

EL MENSAJE DE LAS INSTITUCIONES*

JULIA BARRAGÁN

*Centro de Investigaciones Posdoctorales
Universidad Central de Venezuela*

ABSTRACT

The trust that is a fundamental factor to solve the problem of harmonization of the individual and collective rationality derives from the “common knowledge” which in turn originates in “the public events.” This fact indicates that the normative authority is in a privileged position to generate trust. The means which the normative authority counts on to fulfil this objective are that of the design of the institutions and incentives. In order to generate trust these designs should possess the necessary features to avoid that the practice of altruism makes vulnerable the altruistic players

Keywords: prisoner[s]’ dilemma, games theory, altruism, self-love, cooperation, institution design.

RESUMEN

La confianza que es un factor fundamental para resolver el problema de armonización de la racionalidad individual y la colectiva deriva del “conocimiento común” el cual a su vez se origina en “los eventos públicos”. Este hecho indica que la autoridad normativa está en una posición privilegiada para generar confianza. El medio con el que la autoridad normativa cuenta para cumplir con este objetivo es el del diseño de las instituciones e incentivos, los cuales deben poseer los rasgos necesarios para evitar que las conductas altruistas hagan vulnerables quienes las practican.

Palabras clave: dilema de los prisioneros, teoría de juegos, altruismo, egoísmo, cooperación, diseño institucional.

1. INTRODUCCIÓN

Las decisiones interactivas que responden al modelo del Dilema de los Prisioneros¹ son extremadamente frecuentes, y es muy conocido el enorme

* Aceptación: II / 2002

1 En Luce D. y Raiffa H. *Games and Decisions*, John Willey & Sons, New York, 1957, se encuentra una exposición tan clara como conceptualmente rigurosa del Dilema de los Prisioneros. Originalmente

poder erosionante que ellas ejercen sobre el tejido social, y muy en especial sobre los llamados bienes públicos y bienes comunes².

No importa cuánto nos esforcemos teóricamente en demostrar la conveniencia de asumir comportamientos cooperativos, y la clara ventaja que en términos de utilidad social ofrecen; tampoco cuenta el hecho cierto de que la gran mayoría de nosotros deseamos sinceramente vivir en un mundo en el que prevalezca la confianza recíproca. Las decisiones que en procura de una maximización de las utilidades individuales esperadas llevan el equilibrio a resultados sociales sub óptimos aparecen constantemente, y retornan de manera infinita bajo formas sorprendentemente variadas. Hoy se trata de las pequeñas autoexcepciones que no llegan a constituir un fenómeno social preocupante; mañana se trata de actitudes que rozan (solo rozan) los límites de la legalidad o de la i-legalidad; y en otro momento son acciones que en el nivel individual no vacilaríamos en calificar de “optimizadoras” pero que a la postre conspiran contra la confianza pública.

Mucho se ha discutido sobre las posibilidades que tienen los actores sociales para superar mediante procesos evolutivos la encerrona a que parece condenarlos su propia condición de sujetos racionales³, pero tal como se presentan las cosas hay evidencias de que las dificultades rebasan en mucho la voluntad de acción de estos sujetos. A pesar de la consistencia de los argumentos que se esgrimen, y de la indiscutible deseabilidad moral de lograr una confianza recíproca, al momento en que se construyen las decisiones interactivas en condiciones de incertidumbre que caracterizan el escenario social, los jugadores no parecen encontrar ni en los argumentos ni en la deseabilidad moral razones suficientes para actuar cooperativamente.

Pareciera que un enorme peso impide el desarrollo de las llamadas conductas cooperativas, y aun el uso de expresiones tales como “egoísmo” versus “altruismo”, que implican juicios de rechazo o aceptación social, no logran modificar la tendencia que hace que la estrategia dominante sea la de no-cooperación. Los ojos del análisis teórico se vuelven frecuente-

formulado por Merrill Flood y Melvin Dresher, y formalizado por A. W. Tucker, el Dilema de los Prisioneros ha merecido una gran atención por parte de los estudiosos de las Ciencias Políticas, por considerarlo un modelo extremadamente plausible de las decisiones interactivas en condiciones de incertidumbre que caracterizan la vida social. Asimismo se considera que el mismo expresa con fidelidad las dificultades para articular la racionalidad individual con la colectiva.

- 2 Los demoledores efectos que los comportamientos no-cooperativos generalizados producen sobre los llamados bienes comunes están expuestos en Hardin “The tragedy of the commons”, *Science*, 162, 1968, mientras que los efectos igualmente devastadores de tales comportamientos sobre los bienes públicos se encuentra analizado en Barragán J. “El Costo Social de las Autoexcepciones”, *Τέλος*, Vol VI, número 1, 1997
- 3 Para una completa y actualizada exposición de este enfoque puede consultarse Peyton Young H. *Individual Strategy and Social Structure: an evolutionary theory of institutions*, Princeton University Press, 1998.

mente hacia los actores sociales, y en ellos se buscan diversas razones subjetivas para explicar el origen de dichos comportamientos. Estas razones subjetivas estarían a su vez determinadas por la fragilización de las relaciones interpersonales que se ha producido con motivo de la complejización del entramado social (transformación de la comunidad en sociedad). Desde esta perspectiva, parece normal propugnar un retorno a los valores comunitarios, los cuales podrían a su vez incidir positivamente no solo en la generación de conductas altruistas sino también en el fortalecimiento del altruismo como patrón generalizado de conducta⁴.

Este marcado interés en el comportamiento de los jugadores sociales ha convertido a los mismo en el punto focal de la atención, y contemporáneamente ha permitido que una figura crucial en la definición del juego permanezca en la semipenumbra y aun en absoluta sombra; se trata de la figura del "fiscal", el cual es más responsable de lo que a primera vista parece en el hecho irrefutable de que la estrategia dominante sea no-cooperativa. Esta oculta y sigilosa figura es, sin embargo, la que ha construido el arreglo de restricciones que constituyen el juego, y la que a su vez ha establecido el cuerpo de incentivos para los comportamientos; en otras palabras es el diseñador de la institución conocida como el Dilema de los Prisioneros. En tal condición parece razonable tomarlo en seria consideración, a fin de comprender mejor por una parte las razones de las opciones estratégicas de los actores sociales, y por la otra, someter a evaluación el perfil normativo de la figura del fiscal.

2. LA ACCIÓN DEL FISCAL Y SU RESPONSABILIDAD EN LA CONSTRUCCIÓN DE ESPACIOS MINADOS

La renuencia de los actores sociales para converger en las socialmente óptimas soluciones cooperativas no debería ser atribuida simplemente a deficiencias morales de los actores o a un comportamiento generalizado de desinterés por la cosa pública derivado de relaciones intersubjetivas débiles y mediatizadas, ya que tal atribución puede conducirnos a un callejón sin salida desde el punto de vista teórico. Contrariamente, si aceptamos el carácter autónomo de los sujetos y su calidad de agentes que privilegian las soluciones que satisfacen sus preferencias, dejamos abierta la posibilidad de obtener una mejor comprensión de las razones que cada actor social tiene para optar por cursos de acción que a la postre conducen a soluciones sociales subóptimas.

4 Entre los enfoques que sostienen que los valores comunitarios constituyen un elemento clave para la reinstalación de la confianza necesaria para el florecimiento del altruismo, es particularmente interesante el presentado por Kunz G. en el Capítulo Siete de su trabajo *The paradox of power and weakness*, State University of New York Press, 1998, en el que analiza el llamado "The power of Community".

Por otra parte, si consideramos al fiscal como un diseñador de la institución llamada Dilema de los Prisioneros, no solo podremos comprender aun mejor aquellas razones de los actores sociales, sino que también estaremos en condiciones de ahondar en el significado del mensaje normativo que se encuentra implícito en el diseño.

En su diseño, el fiscal parte de un supuesto acerca de la calidad de los agentes sociales; los supone actores racionales, es decir decisores que son capaces de ordenar sus preferencias de un modo consistente, al tiempo que orientan sus elecciones en búsqueda de maximizar su utilidad esperada. Ya que los supuestos no pueden ser considerados ni verdaderos ni falsos, el único terreno en el que podríamos entrar en polémica con el fiscal es el relativo a la plausibilidad de tal supuesto; en este terreno hay que acordar que el supuesto asumido por el fiscal captura de un modo acertado los rasgos mínimos del comportamiento de los actores en el juego social. Tales rasgos mínimos no son otros que el de la capacidad de los actores para efectuar comparaciones entre las utilidades esperadas de un curso de acción y de su alternativa, y el orientar su elección sobre la base de las expectativas de maximización de su sistema de preferencias.

Bajo el supuesto de ese decisor racional, el fiscal efectúa un diseño de incentivos que de alguna manera es un poderoso orientador de las conductas de los actores sociales. Este cuerpo de incentivos queda expresado en los pagos que se establecen para cada una de las interacciones entre los sujetos. Así, considerando las utilidades del sujeto A se tendrían las siguientes alternativas:

- | | | | |
|---|------------------------|---|---|
| a | A actúa egoístamente | y | los otros actores actúan altruístamente |
| b | A actúa altruístamente | y | los otros actores actúan altruístamente |
| c | A actúa egoístamente | y | los otros actores actúan egoístamente |
| d | A actúa altruístamente | y | los otros actores actúan egoístamente |

El sistema de pagos que el fiscal establece para cada una de estas alternativas hace que las utilidades de A disminuyan desde a hasta d, ya que el diseño propone la siguiente jerarquía: $a > b > c > d$; es decir que la conducta que mejor recompensa según el sistema de incentivos establecido por el fiscal es la de actuar egoístamente en espera de que los demás lo hagan de manera altruista. No resulta entonces extraño ni incoherente el que todos los actores sociales escojan esta alternativa, ya que de alguna manera el sistema de incentivos diseñado los conduce de manera necesaria a ello. Indudablemente y más allá del deseo de todos, al generalizarse tal conducta egoísta, la estructura de toma de decisiones interactivas tiende hacia la solución c cuyas utilidades son bajas. Cabría preguntarse por qué se presenta esta situación que es indeseable tanto del punto de vista individual como colectivo, cuando parecería más razonable tender a la solución b que representa simultáneamente un mundo de buenas

utilidades individuales y de confianza recíproca en el que con sobradas razones todos deseáramos habitar.

La respuesta a tal pregunta se asocia a los terrenos que hay que atravesar para llegar a la solución b; en el caso del actor A, debería él mismo decidir que aun cuando las conductas altruistas en el marco del sistema de incentivos establecido por el fiscal tienen una remuneración menor que las egoístas en cualquier escenario posible, él de todos modos las asume contraviniendo la regla de la maximización de sus preferencias. Pero esto implica que A transite por un espacio minado, en el cual con una alta probabilidad puede quedar aprisionado en la peor situación posible que es la de la solución d.

Los peligrosos terrenos que A debe atravesar han sido sin duda minados por una decisión del fiscal tomada en el nivel de metajuego, según la cual se asigna a los actores un sistema de incentivos que hace supererogatorio el altruismo, y extremadamente conveniente el egoísmo. Ante estos incentivos, no es para nada extraño que A decida seguir la regla de la maximización en lugar de arriesgarse a sufrir pérdidas en sus utilidades, pérdidas que a la postre solo beneficiarían a los demás actores sociales incrementando sus ganancias. En este escenario definido por el fiscal la desconfianza recíproca crece, y no hay ninguna actitud individual que sea capaz de establecer las condiciones mínimas bajo las cuales resulte razonable confiar en el otro.

Adicionalmente, a través del sistema de sus incentivos, el fiscal produce un mensaje normativo que es recibido por los actores; en efecto, si alguien como el fiscal establece un sistema de recompensas que premia abiertamente el egoísmo, es de suponer que ese alguien (autoridad normativa) cree en la superioridad moral de estas conductas, ya que en caso contrario, su diseño sería inconsistente con sus propias premisas normativas. Siendo así las cosas, los comportamientos egoístas reciben un segundo apuntalamiento, ya que no solo son ventajosos como estrategias que maximizan la utilidad esperada, sino que además constituyen la base normativa desde la cual se ha diseñado la institución y su cuerpo de incentivos. Cuando se establece que para alcanzar una solución óptima desde el punto de vista social como es la solución b, es obligatorio atravesar los espacios minados que constituyen la única ruta hacia esa solución, se establece implícitamente la veda de ese camino, o en su defecto la exigen a los actores de asumir comportamientos supererogatorios que no guardan consistencia ni siquiera con formas atenuadas de racionalidad.

Esta característica de inductora de los comportamientos que poseen las decisiones normativas del fiscal, ha merecido una inteligente consideración por parte de Norbert Hoerster⁵, quien se pregunta sobre cómo serían

5 Ver Hoerster, N., *Problemas de Ética Normativa* Editorial Alfa, Buenos Aires, 1975.

los resultados en una estructura (institución) análoga a la del Dilema de los Prisioneros, pero sometida a un diseño de incentivos diferente al del establecido por el fiscal. En busca de una respuesta para su pregunta, elabora un modelo alternativo.

Se trata de un rico anciano que desea que sus dos hijos se reconcilien tras largos años de distanciamiento, para lo cual los invita a la celebración de su cumpleaños, y les comunica que después del mismo redactará su testamento. Si ambos vienen a la fiesta, la herencia se repartirá en partes iguales. Si no viene ninguno de los dos cada uno heredará sólo $1/5$ (mientras los otros $3/5$ los heredará el ama de llaves). Si viene solo uno, el que venga será desheredado, y el que no vino heredará $3/5$ (los restantes $2/5$ irán al ama de llaves). La relación entre los pagos, que constituye el diseño de incentivos presentado por el anciano es la siguiente: $3/5 > 1/2 > 1/5 > 0$; según esta relación, la solución más conveniente sigue siendo la egoísta (no asistir mientras el otro asiste), pero de la exposición del modelo se deriva que el anciano otorga una apreciación valorativa positiva al hecho de que los dos hijos se reencuentren. En virtud de esta premisa valorativa, el anciano procura hacer que cada hijo tenga interés en la concurrencia del otro a la reunión, ya que en caso contrario se vería desheredado; este interés los llevará a establecer contactos, con lo que en principio el reencuentro comienza a hacerse efectivo, y se abre un espacio de interacción que es capaz de generar confianza entre los hijos. Aun si este contacto es insuficiente para generar un clima de "juego limpio" y uno de los hijos piensa en promover la asistencia del otro a la reunión para él dejar de concurrir a último momento y de ese modo ganar $3/5$, tanto la interacción promovida entre los hijos, como su propia racionalidad les llevará a acudir juntos a fin de garantizarse no ser traicionados a último momento con la no asistencia del otro.

El ejemplo de Hoerster aunque conserva la estructura y el sistema de incentivos del Dilema de los Prisioneros incorpora por una parte un relajamiento de la condición de incomunicación establecida por el fiscal, lo cual funciona como espacio generador de confianza, y por la otra estimula el empleo de una racionalidad basada en estrategias conjuntas (acudir juntos). Es indudable que si comparamos ambos decisores normativos (el fiscal y el anciano) podemos ver claramente el modo en que operan los incentivos como mecanismos de estímulos de los comportamientos, y la diversidad de mensajes normativos que de tales incentivos se derivan. Mientras la acción del fiscal está dirigida a instalar minas en el espacio que debe ser atravesado para lograr una solución socialmente deseable, la actitud del anciano procura remover esos obstáculos para hacer posible el clima de confianza recíproca y de respeto por las utilidades individuales que caracterizan las soluciones colectivas estables. Mientras en el primer caso crecen las desconfianza a ser explotado y de manera paralela la voracidad por explotar a otros en búsqueda de una

maximización irrestricta, en el segundo se abonan terrenos para que se instale exitosamente el germen de la confianza, si no hacia el otro actor, al menos si en la calidad valorativa de las reglas que regulan la relación.

3. TRES ENFOQUES CONCEPTUALES RELACIONADOS CON EL TEMA DE LA GENERACIÓN DE CONFIANZA

Actualmente uno de los temas teóricos más relevantes en la ética pública es el de la ausencia de confianza entre los actores sociales, situación que lleva a interacciones como las modeladas en el Dilema de los Prisioneros que producen resultados sociales subóptimos. Adicionalmente, desde el punto de vista práctico, este tipo de comportamientos tiene una repercusión innegable en la calidad de vida de los grupos y en la sustentabilidad de sus bienes públicos y comunes. Ambas razones (teóricas y prácticas) han sido un poderoso incentivo para que importantes autores coloquen su mira en los puntos más relevantes de este problema. Seleccionamos para su análisis a tres autores que por la calidad de sus razonamientos y su carácter influyente han contribuido a construir vías de acceso a los puntos más intrincado del asunto, y al mismo tiempo han logrado establecer algunos avances en la búsqueda de soluciones. Son ellos Robert Axelrod, John Harsanyi y John Nash.

En un escenario en el que el mundo deseaba ardientemente resolver problemas de disgregación social y falta creciente de interés en la cosa pública, el trabajo seminal de Robert Axelrod sobre una eventual evolución y triunfo de la cooperación⁶ produjo un notable impacto teórico y una ráfaga de optimismo en el mundo práctico. La tesis de Axelrod establecía que si el Dilema de los Prisioneros es jugado repetidamente (como de hecho lo es en el juego social), la opción del altruismo resulta estratégicamente conveniente, y lleva finalmente a obtener mejores pagos que con la opción egoísta. Lo interesante para nuestro trabajo es que Axelrod pretende mantener inalterada la estructura de los incentivos establecida por el modelo original, aunque relaja implícitamente la condición de incomunicación, ya que las sucesivas interacciones operan como un mecanismo informacional que hace saber a los actores acerca de los patrones de comportamiento de los demás miembros. En efecto, Axelrod considera que en los casos concretos que se dan la realidad social, el hecho de que los jugadores deban encontrarse reiteradamente, permite generar espacios de confianza recíproca donde la cooperación puede desarrollarse de manera espontánea. Pero la tesis va algo más allá al afirmar que tales comportamientos son “ganadores” en términos de utilidades.

6 El caso fue inicialmente planteado en Axelrod, R., *The Evolution of Cooperation*, Basic Books, New York, 1984, y en una serie de desarrollos posteriores realizados por el mismo autor, entre los que se destaca *The Complexity of Cooperation*, Princeton University Press, 1997.

A fin de probar esta última afirmación Axelrod realiza una serie de simulaciones en las que se enfrentaron diferentes estrategias que pudieran calificarse de manera general como “altruistas” o “egoístas”, cuya formalización en sendos programas de computación fueron provistas por estudiosos del campo de la teoría de los juegos y de la programación. Resultó ganadora la estrategia conocida como TIT FOR TAT, que fuera programada por un gran conocedor del Dilema de los Prisioneros, el Profesor Anatol Rappaport. Esta estrategia consiste en comenzar cooperando (comportamiento altruista) con la intención de generar un espacio de confianza que estimule la generalización de tales comportamientos; a partir de ese momento, el actor social hará aquello que le hagan: es decir será altruista con los altruistas y será egoísta con los egoístas.

Efectivamente esta simple estrategia no solo resulta la más conveniente en el nivel individual, sino que también logra que los comportamientos altruistas se generalicen, lo cual implica la obtención de un activo social indiscutible. Sin embargo la pretensión de Axelrod de haber mantenido la estructura original de lo que hemos llamado la institución del Dilema de los Prisioneros, como así también su cuerpo de incentivos no coincide con lo realizado. Desde nuestro punto de vista Axelrod, tan preocupado por el carácter erosionante de los comportamientos egoístas como lo estaba el anciano de Hoerster, estructura una nueva institución y diseña un nuevo sistema de incentivos, mediante los cuales procura apuntalar los comportamientos “altruistas” y la confianza en la estabilidad de las reglas del juego. Sin lugar a dudas con este trabajo Axelrod logra poner en evidencia los elementos que conspiran contra las estabilidad de las soluciones cooperativas, y ensaya arreglos institucionales que tiendan a corregirlos.

Entre esos elementos hay uno de orden general que es el de la incomunicación, para cuya corrección Axelrod considera suficiente con garantizar que el juego sea iterativo, con lo cual los sucesivos comportamientos hacen las veces de piezas informacionales que ayudan a prever futuras conductas, y sin duda contribuyen a definir espacios de confianza basados en el conocimiento recíproco. Sin embargo, para que se produzca el surgimiento de estos espacios es necesario que el diseñador de la institución garantice por una parte la estabilidad de las reglas, y por la otra la acumulación de las ganancias en el mismo escenario del juego; esto constituye el sistema de incentivos que permite el nacimiento de la confianza, la cual entonces no está primariamente motivada por las conductas de los otros jugadores sino por la existencia de un sistema institucional “confiable”.

Otro de los elementos que conspiran contra la generación de confianza es el del manejo diferencial de la información. En la primera ronda de la simulación Axelrod advirtió que si alguien conocía el momento en el que el juego finalizaba, podía jugar egoístamente en la jugada final, ya que

el contrincante no tendría oportunidad de retaliar (que es el recurso de control sobre el comportamiento del oponente). Para impedir esto, en la segunda ronda incorpora una restricción institucional que establece que el juego finalizará aleatoriamente, con lo que pretende impedir que el manejo de una información privilegiada se constituya en un factor que conspire contra la igualdad estratégica de los actores.

En rigor, el objetivo central de Axelrod en su trabajo es el de demostrar que sin la intervención de factores externos al juego social mismo, es posible alcanzar soluciones cooperativas generales estables. En este aspecto hay que acordar que no logró éxito; pero el propio fracaso de su intento ha contribuido de manera importante a hacer evidente por una parte los aspectos que se deben tomar en consideración en un diseño institucional que pretenda alcanzar la meta de un equilibrio cooperativo, y por la otra, el enorme poder constructivo que puede alcanzar la acción individual en un marco institucional apropiado. En relación con el diseño que hizo posible lo que Axelrod llama “el triunfo de la cooperación”, queda completamente claro el papel que juega la estabilidad de las reglas, el carácter acumulativo de las ganancias y el manejo democrático de la información, factores estos sustanciales en la toma de decisiones interactivas, y que constituyen un buen terreno donde el nacimiento de la confianza recíproca sea posible.

El segundo enfoque que vamos a considerar sobre el tema de cómo generar confianza en el escenario del juego social viene de la mano de John Harsanyi⁷. Para Harsanyi la teoría de la decisión reconoce la existencia de tres planos: por una parte el que concierne a las decisiones individuales que son el objeto de la Teoría de la Utilidad, en segundo lugar el de las decisiones interactivas dentro de un sistema de reglas que son tratadas por la Teoría de los Juegos, y por último el de las decisiones cuya finalidad es la de establecer las reglas del juego, que caen en el terreno de la Ética. Resulta claro que las decisiones del primer tipo afectan de manera directa y exclusiva a quien las toma, y en consecuencia no llegan a constituir un tema central desde el punto de vista social. En referencia con el segundo tipo, éstas son las decisiones que toman los actores sociales en el marco institucional en el que interactúan, y desde luego poseen importancia social ya que siempre es deseable que tales decisiones se ajusten a las reglas establecidas, y contribuyan al mantenimiento y equilibrio del juego; éstas son las decisiones que toman cada uno de los prisioneros involucrados en el marco estructural del ya considerado dilema. Finalmente, el tercer tipo de decisiones son aquéllas que un decisor investido de autoridad

7 Puede verse Harsanyi, J., *Rational Behaviour and Bargaining Equilibrium in Games in Social Situations*, Cambridge University Press, 1977 y “Razón, Moralidad y Teoría Utilitarista” en *Decisiones Normativas- Ensayos en Homenaje a Julia Barragán*, Ed. Sentido, Caracas, 1999.

toma para constituir las reglas que regularán las decisiones que se tomen en el nivel dos; esas decisiones son de una inmensa importancia social, ya que de alguna manera serán las responsables de las opciones de acción que se ofrezcan en el juego social. Estas decisiones que en otros trabajos hemos llamado normativas, tienen la característica especial de que obligan no solo a quien las toma sino también a aquéllos que en virtud del principio de subordinación deben acatarlas aunque no hayan intervenido en su formación⁸. A través de las mismas la autoridad está en condiciones de modificar la estructura de los pagos, haciendo que converger en la casilla cooperativa no constituya un acto supererogatorio, ni haga vulnerable a quien elige cooperar; es más, a través de las decisiones normativas la autoridad puede llegar a disolver el dilema.

El problema del diseño de las instituciones y determinación de los incentivos se ubica para Harsanyi en el tercer tipo, y el decisor está obligado a respetar ciertas reglas en la construcción de tales decisiones. En primer lugar, y siendo fiel a su tradición utilitarista, Harsanyi considera que deben ser tomadas en consideración las preferencias de los sujetos a los cuales afectará la institución; en segundo lugar, y como un modo de garantizar ecuanimidad en el tratamiento de dichas preferencias, se debe asignar a todas ellas equiprobabilidad en la función de bienestar colectivo; y como última condición, el decisor que construye la institución y los incentivos asociados debe ser un decisor bayesiano en el sentido de que las probabilidades subjetivas asignadas a un hecho son siempre condicionales y sujetas a corrección a la luz de toda nueva información.

Como puede observarse, el camino de Harsanyi se orienta básicamente a establecer al decisor normativo algunas restricciones dirigidas a estimular en los actores sociales la confianza en la ecuanimidad de las reglas. Esto a la larga produciría conductas altruistas en virtud de que tales actores verían suficientemente reflejadas sus propias preferencias en las reglas; en teoría, las condiciones establecidas deberían ser una garantía en este sentido. Sin embargo, el escenario del mundo concreto no es tan sereno como parece, ya que al asignar equiprobabilidad a todas las preferencias, en una sociedad con un alto número de preferencias egoístas, podríamos encontrarnos con soluciones éticamente inaceptables; esto sin contar con las preferencias marcadamente perversas o destructivas. Parece un tanto insostenible asignar la misma probabilidad a las preferencias de un sádico, que obtiene la máxima satisfacción con el sufrimiento ajeno, a las de una madre cuyas primeras preferencias se refieren a un buen presupuesto para la educación preescolar. Este ejemplo muestra con claridad que quien diseña una institución y un sistema de incentivos debe

8 Para una exposición más detallada de este tipo de decisiones y su influencia en la vida social me permito sugerir la lectura de Barragán J., "La paradoja de la autoridad y los problemas de formación de capital social" en *Isonomía*, 16, 2002.

necesariamente efectuar ponderaciones y exclusiones de algunas preferencias. Para Harsanyi el problema puede resolverse sin mayores dificultades en un gran número de casos, ya que en ellos toda persona razonable coincidiría en ciertas exclusiones y ponderaciones. No creemos que esto sea tan simple, y aun en casos aparentemente improblemáticos surgen matices y disonancias que llegan a constituir escollos insalvables a la hora de incorporar o desincorporar una preferencia de la función de bienestar colectivo. Harsanyi logra establecer con buenas razones que ciertas preferencias (las heteroorientadas y las desinformadas) pueden ser excluidas. Las primeras, porque no son en propiedad preferencias propias, sino el reflejo de preferencias ajenas, como por ejemplo el placer que se deriva del sufrimiento y alegría de otros. En cuanto a las segundas, al estar basadas en el desconocimiento de una cierta información relevante, no cumplen con la condición de que el sujeto pueda racionalmente establecer las comparaciones necesarias para producir su ordenamiento; este es el caso de las comunidades que prefieren no ser incluidas en un programa de vacunación porque carecen de información suficiente acerca de los efectos tanto de vacunarse como de no vacunarse.

Creemos que Harsanyi cumple parcialmente con el ambicioso objetivo trazado, ya que en efecto deja completamente clara la incidencia de las decisiones de la autoridad en el resultado y la equidad del juego, poniendo en evidencia la fuerza del mensaje normativo de tales decisiones y el carácter ético de las mismas; por otra parte muestra la posibilidad de disolver el dilema mediante la modificación de la estructura de pagos. Sin embargo también se pone en evidencia la fragilidad de su solución cuando debe hacer frente a las dificultades derivadas de la inclusión y/o exclusión de algunas preferencias en la función de utilidad colectiva, lo cual le obliga no solo a desconocer aquello que es base de su propia teoría es decir el carácter autónomo del sujeto de las decisiones, sino que también lo lleva a subestimar el hecho de que los mensajes normativos de la autoridad no inciden de una manera homogénea sobre la totalidad de los sujetos, y en consecuencia su capacidad generadora de confianza tampoco es homogénea.

Finalmente pasamos a considerar el enfoque de los problemas que estamos analizando desarrollado en los trabajos de John Nash y que es ampliamente conocido en la literatura como el Programa Nash⁹; el mismo parte del concepto de negociación, escenario que coloca a los actores en condiciones de intercambiar autónomamente sus utilidades con el

9 Sobre el llamado Programa Nash pueden consultarse los tres artículos seminales: Nash, J., "The bargaining problem", *Econometrica*, 1950, "Non-cooperative games" *Annals of Mathematics*, 1951 y "Two-persons cooperative games", *Econometrica* 1953. Adicionalmente para una cuidadosa exposición de los teoremas que sostienen el Programa Nash, puede consultarse Binmore, K., *Just Playing*, MIT Press, 1998, página 544.

propósito de obtener una solución que sea aceptable por todas las partes; para una mejor comprensión de las intenciones de Nash es bueno recordar que toda negociación es en un cierto sentido un juego, es decir que consta de una serie de interacciones reguladas por un cuerpo de reglas. Según Nash los modelos de cooperación presuponen una negociación prejuego cuya finalidad es la de producir acuerdos *irrevocables y obligatorios* acerca de las reglas que regirán en el nivel del juego; desde este punto de vista el juego “se constituye” en el momento del prejuego, ya que todas las negociaciones prejuego quedan incorporadas en las reglas del juego, las cuales son las piezas normativas consolidadas que determinan tanto las estrategias permitidas y prohibidas, como así también la estructura de pagos o incentivos para cada alternativa. Si observamos detenidamente esta actividad prejuego y sus objetivos es fácil reconocer en ellos la tarea, que sin apelar a procesos de negociación, ha cumplido el fiscal en el modelo del Dilema de los Prisioneros.

En el nivel del juego, la selecciones posibles para los actores quedan circunscritas a aquéllas que son admitidas por las reglas, y el cálculo de utilidades esperadas se encuentra definido por la estructura de pagos asignados. A este tipo de juego que parte del concepto de que todos los acuerdos y decisiones prejuego entre los actores quedan incorporados y consolidados en las reglas, y que en consecuencia no pueden ser tomados en consideración nuevamente en el momento del juego se los denomina “contienda” (contest)¹⁰, y en ellos es posible predecir los resultados.

La pretensión del llamado programa Nash es ambiciosa; si de alguna manera mediante un enfoque no-cooperativo de las negociaciones prejuego fuera posible establecer la solución para cualquier contienda (en el sentido expresado), se habrían resuelto los problemas de los juegos cooperativos; esto implica superar las diferencias entre los enfoques cooperativos y no-cooperativos, en un intento por aprovechar los aspectos más fértiles de ambos. Apoyados en este punto de vista, mediante el análisis de los equilibrios de los modelos de negociación no cooperativos los cuales se evidencian capaces de capturar la verdadera naturaleza de los procesos de negociación, podríamos verificar la plausibilidad de un determinado modelo de solución cooperativa en lugar de confiar en el mismo de una manera ciega. En otras palabras que al momento de proponer un determinado concepto de solución cooperativa en un escenario específico, deberíamos poner a prueba cuanto sabemos acerca de las instituciones de negociación en tal escenario, como un modo de promover la refutación o confirmación de la propuesta.

10 Puede verse Binmore, K., “Nash bargaining theory” en K. Binmore and P. Dasgupta (ed.) *Economics of Bargaining*, Cambridge University Press, 1987.

Creemos que no hay un modo mejor para mostrar el sentido profundo del llamado Programa Nash que el símil que frecuentemente se emplea cuando se compara el uso de conceptos de solución cooperativa para predecir resultados de negociaciones sin haberlos previamente sometido a prueba, con el hecho de saltar de un edificio con unas alas atadas a la espalda esperando volar apropiadamente. Contrariamente, el someter a prueba tales conceptos mediante el uso del análisis de los equilibrios de los modelos no-cooperativos equivale a evaluar la aerodinámica del artificio presuntamente volador mediante el test de un túnel de viento, con el fin de determinar las condiciones en las que efectivamente puede volar. Parece claro que si poseemos mayores conocimientos acerca de los procedimientos concretos de negociación que efectivamente se usan en un cierto escenario, mayor será también nuestra capacidad para establecer si una cierta solución cooperativa es capaz de predecir los resultados.

Sin dudas la solución propuesta por el Programa Nash representa un aporte inestimable al tema de la plausibilidad y de la realizabilidad de los modelos¹¹, además de constituir un recurso metodológico ingenioso y eficiente en la exploración de los mecanismos institucionales de intercambio realmente existentes; pero en lo relativo a la generación de confianza recíproca resulta frágil, ya que la sola existencia de una regla consolidada mediante un proceso de negociación no parece suficiente para garantizar que los acuerdos que han producido las reglas en el nivel prejuego sean efectivamente irrevocables y obligatorios, y sobre todo que en el nuevo nivel del juego sean realmente respetados por los actores sociales.

4. EL CAMINO DE LAS INSTITUCIONES

Los tres enfoques conceptuales que hemos analizado ofrecen respuestas de importancia a cuestiones cruciales para superar los escenarios de desconfianza que impiden atravesar lo que hemos denominado “espacios minados”, y que finalmente obligan a recorrer los más seguros caminos que conducen a las indeseables soluciones no cooperativas. En los tres autores se evidencia claramente el interés por caracterizar los factores que conspiran contra el logro de soluciones que sean aceptables en el nivel individual y al mismo tiempo óptimas desde el punto de vista colectivo.

Axelrod, en primer lugar lograr establecer algunas de las condiciones que permiten a los jugadores confiar en los resultados cooperativos; así, apela al carácter repetitivo del juego social para utilizar la interacción como un mecanismo informacional generador de confianza. Por otra parte descubre que la acumulación de las ganancias en el propio escenario de

11 Para quienes tengan interés en profundizar la función que desempeña la “realizabilidad” en la estructura de justificación de los sistemas éticos me permito sugerir la lectura de Barragán, J., “Ethical Issues in Public Decision Making”, *Rechtstheorie*, 29.Band 1998.

juego constituye otro factor de credibilidad como también lo es el manejo democrático e inseguro de la información relevante. Asimismo demuestra que garantizando estas condiciones en el diseño institucional y de incentivos los comportamientos altruistas son estratégicamente convenientes para quien los practica, y simultáneamente constructivos desde el punto de vista social. Sin embargo, tal como sucede en el clásico Dilema de los Prisioneros, de manera implícita Axelrod mantiene en la penumbra la figura del decisor normativo.

Por su parte Harsanyi intenta que la generación de confianza derive del carácter neutral del decisor normativo, el cual debe colocarse en un plano ético desde donde en la construcción de la función de bienestar colectivo tomará en cuenta las preferencias de los sujetos a los cuales van dirigidas sus decisiones. En este caso la confianza se derivaría de la calidad de las decisiones normativas, las cuales mediante un procedimiento de corrección de la estructura de pagos serían capaces de garantizar la equidad, y a través de un sistema de incentivos estarían en condiciones de ponderar las preferencias, modificando la estructura de las utilidades esperadas. En este caso el decisor normativo abandona la penumbra, pero no así sus métodos de decisión en materia de ponderación y exclusión de preferencias, decisiones estas que quedan finalmente sostenidas por elementos muy frágiles.

Finalmente el Programa Nash muestra analíticamente la posibilidad de que los sujetos sociales realicen acuerdos acerca de las reglas institucionales que más tarde regirán su propio juego social, generando mediante esos acuerdos obligatorios e irrevocables un espacio de confianza para el desarrollo del juego mismo. Sin embargo, no resulta transparente la razón por la cual estos acuerdos sean aceptados como obligatorios e irrevocables y no correrán la misma suerte de otros acuerdos o reglas.

Tal como puede observarse, en los tres autores analizados se hace presente, aunque con diversa transparencia y modalidad, la idea de que la llamada autoridad normativa puede ser considerada como el mediador más eficiente en la generación de la confianza necesaria para superar la encerrona del Dilema de los Prisioneros. En el caso planteado por Axelrod, aunque este punto no se hace expreso, es evidente que Axelrod cuando va corrigiendo las restricciones de su modelo hasta lograr un escenario en donde el altruismo pueda florecer y desarrollarse actúa como autoridad normativa; en el caso de Harsanyi la presencia de la autoridad normativa se hace expresa, y asume claramente su papel de generador de confianza, hasta el punto de poder modificar las estructuras de pagos y excluir preferencias perversas que atentarían contra la credibilidad de la propuesta institucional realizada. Y por último en el caso de Nash, la autoridad normativa asume la forma de pre-jugadores que mediante procesos

de negociación procuran establecer reglas institucionales que sean capaces de generar seguridad y confianza en el ulterior nivel del juego.

Este enfoque que considera a las instituciones como la instancia mediadora que más eficazmente puede ayudar a la generación de confianza colectiva ha cobrado un importante espacio intelectual; sin embargo hay un grupo de pensadores que apuesta por una solución apoyada en la existencia y desarrollo de la llamada confianza individual¹². Tal como se ha hecho evidente en el análisis precedente, parece claro que a pesar de los buenos deseos y del meticuloso trabajo intelectual de Axelrod acerca de que la construcción del espacio de confianza social pueda ser iniciado a partir del nivel de la confianza individual, detrás de todas las soluciones que el propone siempre encontramos la presencia de una decisión normativa que representa un diseño institucional y de incentivos que sirve para apuntalar la confianza colectiva.

La diferencia entre confianza individual y confianza social se vincula a un punto que no debería ser ignorado (aunque con frecuencia lo es) por los enfoques conceptuales; se trata de la diferencia entre conocimiento individual y “conocimiento común”, que fuera brillantemente capturada en un ya clásico artículo por Bob Auman¹³. Este vínculo obedece a que la confianza individual que hace previsible el logro de una solución cooperativa no puede ser considerada como una condición sino que ella es una consecuencia de la existencia de un “conocimiento común” entre los jugadores; este “conocimiento común” (que no es sino una consecuencia de los llamados eventos públicos)¹⁴, se presenta cuando todos los actores sociales o jugadores conocen plenamente todos los a-prioris y a-posterioris del juego, y adicionalmente cada uno de ellos sabe que el otro sabe que el primero los conoce:

“When we say that an event is “common knowledge”, we mean more than just the both 1 and 2 know it; we require also that 1 knows that 2 knows it, 2 knows that 1 knows it, 1 knows that 2 knows that 1 knows it, and so on”.¹⁵

Esto significa tanto como decir que los actores sociales no solo “comparten” toda la información relevante sobre el juego sino que han hecho de ella un conocimiento de todos, lo cual configura un escenario

12 Los más importantes aportes en relación con el tema de la generación de confianza pueden ser revisados en Braithwaite, V. y Levi, M. (eds.), *Trust and Governace* New York, Russell Sage Foundation, 1998 y Zeckhauser, R. (ed) *The Strategy of choice*, MIT Press, 1991.

13 El tema del “common knowledge” se encuentra formalmente caracterizado en Auman, R., “Agreeing to disagree”. *Annals of Statistics*, 4, 1976 y ha recibido tratamiento axiomatizado en Milgrom, P., “An axiomatic characterization of common knowledge” *Econometrica*, 49, 1981.

14 Véase Milgrom, P., op. cit.

15 Ver Auman, R., op cit. Página 1236

propicio para la confianza colectiva, la que a su vez permite que la confianza individual pueda ser ejercida sin que el actor que la practica quede en una situación de alta vulnerabilidad. Es decir que el “conocimiento común” constituye la única base sobre la cual es posible prever resultados y soluciones ya que bajo las condiciones definidas por el “conocimiento común”, no podría suceder nada en el marco del juego que no fuera conocido por todos los jugadores. Si consideramos las propuestas de Axelrod, Harsanyi y Nash que hemos analizado comprobamos que en todas ellas se encuentra implícito el supuesto de un “conocimiento común” que de alguna manera es el que hace posible la previsión de los equilibrios y las soluciones. El hecho de que ese supuesto no haya sido explicitado por los autores obedece a una razón bastante simple y ciertamente relevante; el contenido mismo del supuesto constituye también un “conocimiento común” entre quienes trabajan en la teoría de los juegos, y por tanto efectuar su aislamiento no resulta fácil para quienes lo comparten.

Esta situación se pone en clara evidencia cuando en la realidad se encuentran ausentes las condiciones de “conocimiento común” teóricamente supuestas, produciéndose en términos generales tres consecuencias encadenadas; en primer lugar se efectúan previsiones equivocadas sobre los equilibrios, en segundo lugar se ignora el origen de tales equivocaciones y finalmente se tiende a considerar como arbitrarias o irracionales las selecciones efectuadas por los actores¹⁶.

A esta altura no resulta difícil aceptar que solo la autoridad normativa están en condiciones de promover los arreglos institucionales y los sistemas de incentivos que califiquen como “evento público”; si a su vez tomamos en consideración que el “conocimiento común” deriva de los “eventos públicos”, tendremos que aceptar que solo la autoridad normativa puede crear las condiciones que den lugar al “conocimiento común” capaz de hacer nacer la confianza colectiva, y en consecuencia permitir que el ejercicio de la confianza individual no implique asumir los riesgos de una alta vulnerabilidad y de importantes pérdidas de utilidades.

5. EL MENSAJE DE LAS INSTITUCIONES

Por el momento es posible afirmar sin mayores dudas que la sola actividad o voluntad individuales resultan insuficientes para generar el llamado “conocimiento común” necesario para instalar un clima de confianza colectiva. Aún cuando un grupo social comparta una misma plataforma histórica, y una herencia cultural que le haga a su vez partícipe

¹⁶ Estas consecuencias son especialmente relevantes en casos referidos a predicciones relativas a culturas diversas con las que los que efectúan las previsiones no han desarrollado “conocimiento común”; para quien esté interesado en estos casos, puedo sugerir la lectura de Barragán, J., “El lugar de la tolerancia en las decisiones normativas” en el volumen *Retos de la razón práctica: tolerancia, pluralismo y banalidad*. Universitat de Barcelona, 2002.

de un lenguaje común¹⁷, esto no resulta suficiente para producir una trama de confianza colectiva que permita el ejercicio de la confianza individual. Esta insuficiencia del nivel individual deriva de que la capacidad para producir “eventos públicos” radica fundamentalmente en la autoridad normativa. Pero, si las cosas fueran tan simples, no sería difícil eliminar todas las ineficiencias derivadas de la ausencia de confianza mediante el recurso de seguir unas ciertas reglas en la elaboración de los arreglos institucionales y el diseño de los incentivos de modo que los mismos fueran en todos los casos “eventos públicos”. Sin embargo, a juzgar por los resultados que a diario se observan, es evidente que o bien tales reglas no existen o bien las dificultades de esta empresa son mayores de las que en principio se hacen visibles.

Puesto que no todos los arreglos institucionales ni todos los diseños de incentivos logran el mismo éxito en la importante tarea de generación de confianza colectiva, resulta conveniente revisar analíticamente estos procesos, no con el fin de ofrecer respuestas inequívocas al problema sino con el propósito de definir líneas razonables de perfeccionamiento de las soluciones. A fin de organizar el análisis debemos en primer lugar tomar en consideración el hecho de que la confianza es básicamente una actitud interna de los sujetos basada en percepciones complejas; en consecuencia no sería suficiente con que las instituciones y los incentivos procuren una convergencia generalizada de los comportamientos alrededor de la regla, sino que deben dirigir su objetivo hacia la construcción y consolidación de esa actitud interna. En este sentido no solo se deberían establecer sanciones para las conductas divergentes sino que habrían de tender a crear una atmósfera de defensa de la confianza recíproca, para lo cual hay que tomar en consideración el punto de vista, las percepciones y las preferencias de los jugadores sociales.

Trazado ese objetivo central, lo cual implica asumir un punto de vista que privilegia ese valor sobre otros, se ha dejado suficientemente claro que la autoridad normativa es la destinada a producir los llamados “eventos públicos”, los que a su vez tienen la capacidad de generar el “conocimiento común” sobre el cual se apoya la confianza. En efecto, esta actividad solo puede ser cumplida en el nivel del metajuego, que es el nivel jerárquico propio de la autoridad normativa como productora de las reglas de juego. En tal plano es posible trabajar utilizando los mecanismos de negociación propuestos por el Programa Nash, o bien los mecanismos propios de un decisor ético propuestos por Hansanyi; sin embargo, en cualquier caso las reglas producidas deberán poseer rasgos tales que sirvan al logro del objetivo, es decir deben ser generadoras de confianza. La mayor o menor

17 El compartir una herencia y un lenguaje es sin duda una condición necesaria pero no suficiente para el establecimiento del llamado “conocimiento común”, ya que este no depende solo de la presencia de tales variables sino de un acto normativo consistente en la generación de un “evento público”.

eficiencia para alcanzar tal objetivo servirá asimismo como criterio de evaluación de la calidad de la regla.

Estos criterios para evaluar la calidad de las reglas del juego social conciernen a dos niveles que en la práctica se entrelazan permanentemente: por una parte, está el nivel valorativo asociado a los principios básicos que deben fundar las decisiones normativas (reglas) que aspiren a brindar una solución satisfactoria al problema de generación de confianza; y en segundo lugar, está el nivel instrumental referido a los contenidos de los cuerpos de incentivos que forman parte de las restricciones establecidas por la regla.

En el entendido de que la confianza es una actitud interna basada en percepciones complejas, en el nivel valorativo es necesario sostener en primer lugar el principio de respeto por la autonomía de los sujetos. Acerca de este punto no parece difícil ponerse de acuerdo en un plano conceptual, aunque en el terreno práctico los diseños de instituciones no siempre logran articular acertadamente los mecanismos más idóneos para hacerlo efectivo. Esto tampoco debe sorprendernos, ya que el problema fundamental a que siempre se enfrentan quienes procuran establecer instituciones e incentivos es el de lograr enlaces consistentes entre el plano de los principios que resultan sostenibles y el de las consecuencias concretas que es posible aceptar.

Para darle forma tangible al valor implicado en el principio de respeto por la autonomía de los sujetos, la regla del juego no puede ser considerada por el decisor normativo como una razón que queda plenamente protegida y que será eficaz solo por su capacidad de imponer sanciones, ya que éstas, además de tener un alto costo social de aplicación, se dirigen básicamente a lograr conductas de acatamiento que no servirían al propósito de generar confianza. En efecto, toda vez que el cálculo de utilidades así se lo aconseje, el jugador que está dispuesto a seguir las reglas del juego solo en base a una actitud de acatamiento derivada del temor a la sanción, estará igualmente dispuesto a violarlas, ya que en caso contrario iría en contra del criterio de racionalidad. Esto indica que los mecanismos de sanción no son eficientes para generar actitudes internas, y tampoco preservan adecuadamente el principio de autonomía de los sujetos.

Un sistema de incentivos que intente preservar el principio de autonomía de los sujetos debe ofrecerles a los mismos razones que sean aceptables por un jugador racional el cual por definición debe privilegiar en sus decisiones sus propias preferencias. Para que una solución de este tipo sea posible y estable es necesario que la regla entre a formar parte del razonamiento práctico de los sujetos, circunstancia esta que solo se produce cuando la autoridad normativa incorpora bayesianamente a sus probabilidades subjetivas a priori las correcciones derivadas de las preferencias de los sujetos sociales que participarán en el juego. Asimismo,

en respeto por la autonomía de los sujetos, el cuerpo de incentivos no puede obligarlos a asumir comportamientos supererogatorios o que impliquen exigencias que por razones culturales o materiales no están en condiciones de cumplir¹⁸. En síntesis, el diseño institucional y los incentivos asociados deben constituir un balance de las razones y preferencias de los actores llamados a utilizarlos.

Habría aun un segundo principio normativo que junto al de respeto por la autonomía de los sujetos constituyen las piedras angulares sobre las que reposa un diseño institucional y de incentivos que quiera producir en los jugadores una actitud interna de aceptación de las reglas del juego; se trata del principio que obliga a la regla a ser una generadora de confianza. Tal como sucede con el principio que anteriormente hemos considerado, también con la generación de confianza es más fácil aceptar su importancia en la teoría, que articular los recursos apropiados para llevarla a la práctica. Suele suceder que en muchos casos se supone que la confianza constituye una creencia de origen psicológico que puede ser activada mediante recursos emocionales, y no es infrecuente asistir a “campañas publicitarias” para aumentar los niveles de confianza.

En el contexto teórico que estamos utilizando hemos establecido que la confianza deriva del “conocimiento común” y éste a su vez deriva de los “eventos públicos”; de modo que los arreglos institucionales y de incentivos que dan lugar a las reglas del juego social, deben ser “eventos públicos”. Esto supone que el nivel informacional que proporcionen las reglas debe ser superior al que individualmente poseen los jugadores, y debe ser compartido de manera insesgada por la totalidad de ellos. Este nuevo marco informacional tiene la capacidad de generar en los actores la confianza necesaria para aceptar las restricciones que les imponen las reglas; tal es el sentido de las modificaciones efectuadas por Axelrod a las reglas iniciales, modificaciones que condujeron a una generalización de la confianza y de los comportamientos altruistas.

El otro elemento que en el diseño institucional y de incentivos es consistente con el principio de preservación de la confianza colectiva, es el de que el cuerpo de incentivos debe estimular de manera coherente en los actores sociales los sentimientos de respeto hacia este principio; es evidente que en este sentido el propio diseño de las instituciones y la

18 Una de las herramientas más poderosas que es utilizada en el moderno diseño de incentivos es el de los impuestos; ellos constituyen la contribución que los miembros de una sociedad determinada realizan con la finalidad de satisfacer los costos de los bienes comunes y públicos que son fundamentales para la vida social. En este sentido, para ser eficientes, las cargas impositivas demandan de una aceptación interna por parte de los jugadores, aceptación que está condicionada por la confianza que tales jugadores tengan en la justicia de los montos de las cargas asignadas y en la efectiva equidad de su distribución. Para uno de los más importantes análisis de las implicaciones del problema teórico que nos ocupa puede verse el trabajo seminal de Auman R. y Kurz M., “Power and Taxes” en *Econometrica*, Vol 45, No. 5, 1977

estructura de los incentivos constituye un mensaje acerca de la creencia que tiene la autoridad normativa en la vocación de corrección de tal principio. Si se analiza el cuerpo de incentivos establecido por el fiscal para la versión original del Dilema de los Prisioneros es fácil inferir que el mismo está regido por el principio de incentivación de la desconfianza recíproca, lo cual se expresa en la restricción de la información y la estructura de las utilidades esperadas por los jugadores. Inversamente, el diseño efectuado por el anciano en el caso presentado por Hoerster evidencia su creencia en la fuerza constructiva de la confianza, creencia que se hace evidente en la remoción de los obstáculos para la comunicación, y en la promoción de los acuerdos de cumplimiento obligatorio.

Si tomamos en cuenta todo lo hasta aquí analizado, podemos inferir que además de establecer prohibiciones, permisiones y obligaciones que son reforzadas por los sistemas de incentivos, las instituciones producen un mensaje tácito acerca de aquellos principios en cuya corrección la autoridad normativa cree, y también acerca de su disposición en aceptarlos y contribuir a que los mismos sean aceptados por todos. Si consideramos el mensaje del fiscal en el Dilema de los Prisioneros, no nos puede extrañar que el mismo sea un generador de desconfianzas, como tampoco es de extrañar que la cooperación se haya difundido con el mensaje institucional construido por Axelrod.

6. CONCLUSIONES

Una de las mayores dificultades para articular la racionalidad individual y la racionalidad colectiva deriva de los problemas que conlleva la generación de confianza recíproca entre los actores sociales. Como lo evidencia el Dilema de los Prisioneros tal ausencia de confianza es la responsable de que acciones individuales que buscan maximizar las utilidades esperadas conduzcan inexorablemente a resultados sociales subóptimos. Con el diseño del escenario y el sistema de incentivos establecidos en el Dilema (que modela adecuadamente muchas situaciones de la realidad) no es para nada sorprendente que cada uno de los actores sociales tema asumir actitudes altruistas, ya que esto los haría vulnerables frente a los demás actores, con riesgos de importantes pérdidas. Cuando ese temor se generaliza se producen severos daños para el tejido social, los cuales se hacen particularmente evidentes en los llamados bienes comunes y bienes públicos.

Alguno estudiosos de la materia han insistido en que mediante una serie de recursos, generalmente de origen deontológico, sería posible lograr la modificación de tales comportamientos; este enfoque que intenta la solución del problema sobre la base de destacar el carácter "virtuoso" de los comportamientos altruistas, ha impedido prestar toda la atención que se merece a la figura del fiscal, el cual mediante el diseño de una

institución y un sistema de incentivos ha contribuido a crear espacios minados a través de los cuales es muy razonable no transitar. No es exagerado decir que la intensión del fiscal es la de crear desconfianza entre los actores, y fomentar la no-cooperación. Este mensaje implícito del fiscal se hace muy evidente cuando se analiza la actitud del anciano que según el modelo de Hoerster procura estimular la cooperación y la confianza entre sus hijos.

No es sorprendente que sean muchos los autores que se han interesado en temas relacionados con este problema, ya que encontrar vías de solución al mismo implicaría un importante avance en la tarea de conciliar la racionalidad individual y la colectiva. Hemos analizado tres relevantes contribuciones en la materia que nos han permitido acceder a sendas propuestas. Según Robert Axelrod, con algunas modificaciones introducidas en las restricciones establecidas por la reglas del juego, no solo se logra que el altruismo sea estratégicamente conveniente, sino que se consigue instalar un clima generalizado de confianza; para John Harsanyi un decisor ético, es decir que opera con equidad y respeta las preferencias individuales, puede modificar la estructura de los incentivos, favoreciendo conductas de cooperación; finalmente según John Nash, estos mismos propósito pueden lograrse si el juego es precedido de un pre-juego en el que se fijen acuerdos irrevocables y obligatorios que estructuren un marco previsible de confianza. Los tres autores analizados muestran en diversas maneras el poder que tiene la autoridad normativa para crear las condiciones que hacen posible un juego social en el que el altruismo no constituya un comportamiento irracional.

Si consideramos que la confianza se origina en el llamado “conocimiento común” el que a su vez proviene de los “eventos públicos” y que estos a su vez solo pueden ser producidos para la autoridad normativa, no podemos dejar de admitir el papel crucial que la autoridad juega en la promoción de la confianza. Esto nos lleva a preguntar cuáles son los rasgos que debe poseer un diseño institucional y de incentivos para ser capaz de generar el “conocimiento común” que origine la confianza colectiva.

Los criterios para evaluar tales rasgos aluden a dos niveles: el nivel valorativo y el instrumental, los cuales están íntimamente relacionados y hasta pueden considerarse interdependientes. En el plano valorativo hay dos principios que sostienen una estructura institucional que pretenda generar “conocimiento común”: el principio de respeto por la autonomía de los sujetos y el principio de preservación de la confianza. En relación con el primero, los elementos instrumentales que se inspiran en él y a la vez permiten su concreción pasan por que el diseño institucional y de los incentivos a él asociados constituyan un balance de las razones y preferencias de los actores llamados a utilizarlos; mientras que los elementos instrumentales relacionados con el segundo principio pasan por

que el nivel informacional que proporcionen las reglas del juego sea superior al que individualmente poseen los jugadores, el cual debe ser compartido de manera insesgada por la totalidad de ellos.

Estos rasgos no constituyen, ni pueden constituir, reglas inequívocas que garanticen que un diseño institucional y de incentivos sea efectivamente apto para producir “eventos públicos” capaces de generar “conocimiento común” y en consecuencia confianza, pero sí constituyen mensajes normativos suficientemente claros acerca de las creencia fundamental de la autoridad normativa en la superioridad valorativa de un enfoque de las relaciones sociales frente al otro. A pesar de que muchos expertos se inclinan por artilugios metodológicos infinitos para controlar la eficiencia de los diseños institucionales, la voluntad normativa de hacer que los comportamientos altruistas no sean supererogatorios es el mejor signo de que la confianza no nos hará vulnerables.