘(Nosotros) Te vamos a castigar’. Mediante el uso de la primera persona del presente de estos verbos *prescribimos* castigo y *hacemos* promesas. Estas actividades suponen el cumplimiento de condiciones distintas a las requeridas para la *descripción* de lo que quienes prescriben el castigo o hacen la promesa hagan aparecer como castigos o promesas.” (*Op. cit.*, pág. 60 (A).) Quinton procede luego a comparar el funcionamiento de los verbos ‘castigar’, ‘perdonar’ y ‘retribuir’ (en el sentido de conceder un premio por haber hecho algo valioso).

La crítica fundamental que se puede hacer a Quinton es que ‘castigar’ no es un *verbo realizativo*, según la conocida caracterización de J. L. Austin. ‘Prometer’ sí lo es. Si digo, en condiciones adecuadas, ‘(Yo) Te prometo hacer *x*’, al enunciar dicha expresión llevo a cabo al *acto* de prometer. Pero si digo, ‘(Yo) Te castigo por hacer *m*’ es obvio que no llevo a cabo el *acto* de castigar.

Sin embargo, esta respuesta no resulta satisfactoria. Lo que viene a sostener en definitiva es que el único sentido válido de 'castigo' es aquel del que resulta lógicamente absurdo decir —tal como interesa a la "teoría"— 'A fue castigado pese a que era inocente'. El otro sentido de 'castigo' es espurio y debe ser evitado sustituyendo quizá su aparición en los contextos correspondientes por 'pegar', 'azotar', 'privar de libertad', etc.; es decir, por cualquier expresión que designe el acto o acción específico de producir un sufrimiento a alguien. Y todo esto es muy dudoso, aún como táctica: resulta demasiado *ad hoc*, sobre todo cuando se recuerda que la razón principal que lleva a plantear esta

En todo caso *informo* a quien dirijo la expresión por qué lo castigo, o quizá intento atemorizarlo, etc. (de acuerdo al contexto). Es verdad que Quinton sostiene que al usar la primera persona del tiempo presente de 'castigar' *prescribimos* castigo (es decir, llevamos a cabo el acto de *prescribir* castigo), pero este matiz sólo consigue mostrar que en ciertos contextos formales 'castigar' funciona de manera parecida a 'dictar sentencia', o a 'dar un veredicto', etc. Puede ser que en *este sentido* 'castigar' tenga un carácter preformativo, pero (a) el mismo no resulta relevante para la discusión y (b) no afecta la observación anterior de que al usar la primera persona del presente de 'castigar' no llevamos a cabo el acto de castigar.

En cuanto a las peculiaridades del uso de la tercera persona de los verbos realizativos, es correcto sostener que si alguien (A, por ejemplo) *dice* en condiciones adecuadas 'Prometo hacer x', resulta verdadero que A prometió hacer x, con prescindencia de que haya tenido o no intención de llevar a cabo x. (En términos de Austin: la no satisfacción de algunos de los requisitos fijados en las condiciones F no afecta el hecho de que *al decir* A, 'Prometo hacer x', A prometió hacer x.) Lo que no se ve con claridad es la plausibilidad del paralelo que Quinton desea trazar también en este punto con respecto al verbo 'castigar'. Según Quinton, resulta adecuado decir 'Ellos lo castigaron' —sea que lo hayan considerado culpable o no— siempre y cuando ellos *hayan dicho* 'Te vamos a castigar'. Pero esta última exigencia resulta claramente *ad hoc*: que ellos hayan dicho 'Te vamos a castigar' no constituye el acto de castigar y, *a fortiori*, que hayan enunciado o no tal frase no afecta para nada las condiciones en que se puede decir con verdad 'Ellos lo castigaron'.

De lo anterior parece seguirse, como mínimo, que las expresiones 'Te castigo por algo que no has hecho' y 'Te prometo hacer x' (en donde x es un acto de realización imposible) no padecen del mismo tipo de dificultades. Cabe agregar en cuanto a la segunda expresión que hasta es dudoso que pueda tomarse como un caso típico de promesa. De ahí las dudas fundadas que abrigara Austin respecto de este tipo de casos. Va de suyo que resulta criticable, en consecuencia, la equiparación que hace Quinton entre el caso en que el objeto de la promesa es de cumplimiento imposible y aquel otro en que quien promete no cumple con las condiciones mínimas de sinceridad. (Sobre los temas tocados en esta nota, véase J. L. Austin, *How To Do Things With Words* (Oxford, 1962), esp. Lectures I-III).

manera de superar el antagonismo entre retribucionistas y utilitaristas es evitar que la concepción utilitarista quede sujeta a la crítica de que puede llegar a justificar el castigo de inocentes.

Pero quizá esta objeción no sea la más importante. Llegado a este punto el retribucionista puede contraatacar con bastante eficacia incorporando a su concepción del castigo el punto crítico crucial. Las posibilidades de esta maniobra son interesantes. En primer lugar, el retribucionista tiene que reconocer que en tanto y en cuanto se fundamente en un plano meramente lingüístico la tesis de que el castigo se justifica por la culpa que emana de una ofensa acaecida en el pasado, queda por ofrecer una justificación *moral* del castigo. Pero este reconocimiento no clausura la posibilidad de intentar una reformulación de su tesis básica.* (1) Puede, por ejemplo, ofrecer una versión del retribucionismo en la que las nociones de *culpa* y de *castigo* no aparezcan conectadas de modo expreso *en la teoría*, y en la que tal conexión sólo aparezca como una mera recomendación acerca del uso de 'castigo'. Tal es la estrategia propuesta por Mundle. Su enfoque comprende, básicamente, tres elementos: dos postulados éticos y una recomendación verbal. "P.1:...el hecho de que una persona haya cometido una ofensa moral ofrece una razón suficiente para hacerla sufrir; P.2:...si (o cuando) las personas sufren por las ofensas que han cometido, el sufrimiento impuesto debe guardar proporción con la gravedad moral de las ofensas; *recomendación verbal*:... 'castigo' debe aplicarse sólo en aquellos casos en que se hace sufrir a una persona porque lo merece, debido a la ofensa moral que ha cometido."¹⁸ (2) El retribucionista puede distinguir entre la circunstancia de que quien padece el sufrimiento haya sido *de hecho* culpable o haya sido *declarado* tal, sosteniendo entonces que una expresión como 'A fue castigado por algo que no hizo' no resulta

* A tal efecto, el retribucionista tiene abiertas varias opciones.

¹⁸ C. W. K. Mundle, *Op. cit.*, pág. 71 (A).

incompatible con un enfoque de tipo retribucionista que sólo exija que A haya sido *declarado* culpable (formal o informalmente, expresa o implícitamente) con prescindencia de su culpabilidad efectiva. Tal es, básicamente, la salida propuesta por Baier.¹⁹ (3) Puede distinguir, también, entre un sentido débil y uno fuerte de expresiones del tipo 'A fue castigado por algo que no hizo', argumentándose que en un sentido débil tal tipo de expresiones no resulta contradictorio, mientras que en un sentido fuerte sí lo es. Pero entonces la dificultad pareciera extenderse a cualquier concepción de tipo utilitarista, aunque ya no radicaría en la frase '...algo que no hizo', sino en la frase '...por algo...'. "Es evidente —argumenta Armstrong— que 'algo' se refiere a un acto, a un crimen que alguien ha cometido; pero según la teoría utilitarista —sea que ponga el acento en la reforma o en la disuasión— un individuo no es sometido a sufrimiento *por* un crimen, sino *para* disuadirlo a él o a otros individuos de cometer crímenes en el futuro o *para* eliminar la tendencia que tiene a cometer tal crimen. La única teoría con la que tal oración sería compatible en su sentido fuerte sería aquella que definiera el castigo como 'infligir sufrimiento en una persona debido a un crimen que se ha cometido, por esa persona o por cualquier otra'. Y en tanto sepa, tal teoría no ha sido sostenida por ningún filósofo occidental. Que esto sea así ofrece, creo, una prueba de que dicha oración no posee significado en su sentido fuerte."²⁰

Esta nómina de respuestas no pretende agotar las versiones de tipo retribucionista que se han ofrecido, o que podrían ofrecerse, para salvar la dificultad que surge del planteo que estamos analizando en este párrafo. Tampoco pretende sugerir que estas versiones estén exentas de críticas. Lo que muestran, sin duda, es que sin perjuicio de su importancia teórica, dicho planteo no alcanza a superar definiti-

¹⁹ K. Baier, *Op. cit.*, esp. pág. 132 (A).

²⁰ K. G. Armstrong, "The Retributionist Hits Back", *Mind*, 1961, págs. 147-8 (A).

vamente la opción entre retribucionismo y utilitarismo. Y esta conclusión nos lleva a considerar la tercera de las preguntas que hemos formulado oportunamente.

4. ¿Toman en cuenta ambas concepciones el mismo *tipo de fenómeno* o proceso?

En general, hemos considerado hasta aquí *actos o acciones particulares* de castigo. Nuestros ejemplos preferidos han sido 'A castiga a B', o 'A fue castigado por algo que no hizo', o 'A fue castigado pese a que era inocente', etc. En otras palabras, hemos supuesto que los casos que motivan la discrepancia entre retribucionistas y utilitaristas en cuanto a la justificación *moral* del castigo (frase ésta que ya muestra signos de padecer una ambigüedad crónica) son aquellos en los que un individuo inflige un castigo a otro, o en los que el grupo social decide infligir un castigo a uno de sus miembros, etc. En esto hemos seguido la manera en que se ha planteado clásicamente la discusión. Pero en alguna oportunidad hemos caracterizado al castigo calificándolo de *práctica social*, es decir, como una forma de actividad estructurada —de alguna manera— por un conjunto de reglas, normas o pautas de las que resultan qué tipo de acciones son consideradas ofensas, qué sanciones pueden corresponderles, quiénes pueden aplicarlas, etc. Esta descripción (lo dicho dista mucho de ser una definición de 'práctica social') es suficiente para nuestros fines, a saber, distinguir entre dos maneras en que puede hablarse acerca del castigo: como un acto particular o como una práctica socialmente establecida. Y consiguientemente, de dos maneras en que puede intentarse justificar al castigo (una nueva ambigüedad encapsulada en la expresión 'la justificación *moral* del castigo'): justificar que A castigue a B por haber hecho *m*, o justificar que se inflija un sufrimiento a quien ha cometido una ofensa, como una práctica aceptada dentro de un grupo social.

Esta importante distinción fue sugerida por vez primera

en relación con el castigo por W. D. Ross. “El castigo que inflige el estado —dice— no se produce, ni se debe producir, por generación espontánea. Es precedido por la emisión de una ley en la que se atribuye una pena a un crimen. . . . Debemos distinguir, creo, esta etapa, que consiste en la atribución de una pena, de aquella otra que consiste en su aplicación. Y podemos preguntar en base a qué principios el estado o sus funcionarios deben actuar en cada etapa.”²¹ Más tarde, la distinción fue retomada por Quinton y, especialmente, por Rawls —que intenta proyectarla a un ámbito teórico más extenso.

La pregunta acerca de si el retribucionismo y el utilitarismo toman en cuenta el mismo *tipo de fenómeno* o proceso puede pues, ser respondida negativamente. La distinción que acabamos de trazar sugiere de modo inmediato el por qué. Cuando los utilitaristas sostienen que el castigo sólo puede ser justificado *moralmente* en base a sus consecuencias valiosas tienen presente al castigo como práctica social; y su planteo parece hacer justicia a la única manera plausible en que se puede llegar a justificar una práctica social: por sus consecuencias valiosas. Cuando los retribucionistas afirman que el castigo sólo puede ser justificado *moralmente* por la culpa del ofensor consideran al castigo como lo que resulta de aplicar una norma, regla o pauta a un caso particular; y su planteo parece hacer justicia a la única manera en que tal aplicación puede tener lugar: tomando en cuenta las condiciones antecedentes fijadas en la norma, regla o pauta. La oposición entre ambas concepciones parece pues diluirse como consecuencia de un afinamiento conceptual que comúnmente ha sido pasado por alto.

Claro que la distinción tiene, además, otras consecuencias teóricas. La más importante es que en base a ella puede superarse la reiterada objeción de que el utilitarista puede llegar a justificar el castigo de personas inocentes. El argu-

²¹ W. D. Ross, *Op. cit.*, pág. 61.

mento es relativamente sencillo. En tanto y en cuanto la tesis utilitarista se pretende aplicar a casos *particulares* en los que se inflige sufrimiento a un individuo dado, la objeción parece tener fundamento. Pero una vez que se limitan las pretensiones teóricas de la concepción utilitarista a la justificación *moral* del castigo como práctica social, la objeción se desvanece, sencillamente por falta de blanco adecuado: el ámbito de los casos específicos de aplicación de castigo queda libre a los retribucionistas para que hagan valer en él su intento de justificación.

Pero, ¿no puede el retribucionista volver a plantear nuevamente su conocida objeción? ¿No podría el utilitarista llegar a justificar *moralmente* el castigo *institucionalizado* de personas inocentes si se pudiera lograr, de tal manera, un fin que de acuerdo a sus principios resulte valioso?

Ante este esbozo de objeción puede replicarse inquiriendo cómo podría llegar a institucionalizarse el castigo de personas inocentes: ¿cuál sería, o cómo sería tal “práctica” social? Un ejemplo podría ser éste. Supongamos que paralelamente a la práctica usual del castigo se crea una “práctica del castigo” que consiste en la autorización expresa a una persona o a un grupo de personas o “castigar a cualquier miembro del grupo social (elegido al azar), acusándolo de haber cometido una “ofensa” similar a un tipo de ofensas que en ese momento se extiende en el grupo. Como se ve, las personas autorizadas a poner en práctica esta forma de “castigo” no obran de una manera totalmente arbitraria: tiene que producirse una ola de ofensas que consideren dañinas para el grupo social y deben, además, acusar formalmente al presunto “culpable”. Son libres, en cambio, de determinar cuándo tal ola de ofensas se produce, si son o no dañinas, quién va a ser el “culpable” y qué “castigo” va a recibir. Esta pareciera ser una manera de *institucionalizar* el castigo de personas inocentes.

¿Pero —puede argüir en este punto el utilitarista— qué consecuencias valiosas podría reportar al grupo social esta

curiosa “práctica del castigo”? ¿Qué fin *útil* se obtendría con la implantación de tal sistema que sólo conseguiría producir entre sus miembros la sensación de estar sujetos a una peligrosa inseguridad? Como bien ha señalado Rawls, ¿cómo sabría un miembro del grupo cuándo recibe un castigo *auténtico* y cuándo padece el “castigo” que emana de esta forma extravagante de control social? En síntesis, pareciera imposible poder llegar a ofrecer una justificación utilitarista a esta “práctica del castigo”. Y de ello pareciera seguirse que el argumento en favor de la posible *institucionalización* del castigo de personas inocentes carece de peso.

Claro que en este punto pueden expresarse algunas dudas. Aun reconociendo la evidente importancia teórica que posee distinguir entre la justificación del castigo como *práctica social* y la justificación de un *acto específico* de castigo, puede argumentarse —entre otras cosas— que esta manera de salvar al utilitarismo de la conocida objeción acerca del castigo de inocentes expone al utilitarista a un tipo de crítica de un nivel distinto al de las ensayadas hasta ahora, a saber, la que lleva a cuestionar el contenido y valor teórico de una justificación utilitarista (en general). En otras palabras, lo que puede exigirse al utilitarista es una explicación adecuada de la distinción que propone y, en último extremo, de la teoría moral que sostiene, toda vez que es en estos elementos en los que funda su tentativa de superar la objeción retribucionista y su valorización de la concepción utilitarista del castigo.²²

²² Esta objeción —sumariamente expresada— puede plantearse, en realidad, en dos niveles. (1) La distinción entre justificar una práctica social y justificar una acción particular que pertenece a ella constituye la esencia de una reciente interpretación de ciertos textos de J. S. Mill de la que resulta la posibilidad de ofrecer una versión del utilitarismo que esté exenta de ciertas críticas tradicionalmente formuladas en su contra. Dicha interpretación se basa en textos expresos del propio Mill (Cf. J. O. Urmson, “The Interpretation of the Moral Philosophy of J. S. Mill”, *Philosophical Quarterly*, 1953). Es así como se habla de dos versiones del utilitarismo: el utilitarismo de actos (act-utilitarianism) y el utilitarismo de reglas (rule-utilitarianism). El artículo de Rawls (“Two Concepts of Rules”) intenta mostrar —tomando como ejemplos el tema del castigo y el de las promesas— cómo pueden superarse, en base a la distinción, ciertas críticas al utilitarismo. Aunque Rawls advierte cautamente que no pre-

No debe inferirse de esta observación que el utilitarista no puede llegar a encontrarse en condiciones de responder adecuadamente a dicho planteo crítico. Lo que queremos puntualizar es que *prima facie* tampoco es dable encontrar en el enfoque que acabamos de analizar, una superación definitiva de la oposición existente entre las dos concepciones clásicas del castigo.

IV

1. Hagamos aquí un breve balance crítico.

El análisis de las tres alternativas que planteáramos en III.1 arroja, desde cierto punto de vista, un resultado negativo: la *impasse* a la que llegan las dos concepciones clásicas del castigo no parece superarse en base a la suposición de que formulan preguntas distintas acerca del castigo, o de que intentan dar solución a diferentes tipos de problema, o de que toman en cuenta distintas clases de fenómenos. De una u otra manera, la discrepancia entre los partidarios de ambas concepciones sigue en pie luego de considerar cada una

tende probar que el utilitarismo quede exento, de esta manera, de toda crítica. Y por cierto que su cautela está justificada. Ahora bien. Teniendo en cuenta estos antecedentes, la objeción que formulamos en el texto viene a sostener que la defensa de la concepción utilitarista del castigo basada en la justificación del mismo como práctica social *presupone* el utilitarismo de reglas y que, en consecuencia, la plausibilidad de dicha concepción depende —en última instancia— de las virtudes teóricas de esta versión del utilitarismo. Aún concediendo, por ejemplo, que de tal manera se elimina la posibilidad de justificar el castigo de inocentes, queda por discutir la versión del utilitarismo que permite lograr tal resultado. (2) Al margen de estas disquisiciones, puede ponerse en duda la conclusión de que *nunca* podría justificarse desde un punto de vista utilitarista “la práctica del castigo” (la *institucionalización* del castigo de inocentes). Supongamos un grupo social conmovido profundamente por cierto tipo de crímenes horribles, cuyos miembros son sumamente propensos a abstenerse de realizar aquellas acciones sujetas a castigos muy graves. ¿Por qué no instituir, *en tal caso*, “la práctica del castigo?” ¿No quedaría neutralizada la inseguridad que “la práctica del castigo” proyecta sobre el grupo social, por el posible logro del fin valioso deseado? ¿No se estaría entonces frente a una situación que admite una justificación utilitarista? Puede responderse que no, pero ¿qué fundamento podría darse en favor de tal respuesta? La objeción que formulamos en el texto viene a sostener —en este nivel— que intentar contestar, a su vez, esta pregunta implica internarse en las conocidas dificultades que plantea el “principio de utilidad”.

de estas alternativas —o al menos, sigue en pie en tanto y en cuanto nuestro análisis crítico ha podido señalarla. Todo esto parece claro. Aunque desde otro punto de vista, también resulta claro que a lo largo de la discusión hemos ido acumulando elementos de prueba suficientes como para dar fundamento a la sospecha —meramente sugerida en I.3— de que la discrepancia entre las dos concepciones clásicas reconoce, de alguna manera, un punto de partida equivocado.

Si seguimos esta línea de pensamiento es posible transformar los resultados negativos que acabamos de mencionar, en *resultados positivos*. En efecto, si se admite que de la comparación crítica de ambas concepciones no surgen criterios suficientes como para decidirnos a favor de una de ellas (Cf. III.1) y se acepta que las tres opciones ofrecidas como medio de encontrar una salida a tal situación están expuestas a críticas bastante fundadas (Cf. III, 2-4), pareciera que tenemos ante nosotros una especie de *reductio ad absurdum* de los supuestos teóricos que comparten ambas concepciones del castigo. El más importante y nocivo de tales supuestos está constituido por la idea de que encarar el tema del castigo implica, necesariamente, que hay que lidiar con *un* problema (cuyo contenido no se especifica con claridad) al que debe darse *una* respuesta unívoca. Esta es la “imagen” teórica inadecuada, a la que hiciéramos referencia en I.3.

Sin embargo, arribar a esta conclusión tiene sólo un valor relativo si no se ofrece —o, al menos, se esboza— un enfoque teórico alternativo. Pero a esta altura de la discusión no parece difícil llegar a sugerirlo. Se trata, simplemente, de reconocer que el tema del castigo presenta una multitud de facetas que parecen escapar —al menos en el estado actual de la investigación— toda posibilidad de ubicación dentro de una respuesta unitaria. No debemos preocuparnos, pues, por discutir si la solución correcta se encuentra dentro de la concepción utilitarista o de la retribucionista. Tampoco debemos disipar energías buscando variantes que permitan

dirimir el pleito apelando a sutiles distinciones. Lo que debemos hacer es enumerar una serie de temas no necesariamente vinculados entre sí —al menos, *prima facie*— tratando de fijar su contenido. La segunda etapa debe consistir, obviamente, en un análisis detallado de cada uno de ellos. Esta manera de encarar el tema del castigo difícilmente da cabida a conclusiones grandilocuentes, pero no hay duda que resultará, en definitiva, la más fructífera desde un punto de vista filosófico.

En lo que sigue intentaremos llevar a cabo la primera de estas dos etapas. Enumeraremos varios temas de discusión y formularemos breves comentarios en cuanto a su contenido. Esta tarea, unida a la ya realizada en las secciones II y III, ayudará a aclarar el sentido (o los sentidos) en el que el filósofo tiene algo que decir sobre el tema del castigo. La tarea es claramente preparatoria. Pero no por ello debe minimizársela. En definitiva, desbrozar el terreno es condición necesaria de toda siembra.

2. *La definición de 'castigo'*: 'Castigo' es una palabra cuyo significado no presenta mayores dificultades. El *Diccionario de la Real Academia* la define como "pena que se impone a quien ha cometido un delito o falta". A su vez, 'pena' es definida como "castigo impuesto por superior legítimo a quien ha cometido un delito o falta". Si se deja de lado el obvio "círculo" entre 'castigo' y 'pena', debe reconocerse que el *Diccionario* recoge el sentido básico que estamos dispuestos a otorgar a la palabra (algo que no siempre ocurre, tratándose del *Diccionario*). Flew ha sistematizado las principales notas definitorias de castigo. Para que haya castigo (a) debe infligirse a una persona un daño, (b) debe el mismo ser infligido con motivo de una ofensa (o, supuestamente, por tal motivo), (c) debe ser aplicado al ofensor (o a quien se considere, supuestamente, tal), (d) debe ser el resultado de una acción humana y (e) debe ser impuesto por alguien con autoridad conferida por la institución en contra de cuyas le-

yes o normas se ha cometido la ofensa. Todas estas notas están implícitas en la definición del *Diccionario*. El conjunto de ellas recoge el *uso standard* del término.

Si este es el caso, ¿cómo es que puede surgir un problema de tipo filosófico respecto de la definición de ‘castigo’? Hay dos aspectos que pueden producir tal tipo de problema. El primero tiene que ver con el *carácter de la ofensa*. ¿Debe restringirse a ofensas legales o debe ampliarse a ofensas morales? Hart, por ejemplo, se inclina por considerar que el uso standard de ‘castigo’ tiene relación con el primer tipo de ofensas. Flew, por su parte, opina que el uso standard debe recoger —o, mejor dicho, recoge— ambos tipos de ofensas. La discusión de este punto tiene particular importancia. Entre otras cosas, se conecta directamente con la cuestión relativa a si existe un sentido en que pueda hablarse legítimamente de sanciones morales. Si un amigo no cumple una promesa que me ha hecho, su acto puede describirse —entre otras cosas— como la realización de una ofensa moral. ¿Pero puede decirse que cuando “le retiro mi amistad” lo estoy sancionando? Y aunque esto tenga sentido, ¿puede decirse, *a fortiori*, que lo estoy castigando?

El segundo aspecto se relaciona directamente con el anterior. Si se adopta un criterio amplio en cuanto a la naturaleza de la ofensa se abre un marco de posibilidades en cuanto al empleo adecuado de la palabra ‘castigo’. Pero, obviamente, existen casos en los que se usa la palabra que no parecen ajustarse a las condiciones que determina su empleo standard. Reconstruir el complicado mapa que va desde la privación de libertad que sufre el delincuente, a la paliza que propina el padre al hijo y a la sucesión de mortíferas trompadas que recibe un boxeador en el ring —casos, todos ellos, en los que hablamos de castigo— constituye una importante tarea para el filósofo. La catalogación de los usos secundarios de la palabra puede ser fuente de interesantes revelaciones.

3. *La justificación del castigo*: A través de la discusión

que realizamos en la sección III ha quedado en claro que cuando se intenta encarar el problema de justificar *moralmente al castigo* (y, en general, cuando se trata de justificar *moralmente* acciones humanas), no debe confundirse justificación con definición y debe distinguirse entre justificar acciones individuales y justificar prácticas sociales. Estas importantes distinciones no pueden pasarse por alto. Pero es interesante observar que poseen un carácter negativo: no establecen criterios acerca de la justificación *moral* sino que llaman la atención sobre aspectos que deben distinguirse cuidadosamente cuando se trata de ofrecer, como en el caso del castigo, una justificación *moral*. El problema referente a tales criterios sigue, pues, en pie. Y debe reconocerse que muy poco se ha avanzado sobre el tema, pese a la extensa bibliografía relacionada directa o indirectamente con él.

En III.2 hemos llamado la atención sobre la relación que existe entre la justificación *moral* del castigo que ofrecen los utilitaristas y los retribucionistas y el *tipo* de teoría moral que subyace a cada una de ellas. Y consideramos que no resulta descaminado plantearse el problema de la justificación *moral* (en general) como íntimamente conectado con una tipología de las teorías morales. Es obvio, que la adopción de una posición utilitarista —por ejemplo— determina una concepción acerca de la justificación *moral* de las acciones que va a ser diferente de la que resulta de adoptar una posición de carácter deontológico. El problema parecería trasladarse, pues, a los posibles criterios para decidir entre uno u otro tipo de teorías morales. Desde otro punto de vista, podría sugerirse que, en todo caso, la expresión ‘justificación *moral*’ posee distintos sentidos que dependen del carácter de la teoría moral dentro de la cual se emplea. Estos dos aspectos son, sin duda, importantes y permiten empalmar la discusión acerca de la justificación *moral* del castigo con uno de los temas más trascendentes de la ética.

4. *Responsabilidad*: Íntimamente conectados con la discu-

sión acerca del castigo, tal como la hemos presentado, existen importantes problemas planteados por la familia de conceptos de la que *castigo* forma parte. *Culpa, responsabilidad, perdón, merecimiento* (del castigo) son algunos de los miembros de dicha familia. Nuevamente debemos consignar que el análisis de tales conceptos hace ingresar la discusión del tema en uno de los capítulos importantes de la ética. El caso es evidente cuando se toma en cuenta el concepto de *responsabilidad*. Y pareciera que dudosamente podría realizarse una investigación acerca del castigo sin caer, tarde o temprano, en la discusión de este concepto. A diferencia de lo que ocurre con el problema de la justificación *moral*, existe una importante bibliografía sobre el tema, que comienza en Aristóteles y ha recibido recientemente interesantes aportes por obra de Hart.

5. *Otros temas*: A estos tres temas que, podría decirse, constituyen una íntima trama, deben agregarse varios otros. Algunos poseen carácter filosófico. Otros se relacionan con cuestiones empíricas. Un problema que, en tanto sepamos, sólo ha merecido consideración en la bibliografía reciente por parte de Mabbott es el de la *existencia misma del castigo como práctica social*. No se trata ya de justificar *moralmente* el castigo (pues esto supone que no se lo cuestiona como práctica) sino de preguntarse por las condiciones que lo hacen posible y por su eventual prescindencia. Una discusión de estos tópicos requiere previamente una investigación de tipo empírico: la psicología social y la sociología son candidatos inmediatos para cumplir tal función. Investigaciones de tipo empírico son relevantes también a toda discusión acerca del *monto y efectos de las penas* —un tema al que los criminólogos han prestado especial atención.

Quizá resulte conveniente detener aquí esta enumeración. El objetivo principal que perseguimos en este trabajo es el de introducir la discusión de un tema que no sólo es importante por su valor intrínseco sino, también, por las conexiones

inmediatas que posee con ciertos capítulos tradicionales de la filosofía moral. También asignamos a este trabajo el objetivo adicional de ayudar a romper cierta manera tradicional de encarar tal tema. Aunque negativo, este paso resulta fundamental para lograr una posición teórica que rinda frutos positivos.

SUMMARY

It would be philosophically inadequate to consider that there is just one problem involved in the practice of punishment and, therefore, just one answer to search for. In fact, punishment gives rise to many problems, and it is necessary to appeal to various criteria in order to solve them. However, we cannot deny the possible existence of a common *motivation* which is at the root of all the difficulties concerning the practice of punishment.

The deprivation of life, of freedom, of material wealth are evils that in normal situations deserve our moral disapproval. But why should an act which seems to be the same as another, deserve in one case moral disapproval and in some other, justification or at least the withholding of moral judgement? The answer *prima facie* could be that such an act is morally justified because *A* punishes *B* because *B* did something he *ought not* have done. This would be different than when *A* beats *B* just because it pleases him. In the first case *A*'s act would be punishment of an offender, in the second it would be an offense. But if *B* has committed an offense, and punishing *B* implies inflicting an "evil" upon him, the difficulty is still there; it seems unjustifiable to produce an "evil" just because another "evil" has taken place. The moral justification of punishment and of all the questions surrounding this topic can be regarded as a refinement of the discussion concerning the moral justification of actions in general.

The following section of this paper states two traditional views about punishment: Retributionism and Utilitarianism. The third section deals with three different approaches to the same problem which pretend to overcome the difficulties in reconciling the above mentioned opposing views. The last section considers an interesting family of problems which must be taken into account in any serious treatment of punishment.

Retributionism is principally based on the thesis that punishment inflicted upon a man is morally justified because that man is guilty, i.e., has committed an offense. What follows logically from any Retributionist view is that if punishment is justified because of the offender's guilt, then the severity of the punishment must correspond exactly to the graveness of the offense. This gives rise to the problem of offering the adequate criteria to calculate the just punishment according to the graveness of the offense.

According to Utilitarianists, if one wants to give an adequate mo-

ral justification of punishment, one must look into the future and not into the past; the only way of justifying punishment is by taking into account the valuable consequences that can be derived from its application. Punishment can be applied in order to produce an improvement in the offender, or it can be justified as a means to prevent possible repetition of the offense in the future. Utilitarianism has an answer to the problem of how to calculate the just punishment according to the graveness of the offense: the severity of the punishment must be such that it is not to the offender's advantage to commit the transgression and receive the punishment, or it must be such that the punishment achieves its aim.

Retributionists have an objection against Utilitarianism: if punishment is considered morally justified because of its consequences, then it would be justifiable to punish an innocent if one is certain that the desired consequences would be produced. Moreover, they argue that Utilitarianists make no distinction between punishment properly speaking and other "external" factors (threats, publicity, psychiatric treatment, etc.), and therefore, their claim to justify punishment by its consequences is mistaken.

Utilitarianists also raise strong objections against the Retributionist point of view. Their contention is that Retributionism seems to imply the unjustifiable thesis that punishment is valuable in itself. Punishment can be regarded as the rationalization of the desire for revenge that the offended feels toward the offender. It could also be questioned whether, on one hand, guilt is the only relevant thing to take into account when a punishment is applied, and on the other, whether the punishment given the guilty is always morally justified. Concerning the difficulties involved in the question of how to establish the adequate punishment for any offense, the Utilitarianists raise another objection against Retributionism: either it is impossible to accurately determine the right proportion, or it would only be possible to do so taking into account the valuable (Utilitarianist) consequences that would be derived from the application of the punishment.

At this point of the analysis there are no adequate criteria to decide in favor of one view or the other, as we do not see clearly enough what are the problems to be solved. The expression "morally justify punishment" covers too many things, and so it cannot help to outline the borders of the discussion. However, there are some questions that we can state now:

Do Retributionism and Utilitarianism try to answer the same questions concerning punishment? The first difference (and a trivial one perhaps) that can be noted between these two views is that

Retributionism seems to be asking *why* we punish, whereas Utilitarianism is asking *for what purpose* do we punish. If the questions are different it is obvious that the answers will also be different. If we ask someone "Why do you do *p*?", the answer must necessarily make reference to a past action or past fact which is in some way connected with *p*. If the question is "For what purpose do you do *p*?", the answer will make reference to a fact of the future, i.e., to the consequences that the agent is trying to produce upon carrying out *p*. It would follow that: (a) there is no real theoretical opposition between Retributionism and Utilitarianism, (b) since the two views have a certain degree of acceptability, one can admit that their answers are both correct simultaneously, (c) the basic mistake of both positions lies in their contention that their view offers *the* moral justification of punishment, (d) it would be possible perhaps to integrate both positions into a single general theory of punishment. But these consequences are not acceptable since the original premise is not admissible. The fact that each position gives different answers to the problem of morally justifying punishment does not follow from the trivial fact that they state different questions before the same problem but rather from the fact that they adopt different theoretical positions before the problem of the moral justification of an action. According to the Utilitarianists, punishment, like any other human activity with a moral quality, can only be justified by its consequences. Retributionism, on the other hand, considers that the justification can only be sought in facts of the past, including the intention of the agent. Therefore they each formulate such questions as they consider to be the *kind* of questions that must be stated before that *kind* of problem.

It must be noticed that the distinction we have traced between "Why is *B* punished?" and "For what purpose is *B* punished?" is not clear-cut. It has some exceptions. "Why do you do *p*?" is a question that can be answered by making reference either to an act of the past or to an act of the future; it is therefore an ambiguous question. Thus it is hard to determine the criteria to be used in order to distinguish between the two kinds of questions which are the basis of the dispute.

Let us try another way of overcoming the opposition between Retributionists and Utilitarianists. Do they give a solution to the same problem, in the sense that both of their theses can be considered to belong to the same philosophical compartment? A negative answer would require a much more complicated formulation of the problem. The strongest objection of Retributionists against Utilitarianists is that the latter's view would not exclude the possibility

of justifying punishment of an innocent. Now, they claim it seems logically impossible to punish someone who has not committed an offense. Being guilty is a necessary condition for being punished. Therefore it would be logically absurd to speak of punishment inflicted upon an innocent. Punishment is punishment only when it is deserved, otherwise it would be only an injustice, a crime. All this is in accordance with the standard use of the word "punishment" or of expressions such as "*A* punishes *B*."

This attempt to overcome the opposition between the two views we are considering could yield some consequences. First of all, it would be obvious that Retributionists give a logical answer, i.e., they try to clarify the meaning of the term "punishment" or the conditions of the correct application of such expressions as "*A* punishes *B*." Utilitarianism, on the other hand, would be giving an answer of ethical character: the attempt to morally justify punishment. Secondly, it might provide a distinction between two different aspects of the phrase "the moral justification of punishment", a conceptual one and an *ethical* one. Thirdly, it could result that some sentences used in ordinary contexts are absurd. Finally, the criticism against Utilitarianism would lose all of its support, since it would be logically impossible to punish an innocent.

However, there is a possible objection to this view: in point of fact, there are cases of corrupt judges or sadistic people who know that an accused is innocent and still decide to inflict punishment on him. There is also the possibility of making a mistake concerning the culpability of the accused. In these cases it would make perfect sense to say "*A* was punished even though he was innocent." This objection could be met by replying that the linguistic usages supporting the objection are in fact confused, i.e., that there is a confusion in ordinary usage between (a) using "punishment" to refer to the action of inflicting suffering upon a guilty person, and (b) using "punishment" to refer to the *mere* action of inflicting suffering on anyone. However, this reply will not suffice since it presupposes that the only criterion for the correct usage of "punishment" is precisely the one according to which it would be absurd to say "*A* was punished even though he was innocent." This would be a question-begging device.

Retributionists can still strike back. For this purpose they must admit that it is necessary to give a moral justification of punishment. They have three possibilities open for stating their thesis in a new way: (1) They can offer a version of their view according to which the concepts of guilt and punishment are not explicitly linked together, but where such a link appears in the form of a

recommendation concerning the use of "punishment". (2) They can distinguish between the situation of someone who receives a suffering because, *in point of fact*, he is guilty, and that of someone who has simply been *declared* guilty. Thus the expression "*A* was punished for something he did not do.", would not be incompatible with a Retributionists view according to which the only necessary condition to meaningfully use the above statement is that *A* has been *declared* guilty. (3) It can be distinguished between a "strong" sense and a "weak" sense in such expressions as "*A* was punished for something he did not do." In its strong sense, the expression would be self-contradictory whereas in its weak sense it would not. (However, it would be possible to show that such an expression is deprived of meaning in its strong sense.)

This new attempt to reconcile the differences between Retributionism and Utilitarianism has not been fully successful. Let us consider a third question concerning the two opposing views. Do they both take into account the same kind of phenomena or process? We must distinguish between two different ways of considering punishment: as an individual act and as an established social practice. Therefore one can attempt to justify punishment in two different ways: by justifying the fact that *A* punishes *B* because he did *p*, or by justifying suffering inflicted upon someone who has committed an offense because it is a practice accepted within a social group. Now we can answer the question. It seems that Retributionism and Utilitarianism do not take into account the same phenomena or process. When Utilitarianists maintain that punishment can only be *morally* justified on the basis of its valuable consequences, they are assuming punishment to be a social practice. When Retributionists claim that punishment can only be justified on the grounds that the offender is guilty, they are considering punishment as the result to the application of a general rule to a particular case. Thus, the opposition between these two groups seems to dissolve as a consequence of a conceptual refinement that generally has been overlooked. Another important consequence is that the objection against Utilitarianism, in the sense it makes possible the justification of punishment of innocents, can be overcome. Once the theoretical claims of this view are limited to the moral justification of punishment as a social practice, the grounds for the objection disappear. Retributionism, on the other hand, can validate its attempt to morally justify punishment in those cases of particular applications of punishment.

But is it not possible that Retributionists will again raise their objection against Utilitarianists, in the sense that they can still

morally justify *institutionalized* punishment (valuable by its consequences) inflicted upon an innocent person? The answer is no. According to Utilitarianism there would not be any valuable consequence if this were the case, and therefore it would be impossible to give any Utilitarianist justification for such a "practice of punishment".

Even if one recognizes the theoretical importance of the above mentioned distinction between the justification of punishment as a *social practice* and as an *individual act* one can still question, among other things, what is the content and theoretical value of a Utilitarianist justification in general. In short, one can demand an adequate explanation to Utilitarianists of the distinction they propose, and as a last resort, of the moral theory they maintain. So, *prima facie*, this last approach also fails to overcome the opposition between the two views.

The critical contrast that has been offered between Utilitarianism and Retributionism does not yield sufficient criteria to decide in favor of either of them. The three approaches we have considered in order to overcome the opposition are all subject to criticisms. However, we have sufficient elements to assert that the grounds for this opposition is a mistaken point of departure. Among the assumptions basic to both theories, the most important and pernicious is the idea that there is just *one* problem involved in any treatment of punishment, and, therefore, there must be just *one* answer to such a problem.

Once we have come to this conclusion, it would have only relative value were we not to offer—or at least sketch—an alternative way of approaching the problem. Punishment involves a variety of factors that cannot be easily included within a unitary answer. At the first stage we must enumerate a series of topics connected with punishment but not necessarily related among themselves (at least *prima facie*) and try to see what their real content is. Secondly, it would be necessary to analyse those topics. In what follows we shall try to reach just the first stage.

The Spanish Dictionary of the Royal Academy defines "punishment" as a "penalty inflicted upon someone who has committed a fault or an offense." "Penalty" in turn is defined as "punishment imposed by a legitimate superior on someone who has committed a fault or an offense." According to Flew, the definitory characteristics of "punishment" are: (a) a penalty must be inflicted upon someone, (b) the penalty must be inflicted because an offense has been committed, (c) it must be inflicted upon the offender, (d) it must be the result of a human action, and (e) it

must be imposed by an authority of the institution whose laws or rules have been broken. This gives the standard use of the word "punishment".

There are two philosophical problems that may arise from this definition of "punishment". Must it be restricted just to legal offenses or must it also be extended to moral ones? There is a disagreement concerning this question, and its discussion is connected to the question of whether one can legitimately speak of moral sanctions. The second problem is related to the first one: if we loose the criterion to determine the nature of the offense, then the possibilities of correct usage of the word "punishment" would be multiplied. But, obviously, there are cases in which the word is applied although it does not seem to fit the conditions that determine its standard use.

It now seems clear that when one attempts to face the problem of the moral justification of punishment, one must distinguish between justification and definition, and between justifying individual actions and justifying social practices. Although these distinctions do not establish the criteria for the *moral* justification of punishment they call attention to aspects that must be carefully distinguished before giving any moral justification. The problem of establishing such criteria is still there.

We have previously noticed that there is a relation between the moral justification of punishment as offered by Retributionists and Utilitarianists, and the kind of moral doctrine underlying them. The problem of moral justification seems intimately connected with that of the description of the different kinds of moral theories. Thus the problem would seem to be reduced to that of finding the criteria to decide in favor of one or the other theories. It can also be suggested that the expression "moral justification" can have different meanings according to the theory in which it is used.

There are some other important problems which arise from the family of concepts to which punishment seems to belong. Guilt, responsibility, and pardon are among these concepts. Another problem is that of the existence of punishment as a social practice. The question is not its justification, but rather asking for the conditions which make it possible and the examination of its possible eventual disappearance. A discussion of these topics would require a previous empirical investigation (psychological and social).

The principal aim of this paper has been that of introducing the discussion of an important topic which is connected to other traditional problems of moral theory. This paper has also attempted to break with the traditional way of facing the problem of punishment.