



# Castigo y normas sociales. Acerca de la validez ecológica del castigo en la economía experimental, parte II

## Punishment and social norms. On the ecological validity of punishment in experimental economics, part II

**Maximiliano Senci** ([maximiliano.senci@uns.edu.ar](mailto:maximiliano.senci@uns.edu.ar)) Instituto de Investigaciones Económicas y Sociales del Sur, Universidad Nacional del Sur - CONICET (Bahía Blanca, Argentina) ORCID: 0000-0001-9131-3843

### Abstract

This article corresponds to the second part of the article entitled Punishment and social norms: on the ecological validity of punishment in experimental economics. In the first part I developed a series of arguments that throw uncertainty about the legitimacy of inferring that the cause of the application of punishment in experimental games is univocally the transgression of social norms. In this part, I offer a theoretical defence of a method that has the potential to respond to the problems of normative irrelevance and the illegitimacy of punishment in experimental games, and which consists in the introduction of explicit norms.

**Key words:** punishment, social norms, ecological validity, legitimacy, experimental economics.

### Resumen

Este artículo corresponde a la segunda parte del artículo titulado Castigo y normas sociales: acerca de la validez ecológica del castigo en la economía experimental. En la primera parte me ocupé de desarrollar una serie de argumentos que arrojan incertidumbre acerca de la legitimidad de inferir que la causa de la aplicación del castigo en los juegos experimentales es unívocamente la transgresión de normas sociales. En esta parte me ocupo de ofrecer una defensa teórica de un método que tiene el potencial de responder a los problemas de la irrelevancia normativa y la ilegitimidad del castigo en juegos experimentales, y que consiste en la introducción de normas explícitas.

**Palabras clave:** castigo, normas sociales, validez ecológica, legitimidad, economía experimental.

### Introducción

Como ha señalado acertadamente Francesco Guala, los experimentos de castigo tienen una lectura acotada (narrow) y otra amplia (wide). De acuerdo con Guala: “Under the narrow reading, punishment experiments are just useful devices to measure robust psychological propensities (‘social preferences’) in controlled laboratory conditions. Under the wide reading, they replicate a mechanism that supports cooperation also in ‘real-world’ situations outside the laboratory” (Guala 2012:2).



En otras palabras, mientras que la lectura acotada se refiere a la validez interna de los experimentos, la lectura amplia se refiere a la validez ecológica, es decir, a la pretensión de que los resultados experimentales puedan ser extrapolados a situaciones de la vida cotidiana. Si bien la lectura acotada no está puesta en duda, en cambio la amplia, como el propio Guala sostiene, dista mucho de tener evidencia empírica rotunda que la respalde. En la Parte I del artículo expuse dos argumentos que interfieren con la validez ecológica: la irrelevancia normativa del castigo y la (i)legitimidad del mismo. En esta parte voy a proponer un método que tiene el potencial de solucionar dichos problemas, con el objetivo de que los experimentos de castigo sean más realistas desde el punto de vista psicológico, lo que contribuiría a su mayor validez ecológica. El método recurre a la literatura reciente sobre normas sociales, y particularmente a una línea de investigación reciente que sostiene que los sujetos experimentales exhiben una preferencia por “hacer lo correcto”.

## 6. Marcos, demanda y preferencias morales

Los economistas han buscado estudiar los fenómenos de interés en el laboratorio por medio de la definición correcta de incentivos monetarios, y de su manipulación, de forma tal de poder capturar efectos de tratamiento. Un “tratamiento” es una variación en los incentivos monetarios en relación con un control. Los economistas han intentado capturar diferencias en el comportamiento a partir de la sola variación de la estructura de incentivos monetarios, esto es, de la estructura de pagos de una situación. Como la variación en los pagos es típicamente la manera de identificar efectos de tratamiento, los economistas evitan brindar a los participantes información que pudiera intervenir o sesgar sus decisiones y, por lo tanto, recurren mayormente a instrucciones neutrales o sin contexto. En este tipo de experimentos las opciones del juego y los incentivos monetarios están descritos de manera tal de que se evite cualquier referencia a palabras de la vida cotidiana. Esto es así porque es sabido que el fraseo incidental de las instrucciones que puede acompañar la definición y descripción del juego puede tener efectos no intencionados (y, por ende, no controlados) en el comportamiento de los participantes. Evitar las instrucciones contextualizadas ha sido y es, por ende, un desiderátum de los economistas. Sin embargo, hay razones para pensar, que lejos de ser contraproducentes, las instrucciones contextualizadas pueden ser un recurso valioso. Por ejemplo, cuando el fenómeno tiene componentes sociales ineludibles y puede suscitar reacciones emocionales (e. g., la corrupción), sería deseable que en el laboratorio se pudieran replicar dichas condiciones. Por ende, puede ser no solo adecuado, sino deseable que los experimentos se realicen “en contexto”. Esto se puede lograr de varias formas: 1) las instrucciones pueden asignar roles a los participantes (tales como el de vendedor, comprador, empresa o funcionario público), lo que les permitiría asociar las opciones de juego a lo que se espera de esos roles en la realidad, y por ende a las normas que cada rol o identidad social tiene asociadas como parte de su definición. Esto dotaría al juego de mayor realismo y por lo tanto de mayor validez ecológica; 2) las estrategias mismas pueden ser nombradas tal como lo serían en la realidad y no en un lenguaje abstracto (“castigo o multa”, por ejemplo, en lugar de “opción X”); 3) el juego puede ser denominado de una manera específica (por ejemplo, “Juego de Wall Street” o “Juego de la Comunidad”).

La razón principal por la cual los economistas se han mostrado reticentes a la utilización de marcos consiste en que estos pueden favorecer un efecto de demanda del experimentador (DE a partir de aquí). Los efectos de demanda ocurren cuando el experimentador sugiere la respuesta correcta a los participantes, y estos se comportan de manera tal de complacer al experimentador. La DE, según Zizzo: “refers to changes in behavior by experimental subjects due to cues about what constitutes appropriate behavior (behavior ‘demanded’ from them)” (2010:75). Esto sucede generalmente de manera inadvertida por el experimentador (aunque no siempre es este el caso como vamos a ver).



Recientemente, sin embargo, se ha defendido el uso de la DE como un instrumento más de la caja de herramientas del experimentador. Efectivamente la DE puede ser útil cuando es necesario introducir una presión normativa en el laboratorio que pretendidamente existe en la situación real (esto es, fuera del laboratorio) que se pretende modelar.

En este sentido es muy elocuente la forma en que Capraro y Rand defienden el uso de la DE en el sentido que proponemos: “injecting cues about which option is appropriate creates an undesirable alternative source of variation in behavior, which is particularly problematic when the choice suggested by the demand effect is the same as the one hypothesized to be caused by the experimental manipulation (in which case one cannot tell if the result is due to the manipulation or the demand effect). This is not, however, a problem for the experiments we present in this paper. On the contrary, what we are seeking to study is precisely the effect of giving information about the appropriateness of different options. The demand effect is our manipulation, rather than being a confound” (2018:100).

Como señalan los autores, el problema de la DE desaparece si lo que se estudia es precisamente el efecto de la DE. La utilización de la DE, en forma de normas, apunta al hecho de que hay situaciones que no pueden reducirse, sin más, a sus componentes estratégicos; otra forma de decir esto es que no existe una forma o lenguaje neutral de presentación que pueda expresar la situación real fuera del laboratorio.

Por otra parte, la introducción explícita de normas permite capturar las preferencias por “hacer lo correcto”. Experimentalmente la estrategia para estudiar este tipo de preferencias es sencilla: el experimentador compara dos tratamientos, con idénticas opciones, es decir, con una misma estructura económica de pagos, pero haciendo variar los nombres de las opciones, de manera tal que en un tratamiento la opción relevante sugiera qué es “lo correcto” hacer, mientras que en el otro tratamiento de control la opción sea descrita de manera neutral. Esta técnica consiste en una forma de *framing* específica que hace uso de la DE como un instrumento más que tiene el experimentador a su alcance para modelar la situación experimental.

Además de sanciones, se han empleado otras estrategias para señalar implícitamente la presencia de normas sociales. Algunas de esas estrategias son, por ejemplo, utilizar técnicas de marco (*framing*) y asociar las acciones contra-normativas a la generación de externalidades negativas. El inconveniente que presentan estas técnicas de señalización implícita de las normas es que la interpretación que realizan los participantes de estos factores puede no ser unívoca desde el punto de vista normativo. Por ejemplo, las externalidades negativas pueden ser más o menos toleradas o justificadas según sea la situación interpretada como una interacción de mercado o una negociación, en la que puede haber ganadores y perdedores. Además, existen actividades que no son contra-normativas (manejar en hora pico) pero que, sin embargo, pueden generar externalidades negativas (congestión de tráfico). Y como hemos visto anteriormente, los marcos no están exentos de ambigüedad y de problemas de demanda.

A continuación, voy a exponer un ejemplo para ilustrar de qué manera se pueden implementar en el laboratorio normas explícitas y cómo es posible hacer uso de la DE a favor del experimentador. En el próximo apartado brindaré una discusión teórica acerca de la conveniencia de hacerlo y cómo fundamentar su empleo.

Recientemente Senci et al. ([The influence of prescriptive norms and negative externalities on bribery decisions in the lab](#)) han incorporado la DE como un instrumento para preservar la similitud con la situación de interés en un juego de soborno. En el estudio se exploró de qué manera la interacción entre ciudadanos



y funcionarios se ve influenciada al introducir información normativa explícitamente en las instrucciones. Los participantes fueron divididos al azar en dos grupos (ciudadanos o funcionarios). El ciudadano tenía que realizar una tarea de esfuerzo real para adquirir el derecho a un beneficio monetario. Se implementaron dos tratamientos: en el tratamiento CON\_DERECHO se dispuso un umbral muy bajo para pasar la tarea, lo que aseguraba que virtualmente todos los ciudadanos asignados a ese tratamiento obtuvieran el derecho. En el tratamiento SIN\_DERECHO, por el contrario, se asignó un umbral virtualmente imposible de alcanzar, lo que aseguraba que los ciudadanos en ese tratamiento no obtuvieran el derecho.

Luego de conocer su estatus (esto es, si había adquirido o no el derecho en función de su desempeño en la tarea), debía realizar una transferencia a otro participante en el rol de funcionario público, simulando un gasto administrativo, pero además tenía la posibilidad, si así lo deseaba o creía conveniente, de ofrecer un “excedente” (una coima), esto es, más de lo estipulado en el experimento como pago administrativo. El funcionario, a su vez, tenía que decidir si le otorgaba o no el derecho al ciudadano. La manipulación experimental que nos interesa aquí consiste en la introducción en el tratamiento de una norma explícita que define los deberes y derechos de los roles, esto es, que el funcionario tenía el deber de otorgar el beneficio si el ciudadano había realizado la tarea correctamente, y negárselo en caso contrario.

El tratamiento se comparó con un control en el que ese requerimiento estaba ausente. En este tratamiento el funcionario debía decidir si otorgaba o no el derecho, en función del resultado de la tarea del ciudadano, pero en ningún momento se explicitaba cuál era su deber, ni tampoco el derecho del ciudadano. El argumento del experimento consiste en que los deberes y derechos de las partes que interactúan se encuentran bien definidos en el mundo real, y por lo tanto ese rasgo debe preservarse en el experimento para conservar el realismo psicológico. Tanto el ciudadano como el funcionario (al menos en las sociedades occidentales democráticamente desarrolladas) saben cuáles son las reglas que rigen sus interacciones y los funcionarios, generalmente, forman parte de una estructura estatal en la que existe una cadena jerárquica de mando que deben respetar. Por lo tanto, el requerimiento del experimentador es análogo al requerimiento de la autoridad en la vida real y, por lo tanto, se puede conjeturar que el respeto a la autoridad en el experimento (esto es, al experimentador) guarda similitud con el respeto a la autoridad tal como se da en la realidad entre los funcionarios y sus superiores.

La presión normativa que se introduce en el laboratorio por medio de un requerimiento explícito al funcionario en las instrucciones exhibe un paralelismo con la presión normativa ejercida por la normativa legal sobre los funcionarios. Esta es una forma de preservar el realismo psicológico de la tarea que realizan los sujetos. Por ejemplo, si el ciudadano no ha obtenido el derecho porque no ha pasado la tarea, entonces la obligación del funcionario es negarle el derecho. El experimento mostró que definir el estatus normativo de los participantes no es irrelevante, ya que los participantes en ambos roles fueron sensibles a la presencia de información sobre el estatus normativo del ciudadano y el deber del funcionario correspondiente. En resumen, se puede decir que el cumplimiento de la norma fue casi total, incluso teniendo en cuenta si habían recibido ofrecimientos de coima (95% de los funcionarios otorgaron el derecho en el tratamiento CON\_DERECHO; mientras que solo el 23% lo hizo en el tratamiento SIN\_DERECHO, esto es, cuando los ciudadanos no habían obtenido el derecho al beneficio). Estos resultados respaldan la idea de un mecanismo de conformidad normativo basado en el mero seguimiento de la norma prescriptiva, esto es, independientemente de preferencias económicas.

Veamos entonces cómo podría lucir un diseño destinado a estudiar la vinculación entre normas y castigo, elaborando sobre la base de este juego. Supongamos ahora que el comportamiento tanto de los ciudadanos como de los funcionarios pudiera ser “auditado” por un tercero que tuviera la posibilidad de



castigar (por ejemplo, implementando un descuento de dinero o descalificando del juego a quien no cumpliera las normas). En este caso, el tercero conocería perfectamente los deberes y derechos tanto de funcionarios como de ciudadanos. Es de esperar, entonces, que el auditor pueda utilizar el castigo para hacer cumplir la norma cada vez que ella sea violada. Nótese que si la norma fuera explicitada como proongo, el auditor tendría a su disposición una justificación del castigo que no admitiría ambigüedad.

Hay una línea de investigación sobre cooperación en la que se ha puesto de relieve que, además de preferencias económicas, los sujetos experimentales exhiben “preferencias morales”, que se manifiestan en una preferencia por “hacer lo correcto”. Por “preferencias morales” entiendo un concepto general que comprende modelos de preferencias sociales, modelos que incorporan el respeto a las normas en la función de utilidad de los agentes, y también, irreducible a estos últimos, la noción de una preferencia general por cumplir con las normas. De hecho, un grupo de estudios, entre los que podemos nombrar los recientes de Capraro y Rand ([Do the right thing](#)), muestra que las personas obtienen utilidad por “hacer lo correcto” y que la introducción de un lenguaje normativo que señalice la presencia de normas tiene un fuerte impacto sobre las decisiones en juegos de distribución y de otro tipo, del cual el ejemplo anterior es un claro representante. Este hecho parece señalar que hay situaciones que no pueden reducirse, sin más, a sus componentes estratégicos; otra forma de decir esto es que no existe una forma neutral de presentación, o que no existe una solución para juegos que sea independiente del lenguaje en el que se presenta el problema de decisión. La introducción de normas explícitas puede representar un buen mecanismo para capturar este tipo de preferencias, que no parece que puedan reducirse a un mero cálculo de costo-beneficio.

## 7. Un intento de solución: normas explícitas

En esta sección me propongo fundamentar teóricamente el intento de ofrecer una solución a lo que he caracterizado como el problema de la inferencia normativa del castigo fundamentada en el uso de información normativa. Plantearé una estrategia metodológico-conceptual con el objetivo de dar respuesta a los problemas de la irrelevancia normativa y de la percepción de legitimidad del castigo. Para ello, argumentaré acerca de la necesidad de introducir normas explícitas que estén asociadas a los roles que los participantes representan en los juegos experimentales. Esta estrategia permite, por un lado, ligar el castigo a normas relevantes con contenido que lo doten de un significado preciso y, por otro, desambiguar su implementación. Se trata, sin más, de asegurar la función expresiva del castigo, evitando aquellos problemas que pueden dificultarla o impedirla.

### 7.1 ¿Cómo colaboran las normas explícitas en la solución del problema de la irrelevancia normativa?

En la [Parte I](#) se señaló que uno de los problemas que deben afrontar los experimentos de castigo, si pretenden que sus resultados sean ecológicamente válidos, consiste en brindar una mayor relevancia normativa a la implementación del castigo. Allí habíamos dicho que la implementación del castigo como un descuento de dinero que un participante aplica sobre otro no es ecológicamente válido. Esto por dos razones principales: en primer lugar, porque en la vida real la forma apropiada de sancionar una transgresión normativa no consiste en la sanción monetaria (al menos no para una parte importante de transgresiones); y, en segundo lugar, porque el castigo basado en un descuento monetario tiene el efecto contraproducente de deslocalizar el modo de transacción pretendido por el experimentador y de esa manera modificar la naturaleza de la interacción. La irrelevancia normativa, producto de la utilización de sanciones monetarias en los juegos económicos, tiene el efecto contraproducente de que la sanción ya no transmite el sentido de que la acción no está permitida, sino lo contrario: la acción se permite a cambio



de un costo monetario. De esto se infiere que los sujetos experimentales pueden interpretar el castigo como un precio.

Sin embargo, el castigo es una construcción social compleja, no un evento natural, que solo puede interpretarse a la luz del telón de fondo constituido por las normas y roles sociales que le otorgan significado. El significado que una sociedad otorga al castigo está asociado a la asignación de roles socialmente precisados de acuerdo con ciertas reglas públicas que definen quien está legitimado para ejercer el castigo y quién no. Los roles sociales se encuentran asociados a deberes y derechos bien definidos, es decir, están reglamentados por normas que definen cómo se deben ejercer. Ahora bien, responder a dichos deberes y derechos no es posible si las normas que regulan su ejercicio no son lo suficientemente explícitas. O, dicho de otra manera, no es posible actuar conforme a la norma si no hay en principio acciones que sin ambigüedad alguna puedan clasificarse como transgresiones a la misma.

En el ejemplo tratado en la sección anterior, si al funcionario se le informa que debe otorgar el beneficio al ciudadano que ha obtenido el derecho, y no debe hacerlo en caso contrario, resulta evidente que la norma otorga un significado a la interacción que no admite ambigüedad. Por otro lado, la explicitación de las normas también fijaría con mayor claridad el contexto en el que interactúan los participantes. Esto es, sería sencillo reconocer el modo de transacción de acuerdo con las normas que lo regulan, y se podría advertir cualquier desviación respecto del modo de transacción prescripto en la situación. La asignación de derechos y deberes a través de una normativa explícita confiere contenido a la sanción: si el funcionario que otorga el beneficio al ciudadano que no ha adquirido el derecho es sancionado, la sanción adquiere un contenido que está presente en la norma y que se sustancia en la transgresión del deber del funcionario. En la ausencia de una norma explícita, la sanción corre el riesgo de parecer arbitraria.

## 7.2 ¿Cómo solucionan las normas el problema de la percepción de ilegitimidad del castigo?

Este punto se asocia al de la legitimidad del castigo tratado en la [Parte I](#) que precede a este artículo. Allí habíamos dicho que el problema central de la percepción de legitimidad lo constituía la ambigüedad normativa, y que esta podía tomar tres formas, según esté referida a: 1) la norma subyacente que motiva el castigo; 2) el conflicto entre la frecuencia estadística del castigo y la prescripción normativa y, por último, 3) la distinción entre comportamiento desviado e incorrecto. Voy a sostener que las tres formas de ambigüedad pueden subsanarse a través de la introducción de normas explícitas.

Respecto del primer punto, referido a la norma que motiva el castigo, la información normativa es útil para evitar la ambigüedad propia a las acciones. En la medida en que el consenso respecto de una norma es más unívoco, es probable que se logre un mayor nivel de castigo motivado por esta norma y, a su vez, mayores niveles de aceptación del mismo. Esto por dos motivos: primero, porque introducir información normativa puede ayudar a que la atribución de responsabilidad en un juego económico sea más unívoca; y segundo, porque quien ejerce el castigo tiene una razón a la que recurrir que fundamente la aplicación del mismo. De acuerdo con *Norms Focus Theory* propuesta por Cialdini y colaboradores, un factor crucial que determina la eficacia de la influencia de las normas en la conducta consiste en que la información normativa esté cognitivamente accesible ([A focus theory of normative conduct](#)). Una forma de lograr esto consiste en hacer disponible información empírica y/o prescriptiva acerca del comportamiento de los demás y las normas. Esa es la conclusión que extraen Krupka y Weber de una serie de experimentos de laboratorio destinados a poner a prueba dicha hipótesis: “Drawing subjects’ attention to the likely behavior of others or to others’ prescriptions for appropriate behavior both increase the frequency of pro-social behavior” (Krupka y Weber 2009:314).



Las normas prescriptivas respecto de un cierto tipo de acción tienen una doble función: por un lado, orientan la conducta del actor que realiza la acción y, por otro lado, señala que la transgresión de la norma es posible de ser sancionada. Ahora bien, como he señalado con anterioridad, la información empírica puede ser ambigua y conducir a resultados contradictorios: como es sabido, en ocasiones la información acerca de los niveles reales de cumplimiento de una norma (por ejemplo relativa al pago de impuestos) puede incrementar la conformidad corrigiendo una percepción errónea previa, pero también si la información empírica hace evidente que la norma es transgredida constantemente puede dar lugar a un efecto del tipo “¿qué más da?, ¿por qué habría de pagar mis impuestos si nadie los paga?”. La información normativa explícita, por el contrario, tiene la ventaja de dirigir la atención de los sujetos a la norma sin los consabidos problemas que puede generar la mera información descriptiva. He señalado anteriormente el rol que juegan las expectativas en el cumplimiento de las normas.

La explicitación de una norma en un juego de información completa puede hacer que la norma se transforme en un punto focal y, por ende, que coordine las expectativas de los sujetos. La coordinación en este caso sería doble: por un lado, las normas coordinarían las acciones y, por otro lado, coordinarían el castigo de las transgresiones. Esto es clave para entender de qué manera pueden conectarse el castigo y las normas. Precisamente las expectativas influyen sobre el cumplimiento de las normas, ya que es fuertemente contextual, esto es, depende de lo que el individuo cree que los demás hacen o consideran apropiado. La teoría de las normas de Bicchieri (*The grammar of norms*) ha prestado especial atención a este aspecto condicional de las normas. Según Bicchieri una norma social es “the mapping from context to interpretation, and thus to beliefs and expectations, that elicits a preference for conformity” (2006:57). Los individuos tienen preferencias condicionales, esto es, prefieren cumplir con las normas si creen que un número suficiente de miembros de su propio grupo cumple con las mismas normas (i.e., expectativas empíricas) y, además, si creen que los demás esperan de ellos que cumplan con las normas (expectativas normativas).

El segundo punto se refiere al fenómeno de que una cierta conducta, en este caso el castigo, puede ser objeto de legitimación a partir de su frecuencia estadística, es decir, puede ser fruto de un comportamiento imitativo, y no porque su ejercicio esté fundamentado en la transgresión de una norma. De la misma manera, puede darse el caso contrario, esto es, que a partir de la ausencia de castigo se infiera que este no es legítimo o que no existe un consenso respecto de que la transgresión deba castigarse. Como sabemos, una consideración importante respecto de la información empírica es que esta puede ser manipulada con objetivos egoístas, justamente porque no señala con precisión si una conducta es socialmente apropiada o no lo es. También puede suceder que la aplicación de las sanciones suscite desconfianza porque se perciba que el castigo beneficia de algún modo al castigador, cuya motivación no estaría dada por razones normativas, sino por intereses egoístas. Este tipo de inferencias pueden corromper el cumplimiento de normas tal como ha mostrado experimentalmente Erte Xiao ([Profit-seeking punishment corrupts norm obedience](#)).

La información normativa permitiría desambiguar este tipo de motivaciones y sustraer de la consideración de los castigados motivos egoístas por parte de los encargados de implementar el castigo. Este aspecto es particularmente importante cuando se implementan formas de castigo centralizado, esto es, que dependen una sola institución o individuo. En este caso, los problemas de legitimación pueden ser acuciantes, ya que el castigo puede ser fácilmente corrompido.

El tercer punto quizás sea el más sencillo. En efecto, la incorporación de una norma explícita, si es lo suficientemente transparente, borra la distinción entre comportamiento desviado e incorrecto. En la Parte



l habíamos dicho que un comportamiento puede ser percibido como desviado, sin por ello que se lo perciba como incorrecto. Esta diferencia atañe a la distinción entre información descriptiva e información prescriptiva. Mientras que la primera puede ser objeto de ambigüedad, no así la segunda. Recordemos que el aprendizaje de la norma descriptiva solo puede ser el resultado de la observación del comportamiento de los demás individuos y comporta cierto grado de ambigüedad en la medida en que el conocimiento que obtenemos del comportamiento de los demás es parcial. En la medida en que la motivación no es directamente observable, los individuos pueden formarse ideas equivocadas respecto de la motivación por la cual se aplica sanciones. Es probable que la justificación del castigo discurra con mayor facilidad cuando se trata de comportamiento incorrectos que cuando son solo desviados.

La incorporación de información normativa en un juego que presumiblemente represente de manera más acabada las relaciones normativas que se dan en la vida cotidiana es una manera sencilla de reducir la ambigüedad normativa. Los estudios empíricos muestran que la introducción de información normativa aumenta la conformidad. Posiblemente una de las vías por medio de las cuales esto sucede es porque facilita la convergencia o coordinación de creencias respecto de las normas. Posiblemente el efecto se deba a que las normas explícitas dejan poco margen a la ambigüedad y por lo tanto generan que el comportamiento apropiado sea conocimiento común para todos, es decir, las normas se transforman en puntos focales, que pueden funcionar como soluciones a juegos de coordinación. Ello, a su vez, ayudaría a incrementar el sentido de la responsabilidad (accountability) evitando o disminuyendo la ceguera o desvanecimiento ético (ethical fading), es decir, “the process by which the moral colors of an ethical decision fade into bleached hues that are void of moral implications. If we are correct, educating individuals on moral principles is only useful when individuals perceive that the decision has ethical coloration, but useless when ethical fading has occurred or may occur” (Tenbrunsel y Messick 2004:224).

Si logran satisfactoriamente neutralizar los componentes morales de una decisión, los individuos pueden comportarse de manera egoísta, y aun así mantener la convicción de que han obrado de forma correcta. Resultaría en principio mucho más difícil para un individuo justificar una transgresión normativa aduciendo desconocimiento o interpretando la información disponible de manera sesgada. Esta línea de razonamiento viene precedida también por evidencia empírica que sugiere que cuando la información normativa es ambigua, los individuos tienden a favorecer una interpretación sesgada de la situación en su propio beneficio, esto es, de manera egoísta. De acuerdo con Dana, Weber y Kuang “People may feel compelled to give in some situations -even though they prefer the own-payoff-maximizing outcome- because they do not want to appear selfish, either to themselves or to others. Thus, the underlying motivation driving much fair behavior might be self-interest, coupled with a desire to maintain the illusion of not being selfish” (2007:68). Los autores tienden a conceder que este fenómeno, que se conoce como sesgo egoísta (self serving bias), puede ser favorecido por la ambigüedad normativa, en la medida en que facilita la justificación de la conducta egoísta. Por otra parte, es conocido que la ambigüedad normativa facilita la manipulación estratégica del propio conocimiento: a sabiendas, los individuos pueden evitar verse expuestos a información que, de conocerla, los obligaría a tomar decisiones en contra de sus intereses egoístas; o pueden exponerse selectivamente a información con el fin de poder alegar desconocimiento.

Además de normas explícitas, una forma suplementaria que tiene el potencial de asegurar la función expresiva de señalización del castigo consiste en la utilización de información normativa (que por otra parte normalmente intervienen en las situaciones reales) como mensajes, descargos o advertencias, es decir, formas alternativas por medio de las cuales se puede vehicular verbalmente una disconformidad o malestar. En línea con esto Andrighetto *et al.* realizaron un experimento en el que compararon un



tratamiento en el que los sujetos podían sólo ejercer un castigo material (tratamiento “castigo”) con uno en el que además podían enviar un mensaje (tratamiento “sanción”): “En el tratamiento de sanción, los sujetos cumplen inmediatamente con la prescripción y el posible uso del castigo solo mantiene el nivel de contribución alcanzado. En contraste, en el tratamiento de castigo, en el que la información sobre las normas solo se puede inferir a partir del costo material recibido, el nivel de cooperación alcanzado es sustancialmente más bajo y los costos para lograrlo son mayores que en el tratamiento de sanción. *Cuando se modela solo en términos materiales, el castigo es poco efectivo para ayudar a los sujetos a descubrir la norma. Lejos de coordinar, los sujetos proceden por separado por prueba y error*” (Andrighetto *et.al.* 2013:4, énfasis agregado).

Las comunidades no implementan el castigo solo de manera material, sino que lo hacen a través de formas de recompensas o sanciones no tangibles. De hecho, el castigo material ejercido por los propios ciudadanos que no portan un rol como funcionarios resulta marginal. La aprobación o desaprobación social en la forma de sanciones verbales puede ser una fuerza reguladora de las conductas, aun en ausencia de castigos materiales. Las personas no solo son sensibles a los castigos materiales, sino que también son sensibles a la opinión de los otros. En definitiva, el trabajo conjunto de imposiciones materiales y sanciones verbales es posible que facilite la función expresiva de las sanciones comunicando de manera fehaciente las normas que lo legitiman.

## Conclusión

En el trabajo expuse detalladamente lo que he llamado el problema de la hipótesis de la inferencia normativa del castigo. Dicha inferencia se compone de dos partes: por un lado, la (ir)relevancia normativa del castigo y, por otro, la (i)legitimidad del mismo. Sin atención a estos aspectos, los experimentos de castigo corren el peligro, y este ha sido mi argumento a lo largo del artículo, que sus resultados no puedan ser extrapolados a situaciones de la realidad debilitando considerablemente su validez ecológica.

La conclusión a la que he arribado concuerda con la opinión de otros autores y apunta a que la eficacia del castigo está estrechamente vinculada a la habilidad que tiene para expresar un contenido normativo. Las formas en las que se implemente el castigo en juegos experimentales deben tratar de hacer evidentes para los participantes que las sanciones se asocian inequívocamente a la transgresión de normas y evitar que tengan lugar sanciones que puedan percibirse como arbitrarias. Asimismo, he propuesto una forma de evitar la ambigüedad normativa, y de reforzar la función expresiva del castigo, que consiste en la introducción de normas explícitas. Dicha solución se apoya en estudios recientes que sugieren que la introducción de información normativa aumenta la pro-socialidad y contribuye a que la situación modelada en el laboratorio sea normativamente menos ambigua.

## Bibliografía

- Andrighetto, G. *et.al.* (2013). Punish and voice: punishment enhances cooperation when combined with norm-signalling. *PLoS ONE* 8:6. <https://doi.org/10.1371/journal.pone.0064941>
- Bicchieri, C. (2006). *The grammar of society*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511616037>
- Capraro, V., Rand, D. (2018). Do the right thing: experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality. *Judgment and Decision Making* 13(1): 99-111. <http://journal.sjdm.org/17/171107/jdm171107.pdf>



- Dana, J., Weber, R.A., Kuang, J.X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory* 33: 67-80. <https://doi.org/10.1007/s00199-006-0153-z>
- Guala, F. (2012). Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences* 35(1): 1-15. <https://doi.org/10.1017/s0140525x11000069>
- Krupka, E., Weber, R. (2009). The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology* 30(3): 307-320. <https://doi.org/10.1016/j.joep.2008.11.005>
- Tenbrunsel, A.E., Messick, D.M. (2004). Ethical fading: the role of self-deception in unethical behavior. *Social Justice Research* 17(2): 223-236. <https://doi.org/10.1023/B:SORE.0000027411.35832.53>
- Zizzo, D. (2010). Experimenter demand effects in economic experiments. *Experimental Economics* 13(1): 75-98. <https://doi.org/10.1007/s10683-009-9230-z>

Recibido el 26 Jun 2019

Aceptado el 13 Sep 2019