



Universidad Nacional Autónoma de México

---

Facultad de Filosofía y Letras

Colegio de Filosofía

Racionalidad y emociones:  
Una evaluación crítica para la toma  
de decisiones

TESIS

Que para obtener el título de  
Licenciado en Filosofía

Presenta:

Peralta Huerta Carlos Baruch

Asesor:

Dr. Moisés Vaca Paniagua



Ciudad de México

Abril 2018



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*Escogió incluir las cosas  
que unas en otras se incluyen, la total,  
la complicada, la amontonada armonía.*

Wallace Stevens  
(Notes Toward a Supreme Fiction)

## Agradecimientos

Estos últimos dos años, tiempo que ha tomado la elaboración de este proyecto, podría afirmar con certeza, han sido los años más complicados a los que me he enfrentado, aunque no por ello menos fructíferos —sino todo lo contrario—; una de las palabras de mayor aparición durante este periodo fue la de *ciclo*. Muchos ciclos se cerraron, otros tantos se han abierto. Con ello ha venido la afirmación constante de la idea que se ha ido desarrollando en mi cabeza desde hace ya muchos años, a saber, que el límite de mi *yo* no es únicamente la relación de mi cuerpo y mi mente, sino que éste incluye también todo aquello que me rodea, en el espacio en el que me desenvuelvo, con lo que interactúo y con quienes me relaciono. Por ello agradezco a todos aquellos que se han cruzado en mi camino. Presento a continuación tan sólo algunos agradecimientos especiales.

En primer lugar, agradezco a mi familia por todo el apoyo brindado siempre: Lucía Huerta, madre, gracias por todo el cariño, por siempre estar. Samai Peralta, hermana, gracias por cada segundo que hemos pasado juntos, por tu compañía y consejo. Gracias a mis tíos, Alicia Huerta y Erasmo López, por su ayuda brindada incondicionalmente. Gracias a mis primos, César, Abde, Arath, Renata, por cada risa y cada conversación que hemos podido tener. Gracias por toda su confianza y todo el apoyo que me han dado en cada una de las decisiones que he tomado en la vida.

Por otro lado, quiero darle las gracias a mi segunda familia. Alitzel Pérez, muchas gracias por tu compañía y apoyo en esta etapa de mi vida; gracias por todas las conversaciones que hemos tenido, por los acuerdos y desacuerdos que tenemos con respecto a lo que sea que hablemos y, sobre todo, gracias por las experiencias. Mariana Saucedo, gracias por tu apoyo incondicional, por escucharme y permitirme escucharte; gracias por todas las conversaciones acerca de los temas que, cada uno por su parte, abordamos, y que siempre, por más serias que sean, terminan en risas. Martha Herrera, gracias por haberme permitido verte crecer y haber crecido en tu compañía; gracias por ver cómo la idea de esta tesis se desarrollaba, quizá inconscientemente, desde hace casi ya nueve años atrás. Jenn Dorantes, gracias por cada palabra e idea que pudimos compartir.

Agradezco a la Dra. Margarita Valdés todo su apoyo y compañía que me ha brindado durante estos últimos años. Gracias por todo el trabajo en el que me ha permitido colaborar a su lado. Gracias por todas las conversaciones que hemos tenido, por permitirme discutir mis ideas y ayudarme a construir nuevas; a ella le debo mi concepción actual acerca de cuál es la tarea fundamental de la filosofía. Gracias por impulsarme a hacer lo que quiero y a defender lo que pienso. Gracias por enseñarme a aplicar lo aprendido en la carrera y construir con ello el camino a algo nuevo; gracias también por impulsarme a seguir aprendiendo.

Muchísimas gracias al Dr. Moisés Vaca Paniagua por aceptar dirigir esta tesis, por todos los comentarios, observaciones y consejos alrededor de la misma. Gracias por estar al pendiente sobre su realización; muchísimas gracias por su ayuda. También agradezco el apoyo recibido a través de una beca al proyecto PAPIIT IA 400715 titulado “Por un liberalismo político crítico”, del cual es responsable el Dr. Moisés Vaca.

Agradezco también a la Dra. Maite Ezcurdia Olavarrieta por su apoyo y su confianza. También le agradezco permitirme ser becario en el programa de ayudantes de investigador nacional nivel III o emérito del Sistema Nacional de Investigadores del Consejo Nacional de Ciencia y Tecnología (CONACyT).

Gracias al Instituto de Investigaciones Filosóficas por el apoyo que me brindó a través del programa de Estudiantes Asociados, al ofrecerme algunas de las herramientas necesarias para llevar a cabo mi investigación, así como la oportunidad de seguir desarrollando mis habilidades académicas e impulsar mi conocimiento sobre temas discutidos actualmente en el ámbito filosófico.

Finalmente, muchas gracias, por un lado, a todos mis sinodales que me leyeron y me ayudaron con sus comentarios con respecto a la tesis; por otro lado, gracias a todos aquéllos que no menciono pero que también me han impulsado a salir adelante, por su amistad valiosa y sincera.

# Índice

<b>Introducción General</b> .....	1
<b>1. La teoría de la elección racional</b> .....	5
1.1. <i>¿Qué es la TER?</i> .....	6
1.2. <i>Los límites de la TER</i> .....	11
1.3. <i>Jon Elster y Amartya Sen: Una extensión de la teoría</i> .....	14
1.3.1. <i>Jon Elster</i> .....	14
1.3.2. <i>Amartya Sen</i> .....	18
1.4. <i>Recapitulación</i> .....	20
<b>2. La teoría de juegos</b> .....	23
2.1. <i>La teoría de juegos: Racionalidad o irracionalidad</i> .....	25
2.2. <i>Formalización de la teoría de juegos</i> .....	30
2.3. <i>Juegos cooperativos</i> .....	36
2.3.1. <i>El dilema del prisionero</i> .....	40
2.3.1.1. <i>¿La defección es racional?</i> .....	44
2.4. <i>Las reglas del juego</i> .....	46
2.5. <i>Recapitulación</i> .....	46
<b>3. Las emociones</b> .....	48
3.1. <i>¿Qué son las emociones?</i> .....	50
3.2. <i>Racionalidad y emociones</i> .....	54
3.2.1. <i>El estado cognitivo de las emociones</i> .....	54
3.2.2. <i>Patrones de conducta: normas sociales y valores</i> .....	58
3.2.3. <i>Elección</i> .....	61
3.2.4. <i>Emociones racionales</i> .....	63
3.3. <i>Emociones adecuadas</i> .....	67
3.3.1. <i>Emociones adecuadas en el Dilema del prisionero</i> .....	68
<b>Conclusiones generales</b> .....	72
<b>Bibliografía</b> .....	79

## Introducción General

Uno de los grandes problemas a los que nos enfrentamos día con día es el de *tomar decisiones*. Todo el tiempo llevamos a cabo una elección. Las decisiones pueden ser desde cuál será la primera acción que se realizará, qué ropa se pondrá uno, si se tomará té o café; incluso puede enfrentarse a cuestiones morales, como elegir si se debe ayudar a la persona que se está ahogando junto al río por el que se sale a correr todas las mañanas o esperar a que otra persona lo haga. Cada elección tendrá un nivel de dificultad mayor o menor dependiendo de cuál sea la cuestión a la que nos enfrentemos, pues no podríamos evaluar de igual manera, por ejemplo, la elección entre un helado de chocolate o de vainilla y la elección entre elegir Filosofía o Derecho como carrera profesional. Aunado a esto, en la mayoría de las decisiones interviene también una carga emocional. La mayoría de las veces, las emociones afectan implícitamente la toma de decisiones; no obstante, en muchas ocasiones las personas aconsejan que, cuando se trata de un asunto importante, uno debe guiarse por la razón y dejar a un lado las emociones. ¿Las emociones pueden ser un elemento que, en vez de imposibilitar a la razón de llevar a cabo la mejor decisión, junto con la racionalidad, permita evaluar y llevar a cabo la mejor elección de todas las disponibles dependiendo cada caso? Esta tesis pretende responder afirmativamente esta pregunta a través del ámbito filosófico.

Evaluar cuál línea de acción debemos seguir, o bien, cuál es la mejor decisión para llevar a cabo, tiene como antecedente histórico la discusión de la filosofía moderna sobre la racionalidad en el ámbito moral. Resulta inevitable recordar aquella frase de Hume acerca de que la razón es la esclava de las pasiones; no obstante, algunos filósofos, Kant, por ejemplo, sostuvieron que la razón podría actuar como un eje de motivación para casos como éste, es decir, tomar decisiones. La ética kantiana seguía los mismos principios en los que había basado su teoría del conocimiento, a saber, que la filosofía debía obtener el carácter de universalidad de la ciencia. Para ello, la razón fue el pilar de su ética y con ello sostuvo que la formulación de cualquier juicio moral tenía que darse desde ella, es decir, del interior del individuo, y no de los hechos empíricos.

De manera más contemporánea, dos de las vertientes teóricas que se han detenido a estudiar la naturaleza de la decisión racional son la teoría de la acción racional y la teoría de juegos. Esta tesis pretende ser una exploración situada en tales coordenadas. En particular, me sumo a teóricos como Elster, Sen, Di Castro, Hargreaves y Varoufakis, Sousa, Nussbaum y Camps que han defendido que dichas teorías deben abandonar la idea de que las emociones son una limitante de la razón.

Así en el primer capítulo me detengo en la pregunta: ¿considerar a un individuo como un ser racional es suficiente para evaluar una toma de decisiones? Como se verá, para que una acción se considere como racional, se deben tomar en cuenta los deseos, las creencias y la información recabada por parte del individuo, según la teoría de la elección racional. Dicha teoría se analizará en esta tesis a través de su versión clásica, ya que desde su desarrollo ha tenido muchas modificaciones que han afectado ciertos elementos aquí planteados, lo que podría, a su vez, introducir complicaciones a lo aquí afirmado. El tomar dicha versión clásica permite ver claramente las diferencias que hay de ésta con la teoría de juegos.

El segundo capítulo continúa la discusión acerca de la cuestión racional, y lo hará a través de la teoría de juegos. Si bien el primer capítulo analiza el argumento de que una decisión es racional si las creencias, los deseos y la información del individuo son coherentes, tal y como se verá, no parecen ser elementos suficientes para afirmar que lo sea. Siguiendo a Elster (2014, p. 10), “es racional tomar dentro de nuestros intereses al otro”; y no sólo, sino que también deben tomarse en cuenta los cambios externos que podría haber de un hecho, así como las reglas establecidas dentro del mismo. Por ello, estos elementos se analizarán aquí.

Del mismo modo, algo que se plantea al final del primer capítulo y que terminará de cerrarse en el segundo es la idea de la perfección racional. Las ideas desarrolladas en la teoría de la elección racional parecen sustentar la idea de que el sujeto es perfectamente racional. No obstante, esto no es el caso. Si uno se queda únicamente con la coherencia de los elementos primeramente analizados se puede caer incluso en una cuestión ideológica, pues si todo es consistente, entonces es racional, lo que podría llevar, a su vez, a tratar de imponer dicha evaluación. Los elementos añadidos por la teoría de juegos generan una limitación de preferencias, pues tal y como se dijo, se añade que las preferencias del otro, así como el entorno, afectan también la toma de decisiones y, por ende, lo que se considera como lo



racional. La imperfección racional señalada, tal y como se expresará en el capítulo dos, no puede sustentar la afirmación de que entonces el agente sólo puede tomar decisiones irracionalmente; por el contrario, la imperfección racional mencionada simplemente nos obliga a aceptar que no siempre se podrá llevar a cabo una evaluación completa para decidir cuál es la mejor estrategia.

A lo largo del capítulo dos se desarrollará también la versión formalizada de la discusión presente. Si bien es el único lugar en el que esto se expone, es una manera diferente por la cual se puede comprender lo presentado aquí. Del mismo modo, sirve para reconocer cuál es el trasfondo que hay detrás del ejemplo que se utilizará al final de este trabajo para mostrar la tesis principal, a saber, que las emociones junto con la racionalidad permiten evaluar adecuadamente la toma de decisiones.

Finalmente, el tercer capítulo se centra en hablar de las emociones. En un primer momento se tratará de trazar la idea de qué es lo que se entiende por una emoción, para después centrarse en cómo se relaciona con la cuestión de la racionalidad. Tal y como se afirma a lo largo de la tesis, buena parte de la tradición filosófica ha dejado de lado la cuestión de las emociones, e incluso se las ha tomado como un elemento que impide llevar a cabo una buena elección. Esta idea ha sido una gran influencia en la sociedad, pues actualmente se puede escuchar con facilidad que las emociones están peleadas con la razón. Como ya lo he mencionado, esta tesis pretende defender el caso contrario.

Tres aclaraciones son importantes. Primero, debo decir que la tesis principal que defiende aquí (que las emociones tienen un papel central en la elección racional) no está comprometida con una teoría en particular sobre la naturaleza de las emociones; en este sentido, es compatible con diferentes posturas a este respecto. A pesar de ello, el análisis de las emociones será guiado por las ideas de Elster.

Segundo, en esta tesis se plantean dos maneras de tener la información recabada al momento de analizarla: completa e incompleta. La primera no se analizará dado que la práctica es poco probable que acontezca un hecho en el que esto ocurra.

Tercero, una objeción que podría plantearse a lo largo la tesis es la introducción de conceptos tan variados como individuo, ser humano, agente, persona, entre otros. Dichos

conceptos no tendrán ninguna carga metafísica ni ontológica extra, sino que se usarán de manera indistinta unos de otros; en ocasiones por mor a lo afirmado en alguna cita, o incluso por lo expresado dentro del tema mismo.

En resumen, los dos principales motivos por los cuáles se desarrolla esta tesis son, por un lado, preguntarse acerca de algo que llevamos a cabo a lo largo de toda la vida, la elección, pues siempre se espera elegir la mejor opción de entre todas las posibilidades disponibles; por otro lado, poner a prueba la creencia generalmente admitida de las emociones como una limitante de la razón y presentar un argumento contrario. De ello se seguirá que las emociones adecuadas, acordes con la racionalidad, serán el elemento necesario para que se desarrolle una evaluación crítica sobre las decisiones que se llevarán a cabo.

# Capítulo 1

## La teoría de la elección racional

“No esperamos comer gracias a la benevolencia del carnicero, del cervecero o del panadero, sino a la consideración de su propio interés. No nos dirigimos a su humanidad sino a su egoísmo, y nunca les hablamos de nuestras necesidades sino de su provecho.”

Adam Smith. (Citado en Sen 1987, p. 40)

La preocupación de la filosofía y, en general, de todo el pensamiento humano por responder a la pregunta “¿cómo debemos actuar?” ha generado una gran cantidad de teorías y especulaciones sobre cuál sería la mejor respuesta a ello. El presente capítulo pretende abordar una de las teorías más importantes al respecto, a saber, la teoría de la elección racional (TER). El epígrafe citado señala uno de los conceptos centrales de ésta: el interés. Muchas de las interpretaciones sobre lo dicho por Adam Smith plantean que el autor afirmaba que el egoísmo era el método más adecuado para seleccionar qué es lo que se tenía que hacer y, por lo tanto, esto respondía a cómo se desarrollaba una buena sociedad. No obstante, como señala Amartya Sen en muchas de sus obras, parece ser que ese planteamiento únicamente se centraba en cuestiones económicas, como el comercio.

Este capítulo se centra en desarrollar qué es la *racionalidad paramétrica*, entendida desde el punto de vista particular de la teoría de la elección racional. Esta sentó algunas bases para la teoría de juegos (tema del segundo capítulo del presente trabajo), que comprende los elementos de la TER desde una *racionalidad estratégica*.

En primer lugar, se expondrán los principales conceptos y argumentos que defiende la teoría de la elección racional con el fin de comprender cómo se desarrolla el análisis a través de esta. Es importante destacar que la TER plantea un tipo de racionalidad perfecta, por lo que las cuestiones acerca del error de razonamiento tratan de ser eliminadas a través de elementos como la información y la evidencia.

El segundo apartado del capítulo busca rechazar el entendimiento de la racionalidad como perfecta; esto último es muy importante para el desarrollo de este trabajo, ya que el rechazarlo

permite desarrollar un nivel de racionalidad diferente. Las emociones, como se verá en el tercer capítulo, podrán introducirse gracias a esto último y, por lo tanto, participarán en el análisis racional como un elemento que permita llevar a cabo una mejor evaluación dentro de la toma de decisiones.

El tercer apartado se centrará específicamente en dos de los mayores críticos y, a su vez, defensores de la TER, Jon Elster y Amartya Sen. Esto permite comprender que, aún pese a todos los problemas que se le presentan a la TER, las nociones básicas de la teoría pueden desempeñar un papel importante al momento de evaluar la toma de decisiones; así mismo, presentar los argumentos de estos dos críticos permite sostener que al rechazar la idea de una racionalidad perfecta no debe caerse directamente en la solución de rechazar a la razón y asumir la irracionalidad –tema que se analizará en el siguiente capítulo–, sino, más bien, rescatar lo más importante de la teoría, asumir los problemas generados por los errores de razonamiento y buscar una solución óptima.

A lo largo del primer capítulo, por tanto, defenderé la importancia de la racionalidad, así como algunos de sus elementos, rechazando, a su vez, la idea de una racionalidad perfecta.

### *1.1. ¿Qué es la TER?*

La teoría de la elección racional (TER) surge como una crítica al modelo de la economía del bienestar que se gestó en Europa durante la primera mitad del siglo XX; ésta provocó toda una revolución metodológica en las ciencias sociales (Vidal de la Rosa, 2008). El desarrollo de la TER toma como una de sus bases la afirmación de la economía moderna acerca de que el hombre económico es aquél que es un ser racional (Di Castro, 2009, p. 46). Entender esto puede llevar a muchas interpretaciones; entre las principales se encuentran, por ejemplo, (1) tomar la idea de un ser racional únicamente como un modelo descriptivo, es decir, como una herramienta que permite identificar ciertos aspectos implicados en el objeto de estudio sin que estos conlleven a obtener soluciones de ello; (2) entenderlo como un modelo normativo, lo cual consiste en encontrar soluciones óptimas específicas y, de ese modo, utilizarlo para explicar cómo es que se tienen que alcanzar tales resultados; (3) otro de los modelos más destacados es el hipotético, con el que se pretende explicar, de manera abstracta, hechos que ocurren en el mundo. En este apartado se analizará al ser racional a través del punto (1) y (2); sin embargo, en el tercer capítulo se defenderá el punto (3), pues mi postura con respecto a

la racionalidad sostiene la idea de tomar la relación que hay entre la racionalidad y las emociones como un modelo explicativo.

Di Castro (2009) señala que “en el paradigma de la elección racional, el sujeto económico, descrito por un conjunto de creencias y por un conjunto de deseos o preferencias, dispone de un conjunto de estrategias posibles, dadas determinadas restricciones” (p.46). Bajo esta concepción, se puede entender fácilmente que dicho sujeto económico tiene la capacidad de elegir qué acción posible, de varias, llevará a cabo ante un hecho específico. La idea de un sujeto conformado por creencias, deseos o preferencias<sup>1</sup> permite suponer que su elección implicará que tiene un *interés* particular en satisfacer estos conjuntos. Abitbol y Botero (2006) afirman que el “proceso mediante el cual se articulan causalmente razones y acciones es la decisión. Decidir es elegir una acción de un conjunto de acciones posibles”.

En el marco de la TER, satisfacer los intereses particulares del sujeto económico conlleva a afirmar que la manera en la que éste lo llevará a cabo será *racionalmente*.<sup>2</sup> En términos económicos se puede decir que actuar racionalmente es optimizar o maximizar la utilidad de las creencias, deseos y preferencias a través de la elección de hacer lo que mejor lleve a cabo esta labor.

Por otro lado, en esta teoría los resultados alcanzados al haber realizado una elección muestran que “si se desea lograr y se debe hacer  $x$ ” (Di Castro 2009, p. 49; Elster, 1997, 2008). La TER exige así una cierta correspondencia y congruencia de cada uno de los elementos que la componen entre sí. Esto es, deben corresponder las creencias o deseos del agente con la acción llevada a cabo y, a su vez, con el resultado. La necesidad de este principio normativo evita complicaciones tales como el autoengaño, el *wishful thinking*, la debilidad de la voluntad, entre otros.

No obstante, debe insistirse que el asunto primordial de la teoría es la explicación de la elección tomada, no la finalidad o los resultados en sí. En otras palabras, la TER pretende

---

<sup>1</sup> Siguiendo a Davidson (1994), tanto las creencias como los deseos deben entenderse como eventos mentales, es decir, actitudes proposicionales. Asimismo, se debe asumir que las creencias cumplen ciertas características: tienen un grado máximo de plausibilidad inductiva, la evidencia disponible es lo que las causa y son elaboradas de la mejor manera posible según la evidencia (Dancy, 1985; Elster, 1986).

<sup>2</sup> Para una mayor discusión acerca del término “racional” en la tradición filosófica, así como de los diferentes tipos de racionalidad, véase el artículo de Mora Burgos (2003).

explicar que una acción es racional, en tanto que haya congruencia entre las características del sujeto económico ya mencionadas, como que el agente conscientemente lleve a cabo la elección que más beneficios le brinde. Di Castro (2009) expone lo anterior de la siguiente manera:

Los elementos básicos de la explicación son tres:

- a) El conjunto de acciones factibles, es decir, todas las acciones racionales que satisfacen las restricciones lógicas, físicas, económicas y/o mentales de la situación;
- b) el conjunto de creencias racionales que estructuran causalmente la situación y ligan cada acción con un resultado determinado; y,
- c) el conjunto de preferencias viables, esto es, la jerarquización subjetiva de los resultados esperados de cada acción. (p.50)

Por otro lado, Jon Elster (2014) al plantear ciertos fallos a la TER, expone dos maneras de comprender la racionalidad: la racionalidad paramétrica y la racionalidad estratégica (p.38)<sup>3</sup>. La primera de ellas constituye el esquema bajo el que se ha venido desarrollando la TER; la segunda, por otro lado, sugiere ciertas cuestiones con respecto a la teoría de juegos, por lo que se definirá en el próximo capítulo.

La racionalidad paramétrica consiste en la comprensión que tiene el sujeto acerca de su medio como algo paramétrico; el que algo sea paramétrico significa que ese algo está ya dado en sí mismo. Esto supone que el medio, para el agente paraméricamente racional, se le presenta regido por una serie de leyes causales y que, por lo tanto, es constante. Es importante añadir que, bajo este modelo de racionalidad, los demás sujetos también son para el agente parte del medio. El único objeto variable, por lo tanto, es únicamente el agente mismo.

Dado que para la TER es importante explicar cómo una elección es racional, es importante destacar la cantidad de información que tiene un agente acerca de los resultados posibles que hay al llevarla a cabo. La bibliografía al respecto suele tener muchas variaciones, pues se discute cuánta información se debe adquirir antes de tomar una decisión (Elster, 1983, 1985, 1986, 1989; Kahneman, Slovic y Tversky, 1982; Pears, 1985). Dicho esto, se puede plantear que hay al menos dos maneras de evaluar la cantidad de información:

---

<sup>3</sup> Véanse también Di Castro (2009), Edgeworth (1881), Elster (1997), Simon (1955, 1979).

- Información completa. En esta el agente no tiene mayor problema al momento de evaluar la acción que plantea llevar a cabo según su elección, pues tiene completa certeza de los resultados.
- Información incompleta. Se caracteriza por la presencia de ambigüedad en la información, imposibilitando al agente de manera tal que no pueda llevar a cabo una evaluación concreta acerca de cuál será el curso que tomará su elección. En esta el curso de acción puede plantearse de dos formas distintas: Riesgo e incertidumbre. El riesgo, por un lado, es un modo de información cuantificable, es decir, se pueden evaluar las opciones a elegir asignando probabilidades; la incertidumbre, por otro lado, plantea un estado de ignorancia. En este caso no es posible asignar probabilidades; el agente tiene la posibilidad de eliminar aquello que lo llevaría a obtener los peores resultados y, no obstante, no puede decidir qué acción lo llevará a conseguir la maximización de su utilidad.

El problema más interesante, me parece, radica en las elecciones tomadas cuando se presenta el riesgo o la incertidumbre ya que, si un individuo tiene la información completa entonces no tiene ningún problema en elegir. Por lo tanto, más adelante esta tesis se enfocará en los casos de información incompleta.

A través de las reflexiones llevadas a cabo por Jon Elster (1985; véase también, 1997) puede seguirse claramente cómo se desarrollan los principales argumentos de la TER y cuáles son las condiciones bajo las que se lleva a cabo: Por definición, entiéndase *A* como la acción, *C* como la creencia del agente, *D* como los deseos del agente, y *E* como la evidencia. Las condiciones que debe cumplir la explicación de la TER son:

- (1) Dado *C*, *A* es el *mejor* medio para realizar *D*
  - (1.1) La creencia tiene un grado *máximo* de plausibilidad inductiva dado por la evidencia
- (2) *C* y *D* causaron *A*
  - (2.1) La creencia es *causada* por la evidencia disponible
- (3) *C* y *D* causaron *A* *en tanto que* razones
  - (3.1) La evidencia causa la creencia de *manera correcta*
- (4) El conjunto de creencias *C* es internamente consistente
- (5) El conjunto de deseos *D* es internamente consistente
- (6) La cantidad de información recogida se ubica en una franja cuyos límites superior e inferior se definen por la propia situación

(7) La relación entre *C*, *D* y *E* satisface (1.1), (2.1), (3.1) y (6)

(8) Dado *C*, *A* es la mejor acción con respecto al grupo completo de deseos sopesados (pp. 69-75).<sup>4</sup>

Tal y como puede verse, todas las condiciones se interrelacionan unas con otras. A su vez, es posible agruparlas de tal manera que pueda entenderse qué papel juegan en la explicación; estos grupos se señalarán como los requisitos que cumplen la explicación de la TER, y son tres: la optimalidad, la consistencia y la causalidad.

La intencionalidad se explica a través de las condiciones (1), (2) y (3). Una elección racional implica que el agente es quien decide el curso que se va a tomar; es decir, el agente lleva a cabo una acción *intencional* (Davidson, 1963). La intencionalidad hace referencia al hecho de que las creencias y los deseos de los agentes pueden entenderse como buenas razones para realizar la acción. En este sentido, (1) apela a cuestiones de optimalidad, mientras que (2) y (3) se postulan como argumentos de racionalidad para (1). Es importante señalar aquí que Schick (1997) critica que los deseos y las creencias sean los únicos elementos que se necesiten y se entiendan como buenas razones para llevar a cabo la acción de decidir; añade un elemento extra, las interpretaciones. Para este autor, las interpretaciones “son valoraciones de las acciones consideradas independientemente de los resultados” (Abitbol y Botero, 2006, p.138). En este sentido, los deseos serán valoraciones de los resultados, las creencias una clase de estimación en torno a la probabilidad de que *x* puede llevar a *y*, y la interpretación será el querer que esa acción sea algo que se haga, es decir, una valoración independiente de las acciones.

La consistencia, por su parte, para ser explicada no sólo necesita de (4) y (5), sino de condiciones extra que imposibilitan la entrada a errores, es decir, elecciones irracionales. En principio, la conducta intencional sólo puede ser racional si y sólo si las creencias y los deseos del agente también lo son. Para que una creencia sea consistente, y por ende racional, debe ser creída realmente por el agente y, a su vez, la creencia debe ser cierta; los deseos, por su parte, son consistentes en tanto que sí se cree que puedan cumplirse y que, de hecho, se cumplan al llevarlos a cabo. La razón por la que la consistencia requiere de más condiciones

---

<sup>4</sup> Las cursivas son suyas. El orden con el cual Elster desarrolla su análisis es diferente al tomado aquí.



es debido a la cuestión de justificar hasta qué punto se afirma que una creencia es cierta o cómo se sabe que el deseo se cumplirá al intentar llevarlo a cabo.

La evidencia y la información son las dos nociones añadidas para ello, y pueden reflejarse en las condiciones (1.1), (2.1), (3.1) y (6). La búsqueda de información ayuda a justificar y a afirmar que la acción elegida para llevarse a cabo será racional. Es importante añadir en este punto que la información recabada puede plantearse entre un máximo y un mínimo de necesidad, por lo que queda sujeta a los deseos y a las creencias del agente la decisión de hasta qué punto se dejará de seguir añadiendo información.

La evidencia, por su parte, juega un doble papel. Por un lado, la información recabada, al momento de ser utilizada, se convierte en la evidencia que requiere el agente para justificar y llevar a cabo su elección. Por otro lado, la evidencia cumple su función como tercer requisito de la explicación de la TER, la causalidad: Dado que la relación entre creencias, deseos y evidencia son consistentes, entonces la acción elegida es el mejor curso de acción. Esto se establece en las condiciones restantes, (7) y (8).

Elster (1997, p. 111) resume lo anterior a través de la siguiente figura:

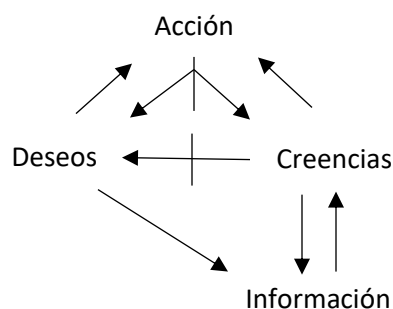


Figura 1.1

## 1.2. Los límites de la TER

Si bien es cierto que los alcances de la teoría de la elección racional lograron muchos avances en cuestiones económicas o en ciencias sociales, es cierto también que la rigidez con la que plantea la racionalidad puede llevar a algunos errores explicativos.

La TER forma parte de un tipo de explicación denominada *individualismo metodológico*, que consiste en analizar algún problema social y explicarlo a través de un individuo, más específicamente de su conducta. Este planteamiento sugiere en primera instancia el problema

de un reduccionismo. Los resultados obtenidos resultan inciertos, dado que de una cuestión general se buscan detallar ciertos aspectos particulares; si bien es cierto que esto permite explicar coherentemente algunas características del sujeto, verlo como un mecanismo imposibilita retratar correctamente todo hecho o fenómeno analizado.<sup>5</sup> Durkheim, por ejemplo, señala que dentro de la realidad social existen propiedades “macrosociales” que no pueden ser reducidas, como lo plantea el individualismo metodológico, en propiedades individuales (Noguera, 2013, p. 74). La discusión del individualismo metodológico, no obstante, no se tocará más a lo largo de este trabajo, ya que no es el tema que aquí concierne.

El segundo problema se centra en la racionalidad paramétrica. Tal y como se dijo en el apartado anterior, ésta plantea a un agente que trata a su medio como una constante, incluyendo aquí a otros agentes. De esto se pueden destacar dos consecuencias.

La primera, en términos de Jon Elster (2014), es el *cambio de preferencias temporales*. “De acuerdo con la definición común, un actor racional sencillamente es uno que tiene preferencias congruentes y completas *desde cualquier punto en el tiempo*” (p.115).<sup>6</sup> ¿Es acaso posible que un agente no sufra cambios de opinión acerca de un hecho particular? ¿No se dan, de hecho, cambios de preferencia con respecto a algo? Basta con mirar un poco al pasado para percatarse de que las creencias y los deseos no permanecen estáticos, pues todo el tiempo están modificándose. Siguiendo el análisis de la racionalidad paramétrica tendría que entenderse que, a partir de las modificaciones que pueden ocurrir con el paso del tiempo, la consistencia se perdería y, por ende, los requisitos de optimalidad y causalidad no podrían sostenerse.

La segunda consecuencia que surge de la visión paramétrica es sobre una cuestión ética. Que el agente paraméricamente racional sea el único objeto variable implica considerar, al llevar a cabo una elección, únicamente sus propios intereses. ¿Resulta realmente una elección

---

<sup>5</sup> Esta primera réplica surge dado el planteamiento que Lars Udehn (2001, p. 356) hace con respecto al individualismo metodológico –tratando de hacer una defensa del mismo–. Éste se basa en tres aspectos: 1) los conceptos sociales deben definirse a través de términos individuales, 2) debe plantearse un reduccionismo de las leyes sociales a leyes que traten acerca del comportamiento de los individuos, y 3) los fenómenos sociales deben poder explicarse a través del comportamiento individual. La importancia de este punto se basa únicamente en los fundamentos de los que se desarrolla la teoría de la elección racional; no obstante, esta es una cuestión que no se desarrollará a lo largo de este trabajo.

<sup>6</sup> Las cursivas son suyas.

racional el plantear una decisión de un yo, sin tomar las consecuencias que podrían afectar al otro? Y, del mismo modo, ¿no es cierto acaso que las elecciones tomadas por otro pueden modificar o intervenir en el propio interés? La TER, desde su enfoque paramétrico, elimina toda la cuestión social para salvar sus condiciones de optimalidad. Las consecuencias de un modelo aplicado de manera particular, para cada uno de los agentes, generaría un esquema de una sociedad egoísta. Uno de los argumentos defendidos en este trabajo consiste en entender que la comprensión del otro debe añadir información relevante al agente para llevar a cabo una mejor evaluación acerca de qué elección es racional.

Jeffrey (1983) trata este punto como una independencia de los estados del mundo con relación a las acciones del agente. Dada la racionalidad paramétrica, hay estados del mundo que no se ven afectados por las acciones o decisiones tomadas por el agente; el clima, por ejemplo, no se ve afectado por las decisiones del agente acerca de si ponerse ropa adecuada para el frío o para el calor, o si éste decide salir de casa con o sin paraguas. No obstante, existen casos en que un punto relevante del estado del mundo se verá afectado por las decisiones llevadas a cabo; un ejemplo clásico de esto se refiere al salir o no con auto, puesto que esto aumenta las probabilidades de que haya tráfico, por mínima que sea.

El tercer problema surge de los requisitos que la TER busca satisfacer. Dadas las condiciones expuestas en el apartado anterior se puede plantear una racionalidad perfecta. Aplicada como una teoría explicativa, desarrolla un argumento en el que el error no puede darse. Errores de evaluación como, por ejemplo, el autoengaño, la debilidad de la voluntad o el *wishful thinking* no encuentran una solución posible a través de la teoría. Si bien la TER puede plantearse como una teoría descriptiva únicamente acerca de los hechos que pueden considerarse una elección racional, no tiene una correspondencia directa con la realidad, en la que las creencias y los deseos del agente se ven estimulados por muchos factores que pueden beneficiar o perjudicar la toma de elecciones. En palabras de Di Castro (2009):

No olvidemos que, para esta propuesta, decidir racionalmente es elegir la estrategia que mejor satisface los deseos del agente (dadas sus creencias y restricciones); por ello, la teoría falla si no se puede establecer unívocamente cuál es la elección racional. Pero, incluso en el caso de que se pueda establecer cuál es la elección racional, la teoría también falla si la conducta realizada difiere de la normativamente prescrita. En otras palabras, la TER puede fallar porque no ofrece los requisitos racionales, o porque la gente no se comporta de acuerdo con ellos. (p. 75)

Otro de los problemas que se le presentan a la TER se da a partir de la relación de las condiciones de explicación de la teoría con la racionalidad paramétrica. Es posible plantear a través de esto que las creencias y los deseos, al no modificarse temporalmente se consideran fijos, lo que podría imposibilitar al agente de ver otro tipo de alternativas a elegir, que le darían mayor beneficio, pues siempre ha preferido una respuesta específica. También puede darse el caso de que el agente no pueda comparar correctamente todas las alternativas dadas las diferencias que se encuentran entre una y otra. Esto podría entenderse como un caso de *decisión difícil* (Levi, 1986). Un ejemplo útil para la comprensión de esto puede ser la elección de una carrera universitaria: historia del arte o administración; cada una de ellas tiene sus propios argumentos para ser elegida sobre la otra, pero no son comparables, ¿cuál es la elección racional?

Otra manera de abordar el problema o el límite se encuentra dentro de la propia ambigüedad del término racional. La TER afirma que los individuos racionales llevan a cabo la toma de decisión en la práctica de manera racional tal y como lo haría de manera teórica, pero cabe preguntar si es así, una *acción instrumental maximizadora* (Vidal de la Rosa, 2008, p.226).

Tal y como puede verse a lo largo de este apartado, la racionalidad paramétrica, vista desde esta perspectiva de la TER, comprende muchos problemas. La racionalidad, tomada como un elemento perfecto, no parece ser algo que conduzca a tomar buenas decisiones, ¿se debería, por tanto, rechazar la racionalidad?

### 1.3. *Jon Elster y Amartya Sen: Una extensión de la teoría*

Amartya Sen y Jon Elster han sido dos de los grandes críticos de la teoría de la elección racional; no obstante, llevan a cabo algunos análisis de ciertos fenómenos a través de ella, tomando en cuenta sus respectivas limitaciones, defendiendo, a su vez, ciertos aspectos de la teoría que son útiles de analizar, pues me parecen útiles para la tesis general de este trabajo.

#### 1.3.1. *Jon Elster*

Elster (2014) afirma que “el *hombre económico* puede ser definido con sus preferencias continuas y su estrecho interés egoísta, pero el *hombre racional* puede tener preferencias no

arquimédicas y ser movido por su interés por los demás” (p. 243)<sup>7</sup>. Tal y como se pudo observar en el apartado anterior, la TER se limitaba a entender lo racional a través del interés egoísta del agente. El punto de Elster es ampliar nuestra concepción de lo racional, aceptando que no sólo lo egoísta es racional, sino que también es racional tomar en cuenta al otro dentro de nuestros intereses. Esto último puede introducirse dentro de la cuestión acerca de las creencias del agente. Es importante recordar que, para la TER, la consistencia de creencias posibilita llevar a cabo una elección racional. Interesarse por el otro implica que creemos consistentemente que este último puede jugar un papel importante dentro de la toma de decisión.

Las implicaciones de esto, a su vez, se ven reflejadas en el cambio de paradigma bajo el que se entiende la racionalidad. La TER consideraba a un agente como paramétricamente racional por tratar a su medio como una constante sin variación; desde la perspectiva de un *interés* por el otro, se puede entender a un agente como estratégicamente racional. La racionalidad estratégica será utilizada como una herramienta del análisis al final del presente trabajo, por lo que se explicará más a detalle en el siguiente capítulo. No obstante, me resulta importante adelantar el planteamiento de Elster con respecto a ella. Considerar una elección planteada desde un inicio como perfecta y actuar conforme a ésta es, como ya se dijo, un error; para Elster, lo óptimo es buscar un punto intermedio, tomar en cuenta las variaciones que pueden presentarse y evaluar sobre éstas si se debe seguir el mismo trayecto de elección y avanzar, o cambiar. En palabras de Elster (2014):

Creo yo que la noción de hombre racional debe extenderse para incluir consideraciones temporales. Para ser precisos, algunos requerimientos de conciencia deben imponerse, a la vez, a la *elección de sucesiones* del actor y a su *sucesión de elecciones*. La no satisfacción del primer requerimiento queda ejemplificada por preferencias incongruentes de tiempo, y del segundo, por un cambio endógeno de preferencias. (p. 115)

Por otro lado, Elster insiste en resaltar que el ser humano en general no es plenamente racional. Los hechos empíricos demuestran las dificultades que se encuentran en un agente para actuar racionalmente. Esto se muestra a través de ejemplos ya mencionados, tales como la debilidad de la voluntad. Ésta, por ejemplo, expresa que ante una situación específica de

---

<sup>7</sup> Las cursivas son suyas.

decisión entre *A* o *B*, el agente tiene la suficiente información para elegir *A* y, no obstante, elige *B*. Dadas las condiciones en las que se encontraba, ¿lo racional no era, precisamente, elegir *A*? La importancia de la posición no radica sólo en el rechazo de la racionalidad perfecta, sino que también rechaza la idea de tomar como correcta la posición opuesta: el hombre irracional.

Algunos autores criticaron que, bajo ciertas condiciones de la TER, podría interpretarse que los animales actuaban de manera racional y que, visto así, el ser humano podría actuar como cualquier otro animal, o bien podría verse como una especie de máquina regida por algún mecanismo. En una postura contraria, puede verse a un animal como un actor que no realiza alguna acción por razones específicas, es decir, es irracional; aceptar al ser humano como un actor irracional tendría graves consecuencias en la concepción que se tiene comúnmente: creencias, actitudes, educación, etc. Esto se analizará un poco más en el siguiente capítulo, ya que hay defensores de la teoría de juegos que pretendieron en su momento defender dicha irracionalidad. Elster (2014) afirma que un aspecto que distingue al hombre de otros animales es la *capacidad de maximización global*, explicada como “la capacidad de relacionarse con el futuro y con lo simplemente posible. [El hombre] puede elegir *la* alternativa globalmente mejor porque es capaz de analizar *todas* las alternativas, todos los futuros posibles” (p. 35). El ser humano tiene la facilidad de adecuarse ante cualquier cambio que se le presenta en un hecho, mientras que en los animales se ha visto que responden así sólo a ciertos casos específicos, y los cambios y adaptaciones pueden darse por una mera evolución accidental.<sup>8</sup>

---

<sup>8</sup> Denton (2009) expone dos ejemplos empíricos que pueden dar cuenta de esto. La importancia aquí debe recalcarse en la posibilidad de comprender la multiplicidad de interpretaciones que pueden realizarse en torno al problema de la racionalidad o irracionalidad en los animales. El siguiente fragmento da cuenta de acciones aparentemente racionales llevadas a cabo como un mecanismo:

“[L]a hembra de una especie de avispa cavadora puede mantener simultáneamente varias madrigueras que contienen crías en desarrollo. [...]. [A través de una] inspección sabe qué madrigueras todavía contienen huevos que no requieren comida, cuáles son ocupadas por larvas jóvenes que necesitarán dos o tres orugas para comer, las que tienen larvas más crecidas, que necesitarán muchas más orugas, y las que contienen crías que acaban de convertirse en pupas, que es necesario sellar para para que el desarrollo se complete. Sobre la base de la visita a estas [...], la avispa sabe cuántas presas ha de capturar y dónde tiene que llevarlas [...]. A primera vista, el comportamiento parece consciente y exigente. Si se intercambian los residentes de las distintas madrigueras durante la noche siguiente [...], la avispa no tiene ninguna dificultad en abastecer adecuadamente cada guardería por la mañana. Sin embargo, si los cambios de los habitantes se hacen después de la inspección matutina, la avispa no es consciente de que en la madriguera hay

La relación que puede plantearse entre la capacidad de maximización global y las consideraciones temporales de la racionalidad permite incluir dentro de la cuestión racional a la *experiencia* y a la *enseñanza*. Es importante señalar que a lo largo del presente trabajo no se desarrollan estos elementos a mayor profundidad; no obstante, sí es importante remarcar su relevancia, pues, como se verá en el capítulo tercero, la experiencia y la enseñanza pueden ser factores determinantes para una emoción. Elster (2014) afirma:

*La experiencia* puede formalizarse como la diferencia entre preferencias *ex ante facto* y *ex post facto*. Un paquete de consumo determinado puede parecer más (o menos) atractivo después del verdadero consumo, que antes. Cuanto mayor sea la parte del espacio de bienes ya explorado, presumiblemente menor será el cambio inducido en preferencias después de cada elección. La *enseñanza* ha sido formalizada mediante la distinción entre utilidad y capital de consumo. La enseñanza, en esta conceptualización, no cambia el disfrute que yo derivo de un tiempo dedicado, digamos, a escuchar música, siendo igual todo lo demás. Sin embargo, sí causa una diferencia para el “todo lo demás” al aumentar el capital de consumo que me permite obtener más disfrute de un determinado tiempo de escuchar la música. (p. 245)

Como se puede ver, la experiencia y la enseñanza juegan al nivel de las preferencias dentro de la teoría.

Amartya Sen, por su parte, centra sus críticas y aportaciones a la elección racional desde un punto de vista ético y, más específicamente, político. Debido a esto, se retomará lo dicho al principio de este apartado. Elster señala que el hombre racional es el agente que, dentro de sus intereses, incluye a otro. Desde la TER, incluir al otro implicaba la introducción de una conducta altruista; ésta se negaba dado que podía conducir a anomalías. Introducir al otro en el análisis de la elección racional, como ya se ha afirmado, modifica la manera de obtener la utilidad; ya sea que pueda maximizarse porque el agente obtiene una ganancia al apoyar al otro, o bien puede reducirse debido a las complicaciones que otro agente puede introducir a través de sus propios intereses (Elster, 2014, p. 235). El análisis explicado así conlleva a que la teoría reduzca al altruismo, o el interés hacia el otro, en el principio egoísta: Ayudo al otro porque me da una ganancia; la ejemplificación perfecta de esto puede verse en el epígrafe

---

ahora un huevo en lugar de larvas. Pasará el día almacenando orugas en él, como si todavía contuviera larvas crecidas. Esto ocurre aunque la avispa pueda tocar los huevos varias veces a lo largo del día. Puede sellar la madriguera que contiene larvas jóvenes porque previamente contenía una pupa. La avispa no muestra ni un atisbo de inteligencia; está distraída. Está programada por el estímulo inicial” (pp. 101-102).

citado al principio de este capítulo, con Adam Smith. Amartya Sen rechaza la idea de seguir viendo al otro únicamente a través del principio egoísta.

### 1.3.2. Amartya Sen

Las inquietudes de Amartya Sen en torno a la cuestión racional, como se dijo anteriormente, se dirigen especialmente a la comprensión de problemas de filosofía política y de ética. Sen sostiene que hay ciertos elementos de la teoría de la elección racional que deben rescatarse.

En primer lugar, rescata el hecho de que la teoría intenta fundamentar las elecciones tomadas a través de una evaluación crítica de las preferencias, objetivos y creencias del individuo; esto implica que lo anterior junto con la información obtenida pueden plantearse como razones para elegir y, así, justificar por qué una acción sería, de hecho, racional. En palabras de Sen (2012):

Tener una razón para hacer algo no sólo es cuestión de convicción sin escrutinio —un fuerte sentimiento instintivo de que tenemos «excelentes bases» para hacer lo que escogemos hacer. En su lugar, se trata de que investiguemos las razones subyacentes de la elección y determinemos si las supuestas razones sobreviven al examen crítico penetrante que puede asumirse si y cuando la importancia de tal escrutinio se entienda. Las bases de la elección tienen que sobrevivir a la investigación basada en el razonamiento riguroso [...], con más información si y cuando sea relevante y accesible. No podemos solamente evaluar nuestras decisiones de acuerdo con nuestros objetivos y valores, también podemos examinar la sostenibilidad crítica de tales objetivos y valores. (p. 211)

Esto, como bien sugiere Sen, parece llevar a entender que todas las *elecciones* realizadas por algún individuo deberían tener un análisis previo para poder sostener la acción; no obstante, lo que se plantea es que, si cierta acción llevada a cabo se sometiera a un análisis, se podría ver a través de éste cómo se sostiene la racionalidad, aunque uno no siempre podría detenerse a hacerlo. La experiencia y el hábito, desarrollados por el ser humano, han permitido esto último, pues “el razonamiento sostenible puede existir sin realizar el escrutinio explícito en cada ocasión” (Sen, 2012, p. 212).

Como se dijo anteriormente, la TER busca reducir la cuestión altruista a un punto donde se sostiene su principio egoísta: Ayudamos al otro porque nos beneficia. Amartya Sen, si bien no rechaza del todo la idea, plantea que una acción elegida que pueda ayudar al otro



puede entenderse, ya sea que el interés propio del individuo se encuentra dentro de un propósito social o bien puede explicarse a través de una separación entre el beneficio social y la importancia que el individuo tome de su propia ganancia (Sen, 2002; Tuck, 2008). Por lo tanto, no es irracional buscar no sólo la promoción del beneficio propio, sino también el beneficio de los demás. El hecho de que en lo racional también entre una preocupación hacia el otro permite ver cómo es racional, a su vez, no impedir que los demás agentes busquen alcanzar sus propios fines (aunque la mayoría de las veces se dé, de hecho, un impedimento mutuo). Del mismo modo, introducir al otro en el escrutinio razonado permite limitar la relación preferencias-racionalidad. Este tipo de argumentación ha tenido fuertes repercusiones históricas; por ejemplo, por diversas razones surgidas de las preferencias de grupos específicos se han llevado a cabo genocidios. La racionalidad entendida desde el punto de vista de Sen permite evitar problemas ideológicos, así como actuar bajo deseos y preferencias ciegos.

La TER, dada la rigidez con la que postula una racionalidad perfecta, paramétrica, conllevaría a pensar una homogeneización de la razón; Sen, por su parte, sugiere un escrutinio razonado, ya que con éste las elecciones tomadas podrían ser de diferentes niveles y aun así ser racionales, v.gr. una persona *A* puede aportar mayores cantidades de dinero a un centro de beneficencia que una persona *B*, y no haber diferencia de racionalidad sólo porque una apoya más que otra. Las personas pueden dar múltiples y diversas razones para elegir una misma acción a realizar y todas podrían ser racionales; lo que Sen (2012) denomina “posibilidad de la pluralidad de las razones sostenibles” (p. 214).<sup>9</sup>

---

<sup>9</sup> Nozick (2003) sostiene que la toma de decisiones de un agente contiene tres tipos de utilidad y, dada la racionalidad de éste, puede darle más peso a una que a otra, dependiendo del caso. Los tres tipos de utilidad son la utilidad simbólica (valores de credibilidad acerca de las creencias), la utilidad evidencialmente esperada (valores que muestran el criterio esperado por el sujeto sobre lo externo) y la utilidad causalmente esperada (valores que muestran el criterio esperado entre lo externo y las creencias del sujeto). Dependiendo del valor asignado por el agente a cada una de las utilidades es como se limita y se asume un marco de referencia acerca de qué es lo racional hacer. Esto muestra a nivel general que las personas podrán tener sus propias razones para asumir una misma cosa como racional y, a nivel particular, que el individuo puede tener deseos incompatibles o creencias que no tengan cierta coherencia entre unas con otras en un momento determinado. Esto asume, por lo tanto, la limitación de la razón. Dado lo limitado de la razón, hay autores que prefieren hablar de satisfacción, en vez de maximización. Véanse Ainslie (2001), Elster (1985, 2002), Kahneman et. al. (1982), Rubinstein (1998), Simon (1978, 1986).

Por lo tanto, la postura de Sen se centra, principalmente, en los siguientes puntos:

- La elección racional debe ser reflejada en aspectos *reales*. Se debe plantear la posibilidad de que no hay una sola respuesta racional a un hecho, sino que a veces es imposible o muy difícil de predecir cuál de las acciones sería la que implicaría una mayor maximización. Así mismo, es posible que se den soluciones incompletas a un hecho; que se intente tomar la mejor decisión no implica que se tengan los suficientes elementos para llevar a cabo la elección (Davoni, Montesano y Lines, 1986; Kahneman et. al., 1982; Levi, 1986; McFadden, 1999; Stigum y Wenstop, 1983; Thaler, 1991).
- La importancia del otro. El otro no tiene por qué implicar un aspecto de irracionalidad dentro del análisis. El otro permitiría dar cuenta, por un lado, de una cuestión ética; no estamos solos en el mundo, y darse cuenta de ello implica ver cómo las decisiones tomadas por uno también afectan al otro. Por otro lado, el otro permite desarrollar una mejor evaluación, porque puede tener diferentes razones u otras opciones de elección igualmente racionales, discrepando, rechazando o aceptando las evaluaciones propias; esto, a su vez, permite demostrar la incompletitud de un análisis, haciendo que el individuo se encuentre abierto con respecto a adquirir más información y diversos argumentos. Esto puede entenderse, en términos de Sen, como *razonamiento público*.
- Autoescrutinio. “La racionalidad es, en efecto, una disciplina más bien permisiva, que exige la prueba del razonamiento, pero permite que el razonado autoescrutinio adopte formas muy diferentes, sin imponer necesariamente ninguna gran uniformidad de criterio” (Sen, 2012, p. 226). El autoescrutinio, visto así, debe desarrollarse constantemente, tratando de alcanzar una mayor consistencia dentro de cada una de las evaluaciones llevadas a cabo al momento de tomar una decisión.

#### *1.4. Recapitulación*

Tal y como se ha visto a lo largo de este primer capítulo, la teoría de la elección racional supone un tipo de comportamiento basado en el interés propio, con el que se busca maximizar la utilidad. La racionalidad, entonces, se basa en la coherencia y correspondencia de los elementos que le componen, a saber, los deseos o preferencias, las creencias y la cantidad de

información requerida. Los únicos requisitos que debe cumplir la TER son la optimalidad, la consistencia y la causalidad.

La TER, entendida desde una visión paramétrica de la racionalidad supone que todo aquello que es externo al sujeto es algo constante, incluyendo a otros individuos. Es decir, que no toma en consideración cómo puede afectar lo que otros individuos realicen o bajo qué circunstancias se presente la toma de decisión. La TER afirma una racionalidad perfecta, por lo que supone que mientras se cumplan sus requisitos se está actuando racionalmente. Esto desde una postura ética puede tener problemas.

La tesis defendida a lo largo de este capítulo consiste, como ya se ha dicho, en rechazar la visión de una racionalidad perfecta, paramétrica, sin plantear una posición contraria de irracionalidad, manteniendo los puntos más destacados de la teoría y repensarlos de manera que pueda abarcar aspectos éticos donde se presente también la importancia de los demás al momento de tomar una decisión; asimismo, se busca llevar al ámbito práctico, es decir, realizar un autoescrutinio de estos elementos que componen la racionalidad y reflejarlo a través de las acciones realizadas.

Uno de los elementos que vale la pena destacar en este punto es la discusión que se centra en la relación de la razón con las normas establecidas socialmente. A lo largo de este capítulo se discutió que hay ciertos aspectos de la racionalidad que resultan ambiguos y esto ha llevado a elaborar diferentes planteamientos con respecto a ella. La teoría de juegos, entendida como una extensión de la TER, puede sostener dos tesis principales: por un lado, la visión estratégica de la toma de decisiones y, por otro lado, los límites previamente señalados por las reglas. Seguir las reglas por sí mismas ya implicaría guiar las acciones de manera racional, ¿por qué eso no es suficiente y, asimismo, porque no se puede asimilar el hecho de acatar una regla con el ser racional?, ¿en caso de sólo acatar las reglas para guiar las acciones, no implicaría ya una cierta irracionalidad? Esto será un poco de lo que se analizará en el siguiente capítulo. Del mismo modo, como se verá en el tercer capítulo, dentro de esta idea cabría preguntarse, ¿cómo se relacionan las normas establecidas socialmente con las emociones?

En el próximo capítulo se desarrollará la racionalidad desde un punto de vista estratégico, que ya toma en cuenta los aspectos cambiantes del ambiente y las decisiones tomadas por

otros agentes. Se verá, en este caso, cómo algunas situaciones pueden llevar al agente a no tomar la decisión que, aparentemente, sería la más racional. Por otro lado, se especificarán algunos de los elementos de la teoría que sostiene esto y se propondrá el análisis de un caso particular, que permitirá relacionar ambos capítulos a la tesis principal de este trabajo.

## Capítulo 2

### La teoría de juegos

[...] el “árbol del análisis”, tan recomendado por los entrenadores, comenzó a crecer de manera monstruosa. No sé por qué, pero en ese momento recordé la célebre poesía infantil de Chokovski: ¡Oh, qué difícil es el trabajo de sacar a un hipopótamo del pantano! No podría explicar en virtud de qué asociación ese hipopótamo se metió en el tablero, pero la verdad es que, mientras los espectadores creían que estaba analizando la posición, yo pensaba en cómo podría sacarse a un hipopótamo del pantano. [...] Después de numerosos intentos no encontré ningún método aceptable de sacarle del pantano, y pensé con amargura: “¡Pues que se ahogue!”

Mikhail Tal (citado en Amster y Pinasco, 2014, p. 116)

Mikhail Tal, campeón mundial de ajedrez a principios de los sesenta, cuenta en una entrevista que se le presentó el pensamiento citado en el epígrafe durante una partida contra Evgeni Vasiukov en 1964. Le tocaba tirar y comenzó a visualizar un montón de movimientos posibles, hasta que su pensamiento se perdió en lo anterior. Al igual que con la imposibilidad de sacar al hipopótamo del pantano, se dio cuenta de que era imposible calcular cuántas variantes había por el sacrificio de una de sus piezas durante su jugada. Tiró, olvidándose de ese asunto y, más tarde, ganó. Tal, en esa misma entrevista, comenta que se sorprendió al otro día, al leer en la prensa que había hecho un sacrificio de pieza muy acertado y bien calculado después de analizar su posición por cuarenta minutos.

Esta anécdota, a modo de introducción, muestra algunos de los argumentos que se analizarán a lo largo del capítulo; tal y como se vio en el capítulo anterior, la TER estándar llegó a verse como una metodología que planteaba la idea de una racionalidad perfecta, y se analizaron algunas ventajas y desventajas que esto planteaba. Debido a esta falta de plausibilidad de una racionalidad perfecta, se llevaron a cabo discusiones en torno a qué tan efectivo era seguir utilizando la definición de la economía moderna del sujeto económico como un ser racional. La teoría de juegos surge, por un lado, como una extensión de la TER

al proponer cambiar el modelo de una racionalidad paramétrica en una racionalidad estratégica; no obstante, por otro lado, ésta también se ha presentado como un enfoque en el que, a través de diversos fundamentos lógicos, se consigue demostrar que a veces “la estrategia más racional consiste en ser irracional” (Amster y Pinasco, 2014, p. 10). El ejemplo citado más arriba demuestra lo anterior a través del pensamiento del campeón de ajedrez, Mikhail Tal, pues antes de llevar a cabo un análisis sobre el juego pensó en una poesía infantil; del mismo modo, el ejemplo presenta, a través del ajedrez, el elemento con el que se llevan a cabo los análisis de la teoría: los juegos.

Al principio de este capítulo se explicará qué es la teoría de juegos y cómo es comprendida en tanto que extensión de la TER. Esto llevará a desarrollar lo que es la racionalidad estratégica y, al mismo tiempo, se analizará la propuesta contraria a la racionalidad, es decir, la irracionalidad. Se rechazará esta última, dado que la propuesta de esta tesis es que, pese a que no existe una garantía absoluta en la racionalidad para llevar a cabo una buena elección, tampoco podemos negarla; además, dado que a través de ella no es posible tener total seguridad acerca de qué es una buena toma de decisiones, necesitamos de un elemento que la complemente en el análisis: las emociones.

La teoría de juegos, como se verá, surge en principio como un modelo matemático para el estudio de la toma de decisiones, por lo que se presenta a través de una serie de fórmulas y matrices. Se desarrollarán, por ello, algunos de los elementos previamente explicados de la TER, a través de estas formalizaciones; se mostrarán en primer lugar, aquéllos que utiliza la racionalidad paramétrica, analizados en el capítulo anterior, y, en segundo lugar, los que pueden comprenderse dentro de la racionalidad estratégica. Dado que la tesis presentada desarrolla un aspecto menos formal, es importante señalar que este apartado sirve, por un lado, como una guía para quienes se encuentren interesados en este aspecto y, por otro lado, como muestra de los antecedentes y características de donde parte esta tesis.

Más adelante se desarrollará un breve análisis con respecto a los diferentes tipos de juego, y se hará énfasis en los denominados “juegos cooperativos de suma no nula”, en los que los resultados obtenidos no tienen por qué ser equitativos para todos los participantes. Dentro de estos últimos se encuentra uno de los ejemplos más famosos de la literatura sobre teoría de juegos, el dilema del prisionero. Dicho dilema ha sido utilizado constantemente en diversos

estudios que van desde la psicología, hasta cuestiones de economía o ciencias políticas. Se explicará en qué consiste el dilema, y se esbozarán las respuestas que han sido planteadas para su solución. Lo central de este análisis es que aquí se replanteará la discusión entre la racionalidad y la irracionalidad, elaborada al principio del capítulo, debido a que la teoría de juegos parece presuponer que la respuesta racional sería una, cuando se trata de resolver el conflicto entre cooperar o no hacerlo, y, no obstante, al evaluar la propuesta parece no ser la respuesta esperada y, por ende, resolverlo a través de la respuesta sugerida como irracional; de ahí que sea un dilema. Del mismo modo, este análisis permitirá comprender lo que se discutirá al final del tercer capítulo, dado que se utilizará este mismo dilema para afirmar que esta respuesta, en apariencia irracional, no lo es si añadimos a las emociones como parte importante de una evaluación sobre las decisiones o acciones tomadas.

Uno de los elementos importantes dentro de la teoría de juegos es, precisamente, que los juegos presuponen ciertas reglas para que puedan llevarse a cabo. En este apartado las reglas de los juegos se propondrán como analogía de las normas sociales. Las reglas en los juegos afectan directamente el tipo de acciones que pueden llevarse a cabo, *verbigracia*, la pieza del caballo en el ajedrez se mueve de una forma totalmente diferente al de un alfil, un peón o una torre. Marcan ciertos límites, tanto para que el objetivo se cumpla como para que no se hagan acciones que no correspondan al juego. Dentro de las sociedades, como se verá, las normas sociales, del mismo modo, afectan directamente las decisiones tomadas por los individuos, logrando que en ciertas ocasiones se comporten de una manera específica y no actúen de tal o cual modo. En el tercer capítulo se discutirá cómo las decisiones tomadas por normas sociales se enfrentan o relacionan con la cuestión de las emociones.

Al terminar esto, se llevará a cabo una breve recapitulación de lo discutido a lo largo de este capítulo.

### *2.1. La teoría de juegos: Racionalidad o irracionalidad*

La teoría de juegos surge aproximadamente en 1944, con la publicación de *The Theory of Games and Economic Behaviour*, de John von Neumann y Oskar Morgenstern; en ésta, ellos definen un juego como “cualquier interacción dada entre agentes que se rigen por un conjunto de reglas, especificando los movimientos posibles para cada participante, y un conjunto de resultados, para cada combinación posible de movimientos” (Hargreaves y Varoufakis, 1995,

p. 1).<sup>10</sup> En otras palabras, la teoría de juegos es una teoría acerca de la toma de decisiones, del cómo se plantean los objetivos que desean alcanzarse, los resultados posibles para cada alternativa, así como la *decisión* llevada a cabo (Davis, 1986; Di Castro, 2009). La teoría de juegos, como se verá en el siguiente apartado, retomó algunos de los conceptos más importantes de la TER, tales como la maximización, la utilidad, la intencionalidad, etcétera. Sin embargo, se enfocó en desarrollar una de las cosas que el sujeto económico dispone para llevar a cabo una decisión bajo ciertas restricciones: un conjunto de estrategias posibles. Davis (1986, p. 27) afirma que “una estrategia significa un plan de acción completo que describe cuáles serán las reacciones de un jugador ante cualquier circunstancia posible”, es decir, cuenta con un conjunto de acciones posibles y, cada uno, tendrá a su vez determinados resultados posibles.

La teoría de la elección racional presuponía que el sujeto era paraméricamente racional, es decir, que el medio en el que llevaba a cabo la toma de decisiones era constante; con la teoría de juegos el individuo se comprende como *estratégicamente* racional, es decir, toma en cuenta que el entorno es cambiante y que, al igual que él, existen otros individuos que también son estratégicamente racionales; a través de este análisis se puede dar cuenta de que los resultados de un análisis paramétrico puede llevar a resultados irracionales, dado que, aún pese a que los resultados sean los racionalmente esperados de manera individual, podrían ser incongruentes a nivel colectivo. Di Castro (2009, p. 58) afirma que:

La teoría de juegos analiza las decisiones estratégicas con base en los siguientes cuatro elementos:

- a) El grupo de agentes,
- b) Las elecciones o estrategias posibles para cada uno de los agentes,
- c) El estado del mundo producido por la elección de los agentes, y
- d) Las preferencias sobre los estados del mundo de cada uno de los agentes.

El individuo estratégicamente racional, por ende, debe tomar en cuenta, por un lado, las intenciones y los resultados que esperan los demás individuos y, por otro lado, ser consciente de que estos individuos basan sus intenciones tanto en sus expectativas individuales como en las que esperan de éste. En los primeros análisis de la teoría de juegos esto llevaba a una cadena infinita, debido a que lo anterior se cumplía de la siguiente manera:

---

<sup>10</sup> La traducción es mía, al igual que todas las citas extraídas de esta publicación a lo largo del texto.



- (a) Cada persona es instrumentalmente racional.
- (b) Cada persona sabe (a)
- (c) Cada persona sabe (b)
- (d) Cada persona sabe (c). Y así *ad infinitum*. (Hargreaves y Varoufakis, 1995, p. 24)

El término utilizado en teoría de juegos para esto es *common knowledge*, y se entiende que, si *X* y *Y* saben que  $\alpha$ , entonces se puede generar un enunciado infinito, dado que “*X* sabe que *Y* sabe que *X* sabe que... *Y* sabe que *X* sabe...  $\alpha$ ”. No obstante, con el trabajo de Nash se llegó a romper con la idea de una regresión sin termino, al introducir la noción de *punto de equilibrio*, pues ésta permitió obtener una solución única y sin regresión definida por el juego. Un punto de equilibrio es un conjunto de estrategias individuales óptima de cada sujeto con respecto de otro o, en otras palabras:

“Definición: Un conjunto de estrategias racionalizables (una para cada jugador), se encuentra en un **equilibrio de Nash** si su realización confirma las expectativas de cada jugador acerca de la elección del otro. En otras palabras, las **estrategias de Nash** son las únicas racionalizables que, si se ejecutan, confirman las expectativas sobre las que se basaron. Esta es la razón por la que, a menudo, se denominan como estrategias de auto-confirmación, porque se puede decir que este concepto de equilibrio requiere que las creencias de los jugadores estén consistentemente alineadas” (Hargreaves y Varoufakis, 1995, p. 53).<sup>11</sup>

En el siguiente apartado se complementará lo anterior a través de la formalización del equilibrio de Nash. Lo importante a rescatar aquí es que el punto de equilibrio, o equilibrio de Nash, permite analizar la formación de las creencias de los individuos acerca de cómo interactúan unos con otros estratégicamente (Elster, 1997, p. 49).

Muchos análisis dentro de la teoría de juegos se detienen a hablar específicamente de qué significa llevar a cabo una elección a través de estrategias. Si bien antes se le ha denominado al individuo como estratégicamente racional, la teoría de juegos supone que las estrategias tomadas no tienen por qué ser racionales. El agente puede disponer de varias estrategias para lograr alcanzar su objetivo, pero unas pueden ser más deficientes que otras o, al contrario, pueden alcanzar de una manera más óptima el resultado esperado. De esto se deduce que

---

<sup>11</sup> Las negritas son suyas.

distintos agentes ante una misma situación, e incluso uno mismo en diferentes momentos, pueden asumir diferentes estrategias, es decir, tomen decisiones diferentes.

Un ejemplo clave para comprender esto se puede tomar de la literatura. En “Las mil y una noches” (Anónimo, 2016), se puede leer la siguiente historia:

Se cuenta de un hombre de Bagdad que vivía en completo desahogo y tenía grandes riquezas. Pero éstas se le agotaron, su situación cambió y se quedó sin nada consiguiendo comer sólo a costa de inauditos esfuerzos. Cierta noche mientras dormía cohibido y amedrentado vio en sueños a una persona que le decía: “¡Tu fortuna se encuentra en El Cairo! ¡Ve, corre a buscarla!”

Emprendió el viaje a El Cairo, llegó al atardecer y fue a dormir en una mezquita. Cerca de la mezquita había una casa y Dios (¡ensalzado sea!) dispuso que una partida de ladrones entrase en la mezquita para asaltar la casa. Los habitantes de ésta se despertaron al oír el movimiento de los ladrones y empezaron a chillar. El valí de la ciudad y sus hombres acudieron a auxiliarles. Los ladrones huyeron. El valí entró en la mezquita y encontró dormido al bagdadí. Le detuvo y le hizo azotar con golpes muy dolorosos hasta que estuvo a punto de morir. Le encarceló y le tuvo tres días en prisión. Después le hizo comparecer y le preguntó: “¿De qué país eres?” “¡De Bagdad!” “¿Y qué motivos te han traído a El Cairo?” “He visto en sueños una persona que me decía: “¡Tu fortuna se encuentra en El Cairo! ¡Vete!” Al llegar a El Cairo me he dado cuenta de que la fortuna prometida eran los azotes que me has mandado dar.” El valí se río de buena gana dejando al descubierto sus molares. Le dijo “¡Hombre de poco entendimiento! Yo he visto en sueños tres veces a una persona que me decía: “Hay una casa de Bagdad situada en tal barrio y cuyo aspecto es éste. En su patio hay un jardincillo y debajo del surtidor se encuentran riquezas enormes. Ve allí y cógelas”. Yo, a pesar de esto, no me he movido y tú, tonto, has emprendido el viaje de una ciudad a otra por una visión que has tenido en el curso de una pesadilla”. Le dio a continuación unos dirhemes y le dijo: “Utilízalos para regresar a tu ciudad”.

Sahrazad se dio cuenta de que amanecía e interrumpió el relato para el cual le habían dado permiso.

Cuando llegó la noche trescientos cincuenta y dos, refirió:

—Me he enterado, ¡oh rey feliz!, de que [el hombre] tomó los dirhemes y regresó a Bagdad, pues la casa que el valí le había descrito era la suya propia. Al llegar a su domicilio cavó debajo del surtidor y encontró una gran riqueza. De este modo Dios le dio un gran tesoro. Éste es un caso prodigioso.

A través de este ejemplo uno se puede preguntar acerca de cuál estrategia fue la mejor, si la utilizada por el valí o la del hombre de Bagdad, independientemente de que se trata de una obra de ficción. El matemático Ivar Ekeland (1992) escribió al respecto que, pese a que ambos habían tomado una estrategia distinta, habían quedado satisfechos, dado que sus análisis y las estrategias tomadas se habían confirmado a través de cómo se les presentaron

los hechos. Por un lado, el valí siempre estaría feliz al burlarse de personas como el hombre de Bagdad que, tras creer en su sueño había viajado en vano; por otro lado, el hombre de Bagdad estaría feliz por haber seguido ese sueño. Desde un punto de vista o del otro, ambos habían tomado una decisión correcta.

Este tipo de casos llevó a algunos a plantearse seriamente la cuestión de la irracionalidad. Amster y Pinasco (2014, p. 10), por ejemplo, afirman en su prefacio que a través de la teoría de juegos se puede demostrar que, en algunos casos, la estrategia más racional puede ser aquella que consiste en ser irracional. Elster (2014), cita a Pascal, quien afirma que “no hay nada tan conforme a la razón como ese rechazo de la razón”. En el capítulo anterior se argumentó que una de las mayores críticas a la teoría de la elección racional consistía en que los seres humanos no son perfectamente racionales, sino que cometen muchísimos errores, tienen muchísimos deseos frustrados, proyectos que han terminado siendo un fracaso, tienen creencias falsas, se auto-engañan, etcétera; el análisis se puede dar en sentido contrario: el individuo estratégicamente racional puede cometer errores, realizar análisis incompletos o no ser completamente racional puesto que si lo fuera, sólo haría falta lo paramétricamente racional y, por ende, no se tomaría en cuenta el hecho de que las estrategias tomadas por los demás, así como el entorno cambiante, afectan también el curso de acción de las decisiones que se llevan a cabo.

La teoría de juegos ha tomado diferentes cursos de acción y, para cada uno de ellos, ha tenido un juego diferente como respuesta. En algunos casos se ha podido responder sin mayor uso de la razón; hay teoría de juegos con respecto a cuestiones evolutivas, por ejemplo, donde se puede explicar un fenómeno en los animales sin que ellos lleven a cabo una toma de decisiones. En otros casos, por el contrario, las respuestas son más difíciles de dar, por lo que se trata de buscar una mejor estrategia para responderla. Pese a que existe esta enorme tendencia por aceptar la irracionalidad, la propuesta planteada aquí es, de hecho, que el hombre puede llegar a ser imperfectamente racional, y eso no ser un problema para tomar una decisión que pueda plantearse como correcta. El hecho de que muchas acciones llevadas a cabo diariamente no son evaluadas racionalmente no quiere decir que no hayan sido evaluadas con anterioridad; se aprende a caminar, se aprende que el fuego quema si acercas la mano. Llevar a cabo acciones que pueden plantearse como irracionales, muchas veces

puede llevar a una mala toma de decisión, aunque es cierto que no siempre es el caso. La imperfección racional es aceptar que no siempre se podrá llevar a cabo una evaluación completa para decidir cuál es la mejor estrategia y, del mismo modo, cabría preguntarse si todavía sería importante hablar acerca de cuál es “la mejor”. Dado que la racionalidad no es suficiente para llevar a cabo una evaluación crítica con respecto a la toma de decisiones, haría falta un elemento, que se analizará a lo largo del tercer capítulo.

## 2.2. Formalización de la teoría de juegos

Formalmente, la teoría de juegos retoma algunos de los elementos de la elección racional.<sup>12</sup> Como se ha dicho anteriormente, la elección racional comienza a partir de la idea de que los individuos tienen preferencias y llevan a cabo una elección de acuerdo a ellas. Entonces:

Supongamos que  $X$  es un conjunto de elecciones posibles. A través de un modelo de elección se puede especificar que  $X \subset \mathbb{R}^n$ , esto es, que dentro de las elecciones posibles hay  $n$  bienes diferentes (v.gr. pizza, tacos, hamburguesas, etc.) y si  $x \in X$ , entonces  $x = (x_1, \dots, x_n)$  especifica cantidades de cada tipo de bien.

Ahora bien, considerando a un agente económico, definamos las preferencias débiles del agente sobre el conjunto  $X$  de la siguiente manera:

$$x \succeq y \Leftrightarrow "x \text{ es al menos tan bueno como } y"$$

Digamos que  $x$  es *estrictamente preferida* a  $y$ , o  $x \succ y$ , si  $x \succeq y$  pero no  $y \succeq x$ . Digamos que el agente es *indiferente* entre  $x$  y  $y$ , o  $x \sim y$ , si  $x \succeq y$  y  $y \succeq x$ .

En la teoría de la elección racional, se tienen dos supuestos fundamentales, a saber, que las preferencias deben ser completas y deben ser también transitivas.

**Definición 1:** Una relación de preferencia  $\succeq$  sobre  $X$  es *completa* si para todo  $x, y \in X$ , ya sea que  $x \succeq y$  o bien que  $y \succeq x$ , o ambos.

La relación de completitud satisface el hecho de que, si un agente se encuentra frente a dos opciones, éste necesariamente tendrá una opinión al respecto, ya sea acerca de cuál

---

<sup>12</sup> La mayor parte de la formalización presentada en este apartado se retoma de Levin y Milgrom (2004), y de Hargreaves y Varoufakis (1995).

prefiere o bien si es indiferente. La definición anterior no excluye la posibilidad de que  $y = x$ , por lo que la completitud implicaría que  $x \succeq x$ , esto es, que la relación  $\succeq$  sea también *reflexiva*.

**Definición 2:** Una relación de preferencia  $\succeq$  sobre  $X$  es *transitiva* si siempre que  $x \succeq y$  y  $y \succeq z$ , entonces  $x \succeq z$ .

La transitividad implica que las preferencias débiles del agente pueden ser cíclicas sólo si son indiferentes, es decir que si se prefiere  $x$  a  $y$ ,  $y$  se prefiere a  $z$ , pero  $z$  se prefiere a  $x$ , entonces  $x \sim y \sim z$ .

Hasta el momento, se han presentado tres elementos de la TER: completitud, reflexividad y transitividad. Cuando todos estos elementos se sostienen, se afirma que existe una ordenación de preferencias bien definida para el agente. Muchos análisis introducen todos estos elementos en un mismo conjunto denominado *maximización de utilidad y elección consistente*, e introducen un cuarto elemento, la *continuidad*, que se puede entender de la siguiente manera:

Para cada  $x, y, z$ , si  $x \succ y \succ z$ , entonces debe existir algún *compuesto* de  $x$  y  $z$ ,  $a$ , el cual da una utilidad igual  $y$ , es decir,  $a \sim y$  y nuestro agente es indiferente entre ellas. Cuando esto se sostiene, esta ordenación de preferencias se puede representar por una función de utilidad, cuya expresión matemática es  $\mathcal{U}(x)$ . Por ejemplo, si  $x$  es “cinco chocolates con almendras,  $y$  es “tres chocolates oscuros” y  $z$  es “diez chocolates blancos”, debe haber una combinación, compuesto, de  $x$  y  $z$ ,  $a$ , como “dos chocolates con almendra y cuatro chocolates blancos”; dicha combinación sería indiferente de  $y$ .

Antes de continuar es importante hacer una acotación. Dentro de un análisis, a cada letra se le asigna un número que indica la utilidad que representa para el sujeto, es decir, el valor de preferencia. Estos números de utilidad se asignan arbitrariamente en tanto que de  $x$  y  $y$  siempre se sostenga que  $x \succ y$ . A estos números de utilidad se les llama *utilidad ordinal*. La utilidad ordinal, no obstante, no habla nunca de la intensidad de una preferencia, por lo que a veces resulta insuficiente.

La *utilidad cardinal* otorga esa información, por lo que ésta se puede entender en tanto que se encarga de analizar la maximización de la utilidad esperada. Esto se lleva a cabo a través de asignar una probabilidad porcentual a cada una de las preferencias presentadas.

La utilidad cardinal, por ende, introduce una modificación al elemento de continuidad y, al mismo tiempo, introduce dos condiciones más como elementos de la teoría.

En la modificación de la continuidad se consideran tres posibilidades,  $a, b, c$ , y se supone que el agente prefiere la primera a la segunda y la segunda a la tercera. Se supone la existencia de una probabilidad  $\rho$ , tal que, para la posibilidad  $a$  se le asigna la probabilidad  $\rho$ , y a la posibilidad  $c$  se le asigna la probabilidad  $1 - \rho$ , entonces el agente podría llegar a estar igualmente feliz con esta situación como lo habría estado con  $b$ .

Con el *incremento de preferencia con probabilidad* se entiende que, si  $a \succ b$  y  $e = (a, b; \rho_1, 1 - \rho_1)$ ,  $f = (a, b; \rho_2, 1 - \rho_2)$ , entonces  $e \succ f$  sólo si  $\rho_1 > \rho_2$ .

En la condición de *independencia*, para tres posibilidades,  $a, b, c$ , si  $a \succ b$ , entonces existe una probabilidad  $\lambda$ , tal que una probabilidad mezclada  $(\lambda, 1 - \lambda)$  de  $a$  y  $b$  debe ser por lo menos tan buena como la probabilidad mezclada  $(\lambda, 1 - \lambda)$  de  $a$  y  $c$ . Escrito de otro modo:  $(a, b; \lambda, 1 - \lambda) \succeq (a, c; \lambda, 1 - \lambda)$ .

Si se cumplen las seis condiciones (completitud, reflexividad, transitividad, continuidad, incremento de preferencia e independencia), el agente cumple el argumento de la TER, es decir, a través de su racionalidad paramétrica cumple la maximización de su función de utilidad esperada.

Dado que la teoría de juegos se enfoca en el análisis de las estrategias, las asume “como si” tuvieran una función de utilidad cardinal, asignando probabilidad, para actuar de manera tal que se cumpla la maximización de utilidad esperada.

Uno de los problemas que se planteó dentro de la teoría de juegos fue cómo se actualizan las creencias a partir de la presencia de nueva información. Para resolver este problema se introdujo la *regla de Bayes* (Hargreaves y Varoufakis, 1995, p. 19). Esta regla afirma que la probabilidad de que el evento A haya ocurrido dado que el evento B ha sido observado, es una probabilidad condicional, es decir,  $\Pr(A|B)$ , y es expresada de la siguiente manera:

$$\Pr(A|B) = [\Pr(B|A) \Pr(A)] / [\Pr(B|A) \Pr(A) + \Pr(B|\neg A) \Pr(\neg A)]$$

Donde  $\neg A$ , dentro de la formalización, quiere decir que el evento A no ocurrió. El evento B es, por tanto, la nueva información.

Otro de los aspectos importantes a destacar de la teoría de juegos es su análisis acerca de cómo las expectativas de los otros afectan las del agente. En el apartado anterior se comentó acerca del *common knowledge*. En la teoría se le conoce por las siglas CKR, *common knowledge of rationality*, y se basa en cómo, a través de la racionalidad instrumental, se comprende la relación de un individuo con otro, por lo que a través de este concepto la teoría de juegos sigue siendo limitada. Para resolverlo, muchos asumen no sólo un CKR, sino también un CAB, o bien *consistent alignment of beliefs* (Hargreaves y Varoufakis, 1995, p. 23). El CAB afirma que una persona no instrumentalmente racional puede esperar que otra persona similarmente racional, que tenga la misma información que la primera, podría desarrollar un resultado similar a través de un proceso de pensamiento distinto. Algunos teóricos de juegos han criticado esto último (Bernheim, 1984; Kreps, 1990; Pearce, 1984); no obstante, parece que CAB puede sostenerse.

En CAB se da el caso de que, si una persona decide llevar a cabo una acción, y la otra persona acepta que la primera es racional, entonces la primera persona ya tiene asumido de antemano la idea de que su movimiento será racional y se encontrará bajo la misma línea de pensamientos que la segunda. En este sentido, si uno conoce los planes del otro, sus creencias estarán en el mismo nivel expresado anteriormente, es decir que, dado ese conocimiento acerca del otro, las creencias de uno no se modificarán, del mismo modo que el otro sabrá que las creencias de uno no se modificarán y, por lo tanto, sostendrá sus planes. Es esto último por lo cual se conoce como alineación consistente de creencias.

La justificación por la cual se sostiene el paso de CKR a CAB, es la *doctrina Harsanyi*, la cual afirma que, si dos individuos racionales poseen la misma información, entonces deben realizar la misma inferencia y, por lo tanto, llegar a las mismas conclusiones de manera independiente (Hargreaves y Varoufakis, 1995, p. 25).

Dos elementos más que son importantes de rescatar para esta tesis son, por un lado, los modelos mediante los cuales se muestra cómo se efectúa un análisis de juego, y el equilibrio de Nash.

Los modelos utilizados por la teoría de juegos son de dos tipos; por un lado, se tiene un *diagrama de árbol* (Fig. 2.1), cuyas estrategias se ven representadas mediante las ramificaciones que se construyen por éstas y, por otro lado, se tienen las *matrices* (Fig. 2.2), son las más utilizadas y se representan mediante una tabla donde se cruzan los resultados posibles para una acción, según la estrategia que deciden utilizar los individuos.

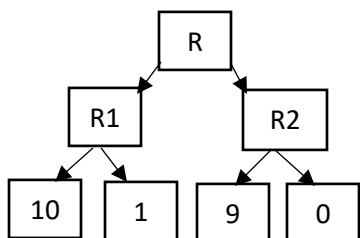


Fig. 2.1

	(B)		
(A)		(a)	(b)
	(a)	1	1
	(b)	0	2

Fig. 2.2

John Nash, en 1950, introdujo el problema de las estrategias racionalizables múltiples. El ejemplo clásico para demostrar cómo se da el equilibrio de Nash, ya antes mencionado, se muestra a continuación (Hargreaves y Varoufakis, 1995, pp. 51-53):

	C1	C2	C3
R1	+100, 100	0, 0	50, 101
R2	50, 0	+1, 1	60, 0
R3	0, 300	0, 0	+200, 200

Fig. 2.3

Jugaré R1 porque espero que C juegue C1. ¿Por qué espero esto? Porque C espera erróneamente que juegue con R3 (en vez de R1, con el que pretendo jugar). Se puede preguntar por qué yo pienso que él piensa eso. Bueno, quizá porque él espera que piense erróneamente que jugará con C3, cuando en realidad espero que juegue C1. De hecho, si él supiera que planeo jugar R1, él podría jugar C3. Pero él no sabe esto y, por esa razón, y dadas mis expectativas, R1 es la elección correcta para mí. Si él supiera que jugaré con R1, no lo haría. Sin embargo, mi suposición es que él espere que jugaré R3, pensando en que espero que jugará C3. La realidad es que espero que juegue C1, y yo jugaré R1.



La línea argumentativa se puede dar de la siguiente manera:

*R: R1, porque*

R porque C: C1, porque

R porque C porque R: R3, porque

R porque C porque R porque C: C3, porque

R porque C porque R porque C porque R: R1 (y se vuelve a empezar)

Una línea argumentativa similarmente consistente podría surgir de R3:

*R: R3, porque*

R porque C: C3, porque

R porque C porque R: R1, porque

R porque C porque R porque C: C1, porque

R porque C porque R porque C porque R: R3 (y se vuelve a empezar)

Siguiendo con el ejemplo, se puede considerar una racionalización equivalente con R2:

Jugaré R2 porque creo que C jugará C2. ¿Por qué creo que jugará C2? Porque él piensa que jugaré R2, pensando que yo esperaré que juego C2. Etcétera.

Este último caso se presenta como un CKR; al mismo tiempo, se puede dar cuenta de la doctrina Harsanyi:

R: R2, porque

R porque C: C2, porque

R porque C porque R: R2 (y se vuelve a empezar)

R2 es, por tanto, la mejor estrategia, porque, además de plantearse con mucha simplicidad, “es la única estrategia que se sostiene por creencias que no suponen que el oponente se equivocará al esperar algo que R no tiene la intención de realizar” (Nash, 1951, pp. 286-95). La manera en la que se puede argumentar que R1 y R3 son racionales, es únicamente a través de que R piense que C ha llevado a cabo un proceso de pensamiento erróneo; esto conllevaría, por lo tanto, a pensar que el otro ha actuado irracionalmente. Por otro lado, Si se acepta CKR y la doctrina Harsanyi, en R2, se puede aceptar CAB, dadas las creencias de los individuos. Si se acepta CAB se aceptará, por lo tanto, el equilibrio de Nash, dado que (R2, C2) sería la

estrategia de ambos individuos; y, por tanto, el resultado de este análisis plantearía a los dos individuos actuando de manera racional. No obstante, es importante preguntarse si eso es, de hecho, el caso.

### 2.3. *Juegos cooperativos*

La teoría de juegos, si bien comenzó siendo una teoría matemática acerca de la toma de decisiones, su nivel explicativo, así como las ventajas que ofrecía, le permitió adentrarse en otras áreas, a saber, análisis en microeconomía, sociología, psicología, filosofía, incluso en biología, como algunas cuestiones evolutivas.

Existe una cantidad de juegos tan amplia, que resulta complicado hablar de todos ellos, incluso de algunos de los elementos de la teoría que se destacan más en unos que en otros. Algunos de ellos son, por sólo mencionar algunos: los *juegos de información incompleta*, en los cuales se analiza el elemento de la racionalidad estratégica que explica cómo se evalúan las estrategias en caso de que los jugadores no puedan recopilar o no puedan acceder a cierta información que podría afectar el resultado esperado; los *juegos dinámicos*, en estos juegos un jugador efectúa una jugada después del otro jugador, y se utilizan para tratar el problema acerca de cómo se modifican las estrategias de los jugadores si se conocen los movimientos previos del otro jugador; los *juegos de negociación (bargaining games)*, en los que se trata de producir un beneficio mutuo, a través de la cooperación, aunque si el acuerdo falla, el beneficio potencial no se lleva a cabo y ambos generan pérdidas; y los *juegos evolutivos*, que se centran en los fenotipos, es decir, los patrones de conducta de las especies (es importante señalar que no se trata del genotipo, es decir, de la base genética de la conducta).

Algunos juegos que no fueron mencionados contienen algunos elementos que se pueden esbozar en aquél que se utilizará para mostrar la tesis presentada en este trabajo: el dilema del prisionero; por ello se mencionan aparte.

Los *juegos repetitivos* se basan en un aspecto social, a saber, la interacción de las personas. En ocasiones es posible que la interacción entre dos o más personas no se dé más que una sola vez; en la vida diaria, por ejemplo, esto ocurre a menudo. Camino al trabajo, a la escuela, en la calle, uno siempre se encuentra con personas que probablemente no volverá a ver nunca

en su vida, pero con las que se puede tener una interacción. También se encuentra el aspecto contrario, es decir, existen muchas personas con las que se convive en varias ocasiones.

El trato en ambos casos suele ser muy diferente. El primer caso puede entenderse como un ejemplo de *juego finito*; uno puede imaginarse, v.gr., que se encuentra atorado en el tráfico, pues está a punto de pasar una carretera de dos vías que en algún momento se hace de una sola vía. Cabe añadir que se le hace tarde para llegar a un lugar importante. Bajo esta condición, y dado que desconoce al otro individuo, puede ocurrir que uno quiera pasar primero que el otro en vez de ceder el paso, siendo el caso. Debido a que es un evento importante, el conductor realiza dicha acción; sabe que esa no es la manera correcta de cruzar esa vía así que sólo lo hace en esa ocasión.

Los juegos repetitivos también pueden considerarse *finitos* cuando se ejecutan varias veces, pero que, no obstante, sólo se juegan un número determinado de veces; existen también los juegos *infinitos*, que pueden jugarse indefinidamente. Puede utilizarse el mismo caso de la carretera de dos vías para ejemplificarlos. En la versión finita del juego puede considerarse que se debe ir a un curso durante tres días, por lo que debe cruzar esa avenida sólo por esos días. Dado lo que pasó la primera vez, puede considerar salir más temprano de su punto de salida y tratar de seguir con el programa de tránsito y ceder el paso, o puede continuar bajo la idea de intentar pasar primero. En la versión infinita del juego se puede recrear esta escena, pero ahora este punto del viaje es necesario cruzarlo para llegar a su trabajo; dado que sabe que si no se sigue la regla de tránsito establecida para ese espacio puede ocasionar un tráfico mayor, y dado que sabe que ese es el recorrido que siempre utilizará, podría tratar de cambiar su estrategia y cumplir con lo establecido (si bien el ejemplo no es bueno para hablar de algo infinito, se puede entender claramente que se trata de que se llevará a cabo constantemente esta acción, sin saber en qué momento terminará). En los juegos repetitivos, por lo tanto, las opciones estratégicas del jugador se pueden ampliar de manera significativa.

En este tipo de juegos se pueden presentar también los casos de *suma cero*, así como los de *suma no-nula*. Los juegos de suma cero, por un lado, consisten sencillamente en que aquello que gana uno de los jugadores es igual a lo que pierde el otro. Es decir que, por ejemplo, si se hace una apuesta de 10 pesos, uno de los jugadores ganará 10 pesos mientras

que el otro perderá esa misma cantidad. El juego de suma no-nula tiene resultados parciales, por lo que lo que gana uno, podría no ser igual a lo que otro pierde, incluso puede ser el caso de que ninguno de los dos gane, o gane uno, pero a costa de algunas pérdidas. El ejemplo clásico de un juego de este tipo es uno denominado *halcones y palomas*. Se trata de que cada jugador tiene la posibilidad de elegir entre dos clases de conducta distinta, por buscar conseguir alimento, ya sea agresiva (halcón) o pacífica (paloma). Si ambos intentan ganar el recurso por la fuerza, cada uno logrará sacar una parte, pero se verá afectado por el esfuerzo ejercido por la batalla; si ambos intentan la vía pacífica, el recurso será dividido en partes iguales; si uno es agresivo y el otro no, el agresivo se llevará todo el recurso.

Los juegos repetitivos pueden considerarse como *juegos cooperativos* o como *juegos no-cooperativos*. Los juegos no-cooperativos consisten en tratar de sacar el mayor beneficio a nivel individual, mientras que los juegos cooperativos tratan de que el resultado final sea benéfico para ambas partes.

A través de estos ejemplos se puede ver cómo los elementos de un juego pueden mezclarse con los de otro. El dilema del prisionero, por ejemplo, se considera un juego cooperativo de suma no nula, que puede ser finito o infinito. Antes de profundizar en éste se explicará un poco más acerca de los juegos cooperativos.

Los juegos cooperativos tienen como base la *teoría de la cooperación*, la cual surge por la necesidad del sujeto de alcanzar sus objetivos individuales, excluyendo mecanismos que obliguen a cooperar con los demás. (Axelrod, 1986, p. 18). Axelrod afirma que postular de esta manera la teoría evita los problemas acerca de si la preocupación por el otro es real; no obstante, resalta que incluso a nivel individual se puede decir que dentro de los intereses propios se puede encontrar una preocupación por el otro. La teoría de la cooperación expone que existen hechos empíricos con los que se puede comprobar que la cooperación puede nacer incluso bajo condiciones hostiles, sin la necesidad de que exista una relación de amistad de por medio, acuerdos normativos o semejantes. Un análisis de esto se expone a través de una de las obras del sociólogo Ashworth (1980), a través de cartas y diarios de personas que se encontraron luchando durante la Primera Guerra Mundial. Durante la guerra de las trincheras se llevó a cabo en algunos frentes un sistema de cooperación que se ha denominado “Vive y deja vivir”. Sin ningún acuerdo previo, cansados de la guerra, muchos soldados

dejaron de disparar al bando enemigo. Se cuenta que incluso podían asomarse y salir a caminar sin recibir un disparo. Los ataques, que debido a los altos mandos no podían dejar de ocurrir, se encontraban calculados, pues la artillería disparaba torpemente haciendo muchísimos estragos, pero ninguna muerte, lo que a su vez demostraba el daño que podía hacer si se rompía esa tregua no acordada; los francotiradores también hacían su parte, disparaban a blancos específicos, demostrando lo que podían ser capaces de hacer. La cooperación se dio a través de la reciprocidad. La estrategia fue una moderación hacia el bando contrario, por un lado, y la respuesta inmediata a cualquier provocación. Podría decirse que había una preocupación recíproca. Al final de la guerra los altos mandos intervinieron en esas zonas en las que no ocurría ningún avance ni retroceso, obligando a llevar a cabo ciertas acciones que provocaran bajas; esto generó otro tipo de reciprocidad, esto es, ya no entre ambos bandos, sino hacia los compañeros que morían por estas causas. Se asesinaba pensando en honrar al aliado muerto; una muerte llevó a otra, y la cooperación desapareció.

Axelrod (1980, p. 172) afirma:

El verdadero fundamento de la cooperación no es la confianza, sino la perdurabilidad de la relación. Cuando las condiciones son idóneas, los jugadores pueden llegar a cooperar uno con otro, descubriendo las posibilidades de mutua remuneración que la cooperación ofrece, ya sea por tanteo, ya por imitación de otros jugadores venturosos, o incluso a través de un ciego proceso de selección de las estrategias más prósperas, que va eliminando las menos venturosas.

Así mismo, Axelrod (1980, p. 96) señala que el desarrollo de la cooperación plantea tres cuestiones importantes: la robustez, la estabilidad y la viabilidad inicial. La primera analiza qué tipo de estrategias pueden plantearse como viables en medio de un contexto en el que ya existen algunas estrategias estables. La segunda analiza cómo puede resistir una estrategia ya establecida ante otras que pretenden introducirse. La tercera evalúa la posibilidad de que una estrategia estable y robusta se pueda sostener en un ambiente en el que predomina la no-cooperatividad.

Amster y Pinasco (2014, p. 221-2), ya desde un lenguaje más adentrado en la teoría de juegos, ofrecen tres respuestas posibles acerca de qué estrategias son posibles desde la cooperación.

*Equilibrios correlacionados:* acordamos qué estrategia usará cada uno, pero no como resultado de dos estrategias individuales, sino un sorteo entre los resultados finales.

*Cooperación sin transferencia de utilidades:* acordamos dos estrategias individuales para utilizar, y cada uno recibe lo que indique el resultado del juego.

*Cooperación con transferencia de utilidades:* acordamos qué estrategia utilizará cada uno y cuánto deberá pagarle un jugador al otro.

Los equilibrios correlacionados sugieren llevar a cabo una propuesta en la que ambos jugadores asignen probabilidades a cada uno de los resultados y se resuelva con cuál se jugará, sin pactar con qué se jugará de manera inmediata. La cooperación sin transferencia de utilidades específica continuar con las reglas del juego, pero pactando con qué jugará cada uno, para saber cuál será la ganancia y cuál la pérdida. La cooperación con transferencia de utilidades es enormemente útil si el objetivo a conseguir es mayor individualmente que lo que se podría lograr si ambos jugadores escogieran una estrategia; en este caso se pacta de antemano quién ganará y quién perderá, para sacar el mayor provecho del premio, dividiéndose la ganancia después de efectuado el juego.

Es importante señalar en este apartado que el tema de la cooperación será de vital importancia para el análisis elaborado al final de este trabajo, debido a que la relación que hay entre la racionalidad y las emociones, al llevar un debido análisis con respecto a qué hacer cuando se presenta una situación en la que los beneficios de dos personas se encuentran en conflicto, permitirá tomar una elección que podría ser cooperativa, tratando de beneficiar en mayor medida a ambos jugadores.

### *2.3.1. El dilema del prisionero*

El dilema del prisionero surge en 1950, inventado por Merrill Flood y Melvin Dresher, formalizado poco después por A.W. Tucker (Axelrod, 1986, p. 19). Consiste en que dos personas sospechosas de cometer un crimen serán interrogadas por separado por un juez. Lo que se sabe acerca del crimen es que no hay pruebas suficientes. Las personas sospechosas, al momento de su captura, portaban armas y, no obstante, se necesita de la confesión de al menos uno para confirmar la participación en el crimen. Para que se lleven a cabo las confesiones, el juez les ofrece el mismo trato:

Si confiesas haber participado en el crimen, y tu compañero no lo hace, entonces como premio te dejo libre, y tu compañero irá 20 años preso. En cambio, si ambos confiesan irán los dos a la cárcel por 5 años. Si ninguno de los dos decide confesar, entonces no tendré pruebas suficientes de que ustedes cometieron el crimen, pero los enviaré a la cárcel por un año por portación de armas. A tu compañero le hago la misma oferta.

La matriz, por lo tanto, queda de la siguiente manera:

	C	$\neg C$
C	(-5, -5)	(0, -20)
$\neg C$	(-20, 0)	(-1, -1)

Fig. 2.4

Antes de ofrecer las respuestas posibles, como se hace en la mayoría de los análisis del dilema, se analizará la forma fundamental del dilema, sin ninguna alteración que pueda modificar la respuesta:

- a) Se excluyen mecanismos con los que se puedan llevar a cabo amenazas. Se excluyen para que el individuo no se vea obligado a optar por una respuesta en específico, sino que debe tomar en cuenta todas las estrategias posibles.
- b) No existe ningún medio por el que se pueda conocer cuál será la decisión del otro participante. Esto imposibilita la idea de tomar una decisión debido a que se conoce la respuesta del otro jugador. La información disponible acerca del otro jugador es únicamente aquélla que se pudo obtener de sus interacciones previas.
- c) No se puede suprimir al otro jugador, ni se puede evitar la interacción, habrá o no cooperación entre ambos.
- d) Los pagos del otro jugador no son controlables. (Taylor, 1976, pp. 69-73)

Siempre que se habla de los posibles resultados en el dilema del prisionero se menciona que es una situación paradójica. Si se hace un análisis profundo de los posibles resultados, la mejor opción es confesar, puesto que si el otro confiesa únicamente serán 5 años en prisión, y no se arriesgará a estar 20 años en prisión si el otro confiesa y uno no. Del mismo modo, si el otro decide no confesar, él estará en prisión y uno será liberado inmediatamente. Siguiendo un análisis a través del equilibrio de Nash se llega al mismo resultado: se debe confesar. La estrategia seguida por ambos jugadores conlleva al resultado del equilibrio (-5, -5). No

obstante, parece que, a pesar de ser la respuesta racional, no es la mejor, pues si ambos jugadores decidieran no confesar, la pérdida sería menor  $(-1, -1)$ . Dado que en el dilema el mejor resultado sería 0 y no  $-1$ , buscar la no confesión podría llevar a una posible “traición”.

La teoría de juegos expone una manera en la que se puede tomar  $(-1, -1)$ , como la respuesta que debería darse del dilema. Lo que se afirma es que en el dilema del prisionero no existe ninguna estrategia dominada, es decir, que lleve a un peor resultado para uno de ellos, sino que es equivalente. Además, los equilibrios no siempre son la respuesta más óptima. Si se responde que la mejor estrategia es  $(-1, -1)$ , es probable que se esté apelando al *óptimo de Pareto*. El óptimo de Pareto es un concepto utilizado en economía que especifica un punto de equilibrio; en este punto ninguno de los jugadores puede mejorar su situación actual sin afectar el bienestar del otro.

Anteriormente se mencionó que los resultados de un juego podrían modificarse si el juego podía repetirse finita o infinitamente. Con respecto al dilema del prisionero se han llevado a cabo torneos con los que se puede explicar fácilmente qué ocurre. En los torneos cada participante lleva una estrategia que cree que puede ganar. Algunas consisten en cooperar siempre (no-confesar), otras en no-cooperar (defección), otras en utilizar la estrategia previa del jugador al que se enfrenta, etcétera. Estos resultados, que surgen a partir de los juegos repetitivos, pueden entenderse por medio de la proposición formal:

Proposición 1: Si el parámetro de actualización,  $W$ , es lo suficientemente grande, no existe una estrategia óptima independiente de la estrategia utilizada por el otro jugador. (Axelrod, 1986, p. 26)

Como ejemplo de esto se puede mencionar que, durante los dos torneos que se llevaron a cabo, la estrategia que más resalto, puesto que ganó ambos, fue la estrategia denominada Tit-for-Tat. Consiste en que su primer movimiento será cooperar y, a partir de ahí, usará la estrategia elegida por el otro competidor durante la ronda anterior. Esto quiere decir que, si se cooperaba con ella, ésta siempre iba a cooperar; pero, si en algún momento se decidía jugar una defección con ella, en el siguiente turno, T-f-T utilizaría la defección.<sup>13</sup>

---

<sup>13</sup> T-f-T ( $\tau$ ) es una estrategia de equilibrio de Nash, y se puede comprobar de la siguiente manera:

1: Existen sólo tres respuestas posibles para alguien que juega ( $\tau$ ).

(a) Cooperar en todas las jugadas futuras.



A partir de los juegos evolutivos, se puede dar otra respuesta posible al dilema del prisionero. Según éstas, las relaciones de parentesco pueden ocasionar un altruismo real. Este altruismo modifica las ganancias totales (costos y beneficios) del causante, es decir, del emparentado. (Hamilton, 1963). Cooperar en el dilema es, por lo tanto, una especie de altruismo (Hamilton, 1971). El parentesco, en este sentido, puede tomarse como la idea de una interacción continua; tal y como se ha visto, en este tipo de interacciones parece tener mayores beneficios la estrategia de cooperar. Entre más interacción exista, la cooperación se podrá sostener mejor.

Un caso curioso comentado en los juegos evolutivos es que las interacciones entre dos individuos son más frecuentes, en tanto que tengas apartado a los demás. En el siguiente capítulo se comentará una cuestión similar a través de las emociones, puesto que con ellas se genera un margen de confianza entre las personas que pertenecen a un mismo grupo (cooperación), y se excluye y rechaza a aquéllos que no forman parte (defección). Cabe añadir que otro de los planteamientos de los juegos evaluativos para los juegos cooperativos es que en ellos se debe incentivar que la cooperación, a largo plazo, resultará con mayores beneficios, que las ganancias inmediatas por una defección en un corto plazo (Axelrod, 1986).

- 
- (b) Alternar cooperación con defección.
  - (c) Llevar a cabo una defección en todas las jugadas.

2: Probar que  $(\tau)$  puede ser la mejor respuesta para  $(\tau)$

Si ambos jugadores emplean la estrategia  $(\tau)$ , entonces cada uno espera la ganancia:

$$E(\tau, \tau) = 3(1 + \rho + \rho^2 + \rho^3 + \dots) = 3/(1 - \rho)$$

Donde  $\rho^\kappa$  es la probabilidad de que el juego se repita  $\kappa$  rondas.

Lo anterior es posible compararlo con la ganancia esperada de la estrategia alterna  $(\lambda)$ :

$$E(\lambda, \tau) = 4 + 1\rho + 4\rho^2 + 1\rho^3 + \dots = (4 - \rho)/(1 - \rho^2)$$

Existe una tercera posible mejor respuesta a la defección absoluta  $(\delta)$ :

$$E(\delta, \tau) = 4 + 2\rho + 2\rho^2 + \dots = 4 + 2\rho/(1 - \rho)$$

A través de esto, se puede afirmar que la estrategia  $(\lambda)$  es siempre inferior a una respuesta de la estrategia  $(\tau)$  que de la estrategia  $(\delta)$ ; y  $(\tau)$  es siempre una mejor respuesta que  $\rho > \frac{1}{2}$ . De acuerdo con esto, se puede concluir que  $(\tau, \tau)$  es un equilibrio de Nash cuando  $\rho > \frac{1}{2}$ . (Hargreaves y Varoufakis, 1995, pp. 173-4)

### 2.3.1.1. *¿La defección es racional?*

En este apartado se analizará brevemente, a partir de filosofías como la de Kant, Gauthier y Wittgenstein, si la defección es racional.<sup>14</sup>

Para Kant, por un lado, la moralidad es deontológica, y esta se sostiene como racional en tanto que el agente guía sus acciones mediante el imperativo categórico; éste afirma que toda acción que el individuo llevará a cabo debe ser tomada como ley universal, es decir, debe ser pensada de manera tal que el individuo pueda plantearse qué pasaría y cómo le afectaría – aunque no sólo a él, sino a todas las personas– si otra persona realizara esa misma acción. Esta moralidad resulta controversial en tanto que no se preocupa de las consecuencias inmediatas; no obstante, en la moralidad kantiana pueden resaltarse las motivaciones morales como herramientas importantes de la razón. En el caso del dilema del prisionero, por ejemplo, puede presentarse una idea altruista de que es importante el bienestar de los demás como una razón suficiente para llevar a cabo una cooperación y no una defección. Las motivaciones morales, por lo tanto, pueden ser una solución racional al dilema, rechazando que la defección sea lo racional, tal y como se presenta desde una postura instrumental de la razón (Hargreaves y Varoufakis, 1995, pp. 155-57).

Para Gauthier se puede plantear la defección, incluso desde lo instrumentalmente racional; para éste, lo racional es la cooperación y no la defección. Afirma que hay dos tipos de maximizadores, un maximizador directo (SM) y un maximizador restringido (CM). El (SM) sigue una lógica de defección, mientras que el (CM) puede utilizarse tanto como una estrategia condicional de la cooperación, en tanto se relaciona con otros (CM), como una estrategia de defección, en tanto se relaciona con (SM). El cálculo se lleva a cabo de la siguiente manera:

	C	D
C	3	1
D	4	2

Fig. 2.5

---

<sup>14</sup> Este apartado se basa en las ideas principales que aparecen en Hargreaves y Varoufakis (1995, pp. 155-164).

Se asume que la probabilidad de encontrar un (CM) es  $\rho$ , y su representación es la siguiente:

$$E(\text{resultado de CM}) = \rho \cdot 3 + (1 - \rho) \cdot 2$$

$$E(\text{resultado de SM}) = 2$$

Para cualquier  $\rho > 0$ , (CM) produce mejores resultados. Lógicamente parece simple; no obstante, se le ha criticado su postura, planteando la pregunta de qué es lo que motiva a una (CM) cooperar con otra (CM), dado que el asunto puede volver a caer en el dilema sobre por qué no traicionar o mentir para sacar una mayor ventaja (Hargreaves y Varoufakis, 1995, pp. 162-64).

En Wittgenstein, por otro lado, se puede concebir una analogía entre las consideraciones que hace del lenguaje con la discusión sobre la doctrina Harsanyi en la teoría de juegos. Se ha de recordar que, según la doctrina Harsanyi, todas las personas son agentes racionales y, en tanto que tal, deben llegar a las mismas conclusiones si tienen la misma información disponible. Wittgenstein, por su parte, problematiza que para que a algo se le atribuya un significado el lenguaje debe tener reglas, y eso parece complicado en caso de que el lenguaje sea un lenguaje privado. En la doctrina se discute si realmente se puede llevar a cabo el mismo procesamiento de información para llegar a un mismo resultado; en Wittgenstein se discute la posibilidad de una base de interpretación para que la comunicación sea posible. Wittgenstein, en este sentido, no discute acerca del lenguaje privado, sino de una cuestión que incluye a este por tratarse de algo más general, es decir, el cómo es posible la comunicación, y habla sobre bases de interpretación, o bien, reglas. En cuanto al dilema del prisionero, esto no se podría presentar directamente en la cuestión de cooperar o realizar una defección, pues a través de la racionalidad instrumental inmediatamente se puede afirmar que lo mejor es hacer una defección; la analogía con Wittgenstein, en este caso debe hacerse desde un punto de vista más amplio, es decir, en las normas que establecen cómo se elabora una matriz y en la manera en la que se efectúan los pagos asignados según la estrategia elegida, pues es a través de estas últimas que surge la problemática acerca de si la defección es o no racional (Hargreaves y Varoufakis, 1995, pp. 157-62).

## *2.4. Las reglas del juego*

En el juego del ajedrez, tal y como mencionan Amster y Pinasco (2014, p. 13), existe un teorema que puede parecer trivial. El teorema afirma que en juego se pueden dar tres situaciones:

Las blancas pueden generar un jaque mate.

Las negras pueden generar un jaque mate.

Cada jugador puede impedir que el otro le genere un jaque mate.

Afirmar alguna de las tres cosas resulta tautológico; no obstante, el teorema trata sobre que esto ya se encuentra establecido previamente en el juego. Las reglas que se establecen en el juego posibilitan que se cumplan los objetivos planteados, y no ocurran acontecimientos ajenos ni se haga un movimiento ilegal que modifique lo que ocurre en él; en el caso del ajedrez es claro, pues cada pieza tiene un movimiento específico.

La analogía con las relaciones sociales parece clara. Las decisiones individuales se llevan a cabo en un contexto donde existen normas establecidas. Dichas normas tienen la capacidad de modificar las consideraciones que se tienen al momento de plantearse la idea de realizar algo. Dependiendo de la sociedad, hay cuestiones que pueden ser incorrectas y ser reprendidas pues éstas se salen de “las reglas del juego”.

No obstante, parece que las normas se plantean a diferentes niveles. Más arriba se ha mencionado que la teoría de la cooperación pretende explicar cómo se da ésta sin necesidad de que entren otros mecanismos que la obliguen a efectuarse; uno de estos mecanismos podría ser el normativo. En el capítulo tres se regresará a este asunto, puesto que se pretende demostrar cómo actúan las emociones en la toma de decisiones, independientemente de las normas sociales.

## *2.5. Recapitulación*

Tal y como se ha dicho al inicio del capítulo, la teoría de juegos se consideró tanto una ampliación de la TER como una herramienta que podía llevar a considerar otro tipo de estrategias donde no se utilizara la racionalidad para ofrecer una buena respuesta. El juego,

en este sentido, se entiende como una interacción cualquiera entre agentes, regida por un conjunto de reglas, estrategias y resultados.

Las estrategias pueden llegar a ser deficientes u óptimas, y ambas pueden llevar a un resultado válido. Dichas estrategias tienen un aspecto formal de la teoría en la que se contienen elementos como el *common knowledge of rationality*, la *consistent alignment of beliefs*, la doctrina Harsanyi, entre otros. Del mismo modo, estas estrategias se presentan a través de un modelo llamado matriz.

En el último capítulo se pretende explicar cómo las emociones, junto con la racionalidad pueden ser un elemento que influya de manera crítica en la toma de decisiones. Se explicará a través del dilema del prisionero. Es por ello que de entre diferentes tipos de juego, el análisis se centró en aquellos que aportaban información relevante con respecto al dilema: juegos repetitivos finitos e infinitos, de suma cero y suma no-nula, cooperativos y no-cooperativos.

Se analizaron las respuestas posibles, bajo las condiciones originales, del dilema del prisionero, centrandó la conversación especialmente en la afirmación de si las estrategias tomadas para resolverlo eran racionales o irracionales. Esto con la intención de conectar los tres capítulos de esta tesis.

Por último, se desarrolló otro de los elementos de la teoría de juegos que entrará en el análisis de las emociones en el siguiente capítulo, a saber, las reglas del juego, o las normas sociales. En el siguiente capítulo se analizará brevemente qué se entiende en esta tesis como emoción. Se hará una comparación acerca de cómo pueden influir las emociones en una toma de decisiones, a diferencia de las normas sociales. Se explicará la relación e implicación de las emociones con la racionalidad y, finalmente, se expondrá, a través de un caso, la posibilidad de hablar sobre *emociones racionales*, desarrollando qué implicaciones, dificultades y virtudes tiene una tesis como esta.

## Capítulo 3

### Las emociones

Si vas hacia el pensar, lleva tu corazón contigo. Si vas hacia el amor, lleva tu cabeza contigo. Vacío es el amor sin pensar, hueco el pensar sin amor.

C.G. Jung. *El libro rojo*.

A través de los capítulos anteriores, se han visto dos posiciones con respecto a la racionalidad; recapitulando, en el primer capítulo se habló de la racionalidad paramétrica, en la que el sujeto que lleva a cabo una toma de decisión es lo único que varía, mientras su contexto y las demás personas —así como sus decisiones— se mantienen constantes y no se toman en cuenta al momento de evaluar el hecho en cuestión; en el capítulo dos se llevó a cabo el análisis de la racionalidad estratégica, por medio de la teoría de juegos, en la que todo el contexto, la sociedad en la que se desenvuelve, las personas y sus acciones con las que interactúa, etc., tienen un peso en la evaluación.

Es importante destacar que la racionalidad paramétrica, a través de la teoría de la elección racional en su forma más básica, llegaba a la conclusión de tomar a la racionalidad como algo perfecto, en lo que no cabía el error. Al final del primer capítulo se plantearon algunas objeciones, de entre ellas, que eso no coincidía con una cuestión práctica, es decir, fácilmente una persona podría percibir que al momento de tomar decisiones ha habido, o toman en cuenta la probabilidad de que exista, un fallo.

Por el contrario, algunos teóricos de los juegos, por medio de la racionalidad estratégica, llegaron a la conclusión de que el individuo no tenía por qué ser siquiera racional al momento de tomar una decisión; esta conclusión se daba debido a que se toma en cuenta todo el contexto en el que el sujeto se desenvuelve y que, debido a las normas sociales previamente establecidas, así como la interacción con los otros, se podría determinar cuál decisión sería la correcta. No obstante, como se discute en el capítulo dos, la irracionalidad tampoco es una respuesta viable al momento de tomar una decisión.

Parafraseando a Pascal (1981), se pueden tomar como dos excesos tanto el excluir a la razón como el sólo admitirla. Por tanto, uno de los argumentos principales de esta tesis

consiste en que la racionalidad, si bien no es suficiente para llevar a cabo una buena evaluación de una decisión, tampoco hay por qué abandonarla. La racionalidad, entonces, ha de tener un elemento que le apoye en la cuestión; el elemento propuesto aquí es el de las emociones.

Si bien es cierto que Pascal escribió sus *Pensamientos* en 1670, la mayor parte de filósofos y pensadores ha tratado a las emociones como un obstáculo para llevar a cabo una buena toma de decisiones. Esto se puede ver claramente, por ejemplo, a través de la tradición kantiana y la utilitarista con respecto a la moralidad; según éstas, las respuestas emocionales representan, principalmente, el “lado animal” de la naturaleza humana como una cuestión evolutiva y, por lo tanto, no son útiles al momento de decidir qué es lo que se tiene que hacer (Goldie, 2010). Incluso en la actualidad aún puede verse esta idea del entorpecimiento de la racionalidad debido a las emociones; pareciera que la idea está muy arraigada en la sociedad, ya que incluso se pueden encontrar imágenes de una conversación entre un cerebro, como análogo a la racionalidad, y un corazón, representando a las emociones, discutiendo sin ponerse de acuerdo, entre otras cosas.

No obstante, en décadas recientes, se ha tratado de modificar esa visión de las emociones como un componente irracional de las personas. El epígrafe citado al inicio de este capítulo, por ejemplo, nos muestra la idea de encaminar las acciones sin separar “la cabeza del corazón”; en otras palabras, el autor considera que el llevar a cabo una acción determinada sin tomar en cuenta esos dos elementos, el pensar y el sentir, conlleva a un análisis incompleto y, por tanto, a una consecuencia no deseada.

Sin embargo, esta postura sobre tener en cuenta a las emociones al momento de tomar una decisión, o bien de relacionarlas con la racionalidad, ha abierto una discusión muy grande al respecto de qué papel juegan exactamente las emociones. Algunos como Damasio (1996) creen que son precondiciones de la racionalidad, es decir, mecanismos que actúan cuando el tiempo es un hecho determinante al momento de decidir qué acción llevar a cabo; esta postura resalta un aspecto más fisiológico, es decir, programas neuronales que dan respuestas corporales inmediatas al tener un mayor vínculo con la percepción. Otros, como Martha Nussbaum (2008), por ejemplo, afirman que las emociones pueden incluso traducirse en contenidos proposicionales, o bien, que pueden ser expresadas; esta postura, por su parte, se

inclina más en relacionar las emociones con un aspecto cognitivo. Esta tesis, no obstante, no se enfocará en la discusión entre las posturas fisiológicas y las cognitivas. En la primera parte de este capítulo se tratará de explicar brevemente qué se está entendiendo aquí por *emoción*, así como la postura que se tomará al respecto.

A partir de esto, en el siguiente apartado, se explicará la relación que tienen las emociones con la racionalidad. En este mismo apartado, y recuperando una sección del capítulo dos, se analizará la relación de las emociones con las normas sociales y los valores, ya que éstas también pueden afectar dicha relación.

Más adelante se argumentará acerca de los beneficios de introducir las emociones en la toma de decisiones, así como algunas posturas que, si bien apoyan dicho planteamiento, insisten que las emociones no pueden ser por sí mismas racionales. Dado este rechazo a la posibilidad de que existan emociones racionales, se hablará, por ende, de emociones adecuadas, con las que se podrá sostener esta tesis de que las emociones, junto con la racionalidad, pueden evaluar críticamente una toma de decisiones, e incluso permitir que el individuo lleve a cabo la mejor decisión.

Por último, se retomará el caso del Dilema del prisionero con sus respectivas modificaciones, introduciendo las emociones adecuadas, para analizarlo a través de éstas. Lo que conllevará a las conclusiones de este trabajo.

### *3.1. ¿Qué son las emociones?*

Debido a la enorme discusión que se encuentra en torno a la cuestión de qué son las emociones, se darán primero unas pequeñas aclaraciones. Por un lado, el término emoción tendrá correlación con el término pasión; esto se debe a que el primer término refiere normalmente a una perspectiva contemporánea, mientras que el segundo fue un término utilizado durante siglos para denominar —con algunas diferencias— lo mismo. Si bien no se hará una reseña histórica acerca de qué es una emoción, es importante aclarar que en ocasiones podría usarse el segundo como sinónimo del primero. Por otro lado, también se diferenciará entre emoción y sentimiento; esta aclaración es importante debido a que algunos autores suelen utilizarlos como sinónimos. Los sentimientos, siguiendo a Solomon (2007), se relacionarán directamente con aspectos fisiológicos más que cognitivos.



Seguendo el análisis de Elster (2001, pp. 34 y ss.), las emociones pueden entenderse, ya sea como un acontecimiento o como disposiciones y, la mayoría de ellas, presentan las siguientes características:

1) Sensación cualitativa:

Las emociones pueden presentarse como acontecimientos únicos para las personas, es decir, se experimentan y se sienten diferentes unas de otras. La mayoría de las veces se puede discernir que cierta emoción es una y no otra, por ejemplo, entre la alegría y la tristeza.

2) Aparición súbita:

Las emociones se desatan ya sea por estímulos cognitivos o perceptuales de manera instantánea. Por ejemplo, bajo condiciones normales, si algo le provoca al sujeto cierta sensación de peligro, éste puede reaccionar inmediatamente ya sea con ira o con miedo. No obstante, esta característica no es universal; existen emociones, como el amor, que pueden ir creciendo de forma gradual, e incluso puede darse el caso de que esto ocurra sin que el sujeto se dé cuenta.

3) Imprevisibilidad:

Aunque tampoco es una característica universal, las emociones suelen considerarse más como eventos, es decir, hechos que le ocurren al sujeto de manera involuntaria, y no como acciones; esto se debe a que las emociones suelen desencadenarse por estímulos externos y no provocadas por el sujeto.

4) Corta duración:

Al desencadenarse una emoción, ésta se desarrolla y después desaparece, haciendo que el sujeto regrese a un estado normal. Las emociones no suelen tener una larga duración; volviendo al ejemplo dado en (2), la ira o el miedo pueden presentarse ante una sensación de peligro, y después desaparecer junto con esa sensación. Sin embargo, esta característica tampoco es definitoria para las emociones, puesto que existen algunas que pueden durar varios días y algunas, incluso, toda la vida.

#### 5) Desencadenamiento cognitivo:

Las emociones pueden darse por diversas creencias que sostenga el sujeto. Las creencias pueden estar relacionadas con otras personas, con el sujeto mismo, por hechos temporales, contextuales o culturales, y pueden tener ciertos grados de valoración, desde un hecho tomado como cierto, hasta lo meramente probable. Si bien éste ha sido el análisis clásico, actualmente se considera que las emociones pueden originar la creencia que ha de justificarla. Un ejemplo clásico al respecto es que el miedo puede originarse debido a un mero hecho perceptual o sensorial, y no cognitivo; y a través de ella, generarse la creencia.

#### 6) Dirigidas hacia un objeto intencional:

Las emociones son sobre algo, es decir, se relacionan directamente con una cosa, una persona o un estado de hechos. Un ejemplo básico es que uno puede enfadarse *con* alguien por *x*. Budd (1995) defiende que existen contraejemplos a este punto, y utiliza la expresión musical de las emociones como uno; explica que la emoción en este caso es abstracta, pues no refiere a nada en concreto, por un lado, y tampoco refiere a un sujeto definido. Esta discusión es larga y ajena a los propósitos de este trabajo, por lo que no se desarrollará más, sin embargo, sí ha de añadirse que no me parece un contraejemplo claro a este punto, debido a que a través de éste se pueden tomar en cuenta las intenciones del autor al momento de realizar la composición, entre otras cosas.

#### 7) Expresión fisiológica y fisonómica:

Las emociones vienen acompañadas de un aspecto fisiológico. Pueden percibirse como posturas corporales que adopta el sujeto, como cambios en el tono de la voz, en el acto de ruborizarse, sonreír, llorar, entre otros. No todas las emociones tienen una expresión característica, por lo que pueden llegar a confundirse si sólo se pretende definir una emoción a partir de lo observable.

#### 8) Inducción de cambios fisiológicos:

Dentro del aspecto fisiológico también se puede destacar que la emoción se puede presentar en diversas intensidades, según el nivel de excitación que se haya generado en el sujeto; asimismo, estos niveles se presentan de manera autónoma, por lo que el sujeto no tiene control sobre éstas. Pueden distinguirse, por ejemplo, ciertas variaciones en el ritmo cardíaco,

náuseas, sentirse como un “nudo en la garganta”, entre otras. No obstante, las emociones podrían presentarse sin ningún signo de excitación.

#### 9) Inducción de tendencias a realizar determinadas acciones:

Las emociones se relacionan con tendencias a realizar ciertas acciones; esto no quiere decir que las emociones lleven a un sujeto a actuar, sino que se limita a la disposición que se tiene de hacerlo. Una emoción como la ira, por ejemplo, puede generar la tendencia a lastimar o hacer daño a lo que nos la ha causado. Dado que se habla de una tendencia, esto puede reducirse al sólo querer hacerlo e incluso a suprimirlo y, por lo tanto, a no llevarlo a cabo.

#### 10) Valencia

Las emociones pueden experimentarse tanto de manera positiva como negativa, es decir como placenteras o como dolorosas, o como deseables o indeseables. Por ejemplo, se puede clasificar que la alegría se presenta como algo positivo, mientras que la tristeza como algo negativo o indeseable. Existen emociones que pueden experimentarse con una valencia neutral, por ejemplo, la que puede llegar a experimentarse al conocer que un buen amigo tiene mucho éxito, ya que puede experimentarse cierta envidia al mismo tiempo que alegría.

Estas diez características no tienen por qué encontrarse en todas las emociones, es decir, puede variar su presencia entre una y otra. No las define, aunque sí que pretende señalar aproximadamente qué se puede comprender como una emoción. Algo importante a destacar de los diez puntos anteriores es que las emociones se pueden relacionar directamente con dos aspectos: el aspecto fisiológico y el aspecto cognitivo. El primero se refiere, como se ha dicho ya, a los aspectos visibles de una emoción; dentro de éste se puede encontrar una manera de analizar las emociones a partir de ciencias como la neurobiología o la psicología, que han desarrollado teorías acerca de dónde se aloja cada emoción en el cerebro, así como la idea de que las emociones tienen mayor relación con la percepción y, por ende, éstas pueden entenderse como programas de redes neuronales, entre otras cosas (Ekman, Levenson y Friesen, 1983; Damasio, 1996) . El segundo aspecto se refiere a la relación que tienen las emociones con las creencias y los deseos del sujeto que las experimenta; esta vía ha sido la manera de explicar las emociones de muchos filósofos, y se ha argumentado a través de teorías como la de la valoración, en donde las emociones adquieren un significado

dependiendo de cómo se percibe cierta información recibida, como benéfica o perjudicial al sujeto (Arnold, 1960; Frijda, 1995); algunos otros han expresado que las emociones son cognitivas porque se pueden traducir en contenidos proposicionales tales como “X experimenta la emoción *p*”, o bien, “Baruch está enojado porque...” (Nussbaum, 2008). Si bien definiendo la relación que existe entre ambos aspectos, es decir, que no existe un análisis completo de las emociones sin tomar en cuenta a los dos, este trabajo aborda la relación que existe entre la racionalidad y las emociones, por lo que el aspecto fisiológico se dejará de lado para poder llevar a cabo un análisis más exhaustivo de lo cognitivo, ya que este último tiene mayor peso con respecto a la racionalidad.

### *3.2. Racionalidad y emociones*

Tal y como se dijo al principio del presente capítulo, la idea de que las emociones pueden llevar a tomar mejores decisiones al relacionarlas con la racionalidad es muy reciente. Si uno echa una pequeña ojeada a la historia de la filosofía se puede dar cuenta que las emociones estuvieron conectadas con el actuar irracional de las personas por mucho tiempo. La filosofía estoica, por poner un ejemplo, pretendía llevar a cabo una evaluación de las emociones con el fin de suprimirlas ya que éstas podían nublar la razón. Si bien ocurría esto, es cierto que algunos autores, en mayor o menor medida, rescataban algunos aspectos de las emociones. Spinoza, por ejemplo, expresaba que tenían que evitarse las pasiones negativas y buscar las positivas, como la alegría, ya que éstas podían ayudar a la persona a mejorar, mientras no hubiese una alegría excesiva. Estas dos ideas, la estoica y la de Spinoza, comparten la idea de una ética, de un mejoramiento del individuo, lo cual es una de las motivaciones de este trabajo. Siguiendo a Victoria Camps (2011), “resaltar el papel de las emociones en la ética es un modo, quizá el único, de abordar el poco tratado problema de la motivación moral”. Para ello, debe hablarse primero de la relación de la racionalidad con las emociones.

#### *3.2.1. El estado cognitivo de las emociones*

En el primer capítulo se afirmó que el sujeto podía ser descrito a partir del conjunto de creencias, así como por el conjunto de deseos o preferencias que éste sostuviera, entre otras cosas. Al explicar el modelo estándar de la teoría de la elección racional se presentó un pequeño diagrama en el que se podía comprender visualmente la relación que tenían tales

elementos con el llevar a cabo una acción. En “Economics”, Elster (1997; p. 129) presenta un diagrama alternativo introduciendo a las emociones en el análisis:

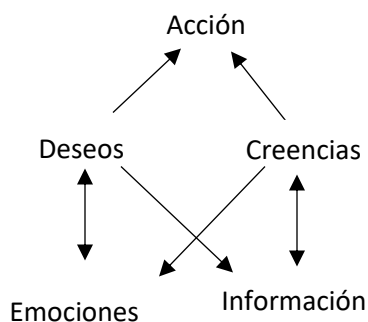


Figura 3.1

En el punto cinco del apartado 3.1, desencadenamiento cognitivo, se afirmó que las emociones tenían a las creencias como antecedente del sujeto, es decir que las primeras dependen de las segundas en tanto que para su reacción se debe tener cierto conocimiento sobre lo que está aconteciendo; esto se puede ver claramente en la figura anterior. No obstante, como se explicó también, es posible que ocurra un movimiento contrario, es decir, la emoción y después la creencia. Este movimiento no se ve en el diagrama ya que éste se centra en la versión clásica de explicar las emociones; sin embargo, se rescata aquí para ver cómo pueden las emociones introducirse dentro del paradigma de la elección racional.

Esta relación entre las emociones y las creencias es ampliada en *Alquimias de la mente* (Elster, 2002), puesto que existe otra manera en la que se vinculan. Es posible tener conocimiento acerca de una emoción y, por lo tanto, se puede generar una creencia acerca de ella. Esto se puede desglosar de la siguiente manera (p. 309):

1. Creencias acerca de las emociones de la propia persona.

Este punto puede entenderse mejor a través de una vía negativa; el que haya creencias acerca de las emociones propias implica también que es posible que existan emociones de las que no se tenga conocimiento y, por ende, no se puedan establecer algunas creencias sobre ellas; Elster (2001, p. 104) denomina a éstas como *protoemociones*. Existen tres tipos de protoemociones: fuertes, semifuertes y débiles. Las primeras son protoemociones que el sujeto puede darse cuenta de que tiene, pero que no ha terminado de conceptualizar o hacer

consciente; las segundas son aquellas de las que sí existe un concepto dentro del entorno sociocultural del sujeto, y que las demás personas pueden notar que tiene, pero de las que no es consciente que tiene; las últimas se definen como aquellas que el sujeto no puede reconocer debido a que el concepto no existe en su entorno sociocultural.

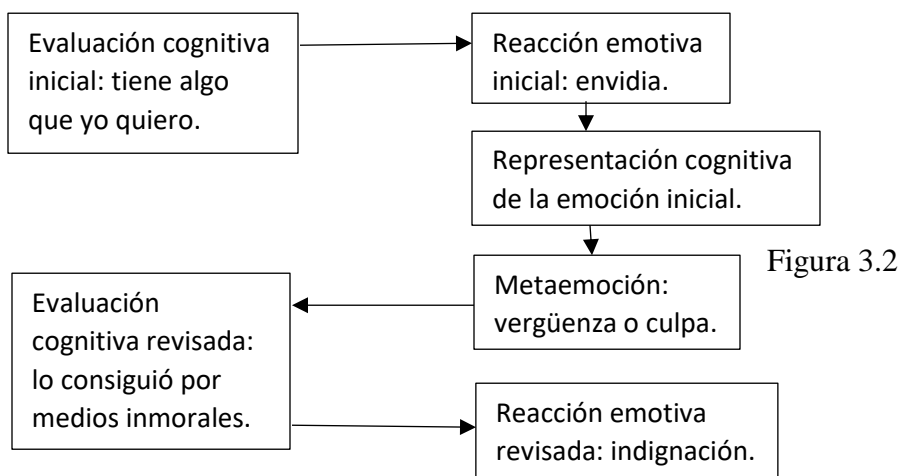
Ahora bien, al analizarse una emoción de la que sí se puede tener alguna creencia puede generarse una nueva emoción, que servirá para modificar a la primera o bien para modificar la conducta de la persona. A esta emoción de segundo orden se le denomina *metaemoción*. Ésta, en otras palabras, conlleva un proceso de autorreflexión por parte de la persona.

2. Creencias acerca de las emociones de las demás personas.

El reconocimiento de las emociones en otras personas ya sea que éste logré llevarse a cabo satisfactoriamente o no, es decir, que las creencias que tiene la persona con respecto a las creencias del otro sean o no correctas, puede modificar sus actitudes, sus emociones e incluso, como se verá al final de este capítulo, sus decisiones.

3. Creencias acerca de las motivaciones de las demás personas.

Esto podría comprenderse como un tercer nivel en las creencias: Uno puede tener una emoción inmediata sobre la acción de una persona; posteriormente darse el caso explicado en el punto dos, es decir, tener una creencia acerca de la emoción que tiene el otro, y después modificar esa creencia al evaluar cuáles fueron las razones, o motivos, que llevaron a la otra persona a actuar de esa manera, generando una reacción, o emoción diferente a la anterior. El siguiente diagrama muestra claramente cómo se da este punto (Elster, 2001; p. 108):



#### 4. Creencias sobre probabilidades.

Estas creencias son aquéllas que evalúan un estado de cosas determinado y lo relaciona con aquello que podría ocurrir, es decir, con algún hecho futuro. Este tipo de creencias se relacionan directamente con las emociones del miedo y la esperanza, en las que se piensa en algo que aún no ha ocurrido, pero se cree posible, en mayor o menor medida.

#### 5. Creencias contrafácticas.

Este tipo de creencias se relacionan con las emociones que surgen a partir del pensar en cómo le afectaría a la persona una situación en caso de que le hubiese ocurrido a ella. Se puede comprender fácilmente a través de la oración “Podría haber sido yo”. Se relacionan con emociones como la empatía, que surge a través de ver lo que le ocurre a otra persona.

#### 6. Creencias “como si”.

Éstas son muy similares a las creencias contrafactualas; la diferencia es que en este tipo de creencias se pueden hacer relaciones directas incluso con hechos no reales o ficticios. Del mismo modo que las anteriores, se puede plantear la idea de que algún hecho puede verse “como si” le hubiera pasado a uno y, a través de eso tener una emoción y una reacción específica.

Existen otras posturas con respecto a los antecedentes cognitivos de las emociones que sería interesante destacar, aun pese a que no se profundice en este trabajo con respecto a ellas. Por ejemplo, en “Y el cerebro creó al hombre” (Damasio, 2010), el autor trata a las emociones como programas complejos localizados en alguna parte del cerebro que se relacionan directamente con acciones efectuadas en el cuerpo, como las expresiones faciales. En “El error de Descartes”, Damasio (1996) afirma que, sin las emociones, la toma de decisiones se llevaría a cabo a través de una racionalidad incompleta, debido a que éstas ayudan a hacer una evaluación de prioridades. Otros a través de las *teorías de la valoración* afirman que la interpretación de los acontecimientos como eventos de significación personal tienen como resultado a las emociones. Frijda (1986), por ejemplo, fundamenta a las emociones en el placer y el dolor. Modzelewski (2017) resume muy bien la postura de Frijda al afirmar que éste “sostiene [...] que las emociones son respuestas a sucesos importantes para el individuo, experiencias subjetivas cuyo núcleo es el placer o el dolor y que incluyen una evaluación

sobre el significado situacional. Agrega que siempre que se produce una emoción hay alguna forma de cambio en la tendencia del sujeto a la acción” (p. 101).

En el diagrama 3.1 se puede apreciar que las emociones no sólo se relacionan directamente con las creencias, sino también con los deseos. Las emociones pueden depender de los deseos y, a la inversa, las emociones pueden provocar algunos cambios en los deseos. En el primer caso, las emociones surgen a partir de que se lleve a cabo la satisfacción o la frustración de un deseo y, en el segundo caso, la emoción generada por dicha satisfacción o frustración de un deseo puede modificar dicho deseo. Tal y como se explicó en el capítulo uno, debe quedar claro que los deseos son el cálculo sobre qué tan importante es la decisión que se tomará.

### *3.2.2. Patrones de conducta: normas sociales y valores*

Las emociones, así como las creencias, se ven influidas por dos factores: las normas sociales y los valores. Desde un acercamiento general, esto es debido a que, de una cultura<sup>15</sup> a otra, éstas se evalúan de distinta manera, es decir, que mientras para alguna cultura la cuestión del honor, por ejemplo, es un punto clave para evaluar si alguien está actuando moralmente o no, para otra cultura esto podría carecer de importancia. Asimismo, debido a los diferentes grados de valoración que existen sobre las emociones, éstas pueden desencadenarse, a su vez, a diferentes niveles; estas diferencias también son un factor para las protoemociones explicadas más arriba, ya que en una cultura podría no existir una emoción específica que en otra sí, o que no se haya conceptualizado. Debido a ello, la conducta presentada frente a las emociones también puede variar.

Las normas se encargan de regular las conductas de las personas. Las normas sociales, por lo tanto, son regulaciones de la conducta que se comparten con las demás personas, es decir, *socialmente*; además, éstas deben estar aprobadas por la comunidad de personas en la que actúan. Para que las normas sociales se desarrollen por completo, las emociones que se generan al infringirlas deben compartirse; es decir que, si uno comete un acto en el que se

---

<sup>15</sup> Cultura la utilizaré aquí en un sentido muy general, en el que se incluye la idea de que existen comunidades o grupos sociales con diferentes sistemas de creencias, valores, y prácticas que pueden diferir entre una y otra.



infringen las normas y esto provoca el sentir vergüenza o culpa, esto se debería de sentir en cada caso similar.

Las normas sociales pueden integrarse en tres formas diferentes de ser emitidas (Elster, 1997; p. 114). La más sencilla se puede enunciar como “Haz *x*” o su contraria, “No hagas *x*”. La segunda forma se expresa con relación a una conducta ya realizada, ya sea por parte de la persona misma o de otras, es decir, “Si haz hecho *y*, entonces haz *x*”, o bien “Si otros han hecho *x*, entonces haz *x*”. La tercera forma podría considerarse como una versión kantiana de las normas: “Haz *x* si fuera algo conveniente que todos llevaran a cabo *x*”.

La relación que tienen las emociones con las normas sociales se da de dos maneras. Por un lado, las emociones pueden funcionar como factores de imposición, es decir que a través de ellas se puede asegurar que se respeten y se cumplan las normas sociales; por ejemplo, que una persona sea juzgada por lo que pretende llevar a cabo y que no coincide con las normas puede provocarle culpa o vergüenza, que desencadena a su vez el hecho de que al final decida no realizar tal acción y, por tanto, respetar la norma social; en otras palabras, Elster (2001, p. 91), afirma que “quienes se desvían de la norma se arriesgan a recibir un castigo o a verse reducidos al ostracismos por parte de otros miembros del grupo, [ya que] las normas se acatan para conseguir el beneficio directo de no verse penalizados”. Por otro lado, las normas sociales pueden tanto estipular qué emociones deben expresarse en circunstancias específicas como regular las emociones por sí mismas; en este punto se suelen introducir cuestiones más complejas como la educación o las costumbres que suelen existir en una sociedad, ya que a partir de lo que se aprende o se observa se limita quién puede expresar cierta emoción, hacia quién se dirige, así como en qué momento ocurre. Tomando un ejemplo de Ekman (1980, p. 87), en un funeral hay una especie de derecho a luto, con el que se expresa que la esposa de un hombre fallecido tiene que verse mayormente afligida en comparación a la secretaria de éste.

El segundo factor que influye en las emociones son los valores; éstos se relacionan con las preferencias y los compromisos individuales que, a su vez, se comparten con otros individuos, aunque, a diferencia de las normas sociales, no dependen de la aprobación o no aprobación de los otros individuos (Elster, 1980; p. 91). La mayoría de los valores que sostiene un individuo son transmitidos generacionalmente, es decir que son aprendidos. Los

valores afectan directamente en las emociones debido a que el individuo se sentirá cómodo o no con el hecho en cuestión dependiendo de si este último encaja o no con esos valores.

Debido a la estrecha relación que tienen las normas sociales y los valores, su distinción resulta compleja de elaborar. Los valores pueden sostener, a través de los intereses de la persona, algunas convenciones establecidas socialmente; las normas sociales, por su parte, hacen que lo establecido socialmente se cumpla a través de la reprobación y la sanción. Aunque los aspectos morales, y junto con ello lo normativo, resulta de vital importancia para las conclusiones que se pueden inferir de este trabajo, profundizar más en ello llevaría a una posible desviación del tema central, por lo que no se ahondará más en ello.

Las emociones que se relacionan con los dos aspectos ya mencionados, las normas sociales y los valores, son aquéllas que se podrían definir como *emociones sociales*. Éstas se comprenden a través de una evaluación positiva o negativa que hacen algunas personas acerca de la conducta o el carácter de otra. Siguiendo el análisis de Elster (1980), hay principalmente ocho emociones sociales, a saber:

Emociones sociales *negativas*:

- *Vergüenza*: emoción negativa que surge por una creencia respecto al carácter de la persona que la siente.
- *Desprecio u odio*: emociones negativas dadas por creencias acerca del carácter de otras personas.
- *Culpa*: emoción negativa generada por una creencia de la acción de la propia persona.
- *Ira*: emoción negativa dada por creencias acerca de la acción de otra persona.

Emociones sociales *positivas*:

- *Dignidad*: emoción positiva que surge por una creencia respecto al carácter de la persona que la siente.
- *Simpatía*: emoción positiva dada por creencias acerca del carácter de otras personas.
- *Orgullo*: emoción positiva generada por una creencia de la acción de la propia persona.
- *Admiración*: emoción positiva dada por creencias acerca de la acción de otra persona.

### 3.2.3. Elección

El último elemento que se analizará aquí para hablar sobre la relación de la racionalidad con las emociones es el de la *elección*. La elección aquí se entenderá como la situación previa (el momento en el que un individuo, a través de la racionalidad y las emociones, toma una decisión) del llevar a cabo una acción. Elegir tiene como supuesto la idea de que la persona se guía por una estructura de recompensa que le dan las opciones entre las que se encuentra decidiendo, es decir, si tiene que escoger entre  $x$  y  $y$ , si escoge  $x$  es porque un cambio de la estructura en cualquiera de las opciones le hizo decidir, independientemente de si el resultado es el esperado o no.

La elección se explicará a través de tres niveles diferentes: la acción sin elección, la elección mínima y, la más importante para esta tesis, la elección racional.

La acción sin elección es algo muy discutido ya que existen cuestiones meramente fisiológicas en las que podría decirse que el sujeto ha realizado una acción voluntaria porque *no tenía elección* (Frijda, 1986). Un caso difícil de analizar sería, por ejemplo, la micción; supóngase que una persona tiene que trabajar en algún lugar que queda a seis horas de distancia en carretera de su casa y por la tercera hora comienza a tener ganas de hacer del baño. Éste puede elegir detenerse en alguna gasolinería y llevar a cabo la acción o aguantarse hasta llegar a su destino. En primera instancia se podría considerar que *realmente* está tomando una decisión al respecto; no obstante, lo que se señala es que cualquiera de las dos elecciones puede no ser una acción, tal y como se dijo más arriba, por ser una urgencia corporal, un acontecimiento, y no algo voluntario. A través de este ejemplo se puede notar que a lo que se refiere el hablar de acción sin elección es a la idea de que se realice una acción de manera voluntaria, pero que el individuo sea insensible a la recompensa obtenida entre una cosa y otra. Sostengo aquí que aunque la micción es una reacción más fisiológica y que, por ende, no se puede impedir que se lleve a cabo su realización (es decir, que no se elige ni es una acción voluntaria) y, por tanto, se considere una acción sin elección, esta ansia, este deseo, delimita tanto las opciones como el horizonte de objetivos que pueden darse de manera cognitiva en el individuo al momento de ocurrir, haciendo, de este modo, que se tenga, en cierto sentido, sensibilidad hacia la recompensa y, por ende, se efectúe como una acción con elección mínima, aún pese a que la micción por sí misma no lo sea.

La elección mínima, por otro lado, ya comprende la idea de ser sensible a la recompensa, aunque ese es el único requisito. La elección mínima está enormemente mediada por el deseo del individuo, sin tomar en cuenta ninguna creencia acerca de si le conviene o no llevar a cabo esa elección. Por ejemplo, supóngase ahora que el individuo tiene la posibilidad de elegir entre dos opciones,  $x$  y  $y$ . Asimismo, supóngase que una de las variables es constante mientras que la otra puede variar, es decir,  $y_i$ ; se podría decir que  $x$  es ahorrar dinero y  $y_1$  comprar un helado,  $y_2$  comprar unas galletas,  $y_3$  comprar unas papas, etcétera. Límitese en este ejemplo a  $x$  en tanto mera acción de guardar dinero para después comprar algo que le guste para comer. El individuo podría no desear en ese momento consumir algo y elegir  $x$  o encontrar algún producto que desee consumir en ese momento y escogerlo,  $y_i$ . El individuo, sensible a la recompensa, por tanto, elige realizar una elección de acción debido a su deseo, sin importar cuál le beneficia más, o cuál le ofrece más o menos producto, por seguir con el ejemplo.

El caso de la elección racional se ha discutido ampliamente en los dos capítulos anteriores. Para que una elección pueda ser considerada como racional, los conjuntos de creencias, deseos e información recabada por la persona deben tener cierta coherencia entre sí. Además, como señala Davidson (1995), se debe asumir que el agente es racional.

Como ya se ha dicho anteriormente, que una acción sea racional significa que, dadas ciertas creencias –racionales– del individuo, es la mejor manera de satisfacer sus deseos. Para que una creencia sea racional debe basarse en una cantidad de información óptima (Di Castro, 2009; Elster, 2001, 2014; Winter, 1964).

Con respecto a la relación que tienen las emociones con la idea de elección, las ideas también suelen ser muy diversas. Solomon (1993), por poner un ejemplo, afirma que todas las emociones son una estrategia subjetiva que tiene como propósito maximizar la propia dignidad y, por lo tanto, éstas pueden elegirse racionalmente. Elster (1997, 2001, 2002), por el contrario, considera que las emociones se experimentan involuntariamente y, por lo tanto, no pueden elegirse. Desde esta perspectiva se pueden llevar a cabo ciertas acciones que permiten ver la relación aquí planteada. Primero, si bien no se puede elegir qué emoción tener según sea el caso, se puede elegir la situación en la que se produce; esto se lleva a cabo cuando, por ejemplo, se prefieren evitar ciertos hechos para que las emociones que pudieran

surgir al respecto no se produzcan. En segundo lugar, se pueden generar emociones a través de imaginarse algunas situaciones, tanto que han ocurrido como que podrían ocurrir (Hochschild, 1983). Tercero, es posible evitar ciertos aspectos de la emoción controlándolas; cuando ocurre la aparición de una emoción, es posible desviar la atención o simplemente aceptar que se está produciendo la emoción y dejar que ocurra suprimiendo las expresiones fisiológicas. Por último, se cree que es posible controlar la disposición que se tiene previamente antes de que se genere una emoción; este aspecto es el sugerido normalmente por diversas corrientes filosóficas, como el caso de los estoicos, por prácticas religiosas como el budismo e incluso por tratamientos como la psicoterapia.

Tal y como se ha mencionado anteriormente en este capítulo, por siglos se ha aceptado la idea de que las emociones dificultan el llevar a cabo una elección racional; en décadas recientes se ha pretendido defender la posición contraria, argumentando que las emociones permiten preparar al individuo a situaciones complejas en las que el análisis racional resulta insuficiente. Un ejemplo de esto puede ser el caso en el que se requiera necesariamente tomar una decisión; si ninguna de las opciones implica mayor optimización, las emociones que provoca una u otra podrían ser el factor que determine la elección. En este apartado se ha profundizado un poco más en el asunto de que se puede tener cierto dominio sobre las emociones, teniendo como resultado que las emociones podrían dejar intacta la capacidad de tomar una decisión racional, o sea, que las emociones no interfieran en la toma de decisiones. ¿Pero, las emociones pueden realmente ser un factor que, junto con la racionalidad, ayude a tomar la *mejor* decisión?

#### 3.2.4. Emociones racionales

Para responder la pregunta anterior, se analizará primero la idea de si las emociones racionales son plausibles. Di Castro (2009, p. 133; Elster, 1997, p. 130) resume claramente qué es lo que podría entenderse como una *emoción racional*:

- a) Las emociones racionales son un tipo de acción racional.
- b) Las emociones son racionales si se dan adecuadamente a la situación en la que se presentan.
- c) Las emociones son racionales si su intensidad corresponde a la situación en la que se presentan.

- d) Las emociones son racionales si se fundamentan en creencias racionales.
- e) Las emociones son racionales si se fundamentan en deseos racionales.
- f) Las emociones son racionales si hacen feliz al individuo a quien se le presentan.

a) Como se mencionó en el apartado anterior, Solomon (1980) sostiene que las emociones podrían comprenderse como cierto tipo de acciones, al considerarse racionales e intensionales. Recuérdese el diagrama 3.1.; se puede ver que la acción es racional en tanto se sostenga de la coherencia de los deseos y las creencias del individuo. Del mismo modo la información, al ser coherente con estas últimas, puede considerarse, por tanto, como racional y, a su vez, como una acción: la acción de recabar información. Siguiendo lo afirmado por Solomon podría inferirse una cuestión similar, dado que las emociones podrían comprenderse como intencionales, es decir que estarían tratando de alcanzar algún fin en específico, y por ende las emociones como acción podrían reflejarse a través de la conducta.

b) Tomando una cita de Elster, Broad (1971) afirma:

Algunos tipos de calidad emocional son *adecuados* y otros *no lo son* respecto de un objeto epistemológico dado. Es adecuado percibir con algún grado de *temor* un objeto que uno considera *amenazante*. Es inadecuado percibir con *satisfacción* o con *alegría* a un congénere *que sufre dolor o pena inmerecidos*. Luego, una emoción que se adecua *en cuanto al tipo* a su objeto epistemológico puede no adecuarse *en cuanto al grado*, o sea que puede ser excesiva.

Es decir, si la emoción que se presenta ante una situación es adecuada, se puede afirmar que la emoción es racional. Esta emoción, por lo tanto, no sólo debe ser adecuada de la manera ya mencionada, sino que también debe guardar cierta coherencia con los deseos y las creencias del individuo, puesto que si sólo lo hiciera con la situación cabría preguntarse si es adecuada con relación a si los demás individuos aceptan o rechazan dicha emoción; en caso de ser así se estaría hablando, ya no de racionalidad, sino de normas sociales. Asimismo, tampoco debe entenderse que exista una relación necesaria, es decir, que dada una situación *x* exista necesariamente una emoción que se dé y sea adecuada, porque de ser así volveríamos al problema planteado en el capítulo uno acerca de una racionalidad perfecta.

c) Este aspecto se centra en la idea de que el nivel de intensidad que presenta una emoción es lo que especifica si es racional o no; debe aclararse que no existe algo así como un nivel “exacto” de intensidad. Esto se puede entender claramente a través de un ejemplo clásico.

Supóngase que una persona está por cruzar la calle; puede sentir miedo por el riesgo de ser atropellada. Este miedo puede ser coherente en ese sentido. Si se analiza un poco más se puede hablar de niveles de intensidad. Si su miedo es bastante fuerte, al punto de que le impida cruzar la calle, e incluso paralizarla, ese grado de intensidad sería el inapropiado y, por ende, irracional; por el contrario, si ese miedo se da en un grado menor, por la simple creencia de que es una situación posible, pero no le paraliza y sólo le hace tener precaución de fijarse correctamente si no se acerca ningún auto, entonces puede decirse que ese grado de intensidad fue racional.

d) Este punto se puede analizar a través de su vía negativa, es decir que las emociones podrían ser irracionales si se basan en creencias que son irracionales con respecto a la situación en la que se plantean. Un ejemplo de esto sería la idea de querer ir a una obra de teatro, la persona que desea hacerlo decide investigar sobre la adquisición de boletos y se encuentra con que ya no hay disponibilidad de adquirirlos; pese a esto, el deseo de querer asistir le genera la creencia de que aún podría conseguirlos. En este punto ocurre que tanto la creencia como el deseo generan la emoción de esperanza de conseguir las entradas. El caso contrario sería que, dado que una creencia en la que se basa una emoción es racional, esta última también lo es; buscar información sobre la obra que se quiere ver, y que afirmativamente aún hay disponibilidad genera la creencia de que, de hecho, puede asistir, y de ello surja una emoción de alegría, que sería, por ende, racional.

e) Aunado lo anterior, la racionalidad de la emoción podría caer en la idea de que se fundamentara en la racionalidad del deseo. Para Elster (1997, p. 137), el que un deseo sea irracional sólo podría entenderse de dos maneras; por un lado, un deseo sería irracional sólo si surge por un mecanismo psíquico que conduzca a situaciones como el conformismo o su contrario (si hay conformismo el deseo simplemente es adaptado, si hay inconformidad el deseo surge únicamente por querer llevar la contraria, sin mayor fundamento), y por otro, sería irracional aquél deseo que conduzca a sufrimientos o desilusiones sin ninguna utilidad (desear lo que no se tiene y, de hecho, no se puede tener).

f) La hipótesis de que las emociones racionales son la que nos hacen felices tiene, claramente, la implicación directa de que todas aquellas que no producen felicidad son irracionales. Esto se vincula directamente con la disposición emocional más que con la

emoción que surge. Esta hipótesis en primera instancia podría parecer plausible, dado que se busca la optimización y, en general, el bienestar del individuo. Puede considerarse como ejemplo una contraposición entre la vergüenza y la culpa. La vergüenza es una emoción negativa que puede llegar a ser dolorosa para el individuo, lo que lo llevará a intentar evitarla a través de echarle la culpa a los demás por sus propios actos. La culpa, por su parte, normalmente induce al sujeto a actuar sobre sus propios actos, intentando reparar los errores o daños cometidos. Si se sigue este argumento, la culpa se consideraría una emoción racional. La vergüenza al ser, en este sentido, irracional, ¿tendría que eliminarse del individuo? Parece ser que no, y más aún, la vergüenza parece ser útil en otros aspectos, como en el cumplimiento de las normas sociales. El argumento puede ampliarse a todas las emociones, ya que no parecería viable para un sujeto mantener la alegría y eliminar la tristeza, dado que, de un modo u otro, estas se complementan.

Tras el análisis de cada uno de los puntos, aunque cada uno tiene una aportación positiva que podría considerar a las emociones como racionales, parece que de manera individual ninguna podría sostenerse. Algo que se puede inferir de esta discusión es que las emociones tienen muchísimos beneficios al momento de entrar como un factor en la toma de decisiones de acción. Citando a Elster (2001, p. 52), por ejemplo, “las tendencias a la acción provocadas por la ira y por el miedo inducen estados de alerta que capacitan al organismo para tratar con peligros o amenazas del entorno”. Sin embargo, las emociones también pueden inducir efectos en las que haya fallos e incluso interferencias tanto en la toma de una decisión como al momento de llevar a cabo una acción; Elster, citando a Frijda (1986), señala, por ejemplo, que “abundan las anécdotas referidas a la incapacidad de moverse o pensar cuando nos confrontamos con el peligro, [...] o cuando se nos apaga la última cerilla debido a los torpes movimientos de nuestros dedos”. Para Elster (2010, p. 264), de esto se sigue que las emociones pueden ser tanto útiles como adaptativas biológicamente, pero no racionales. Nussbaum (2006), sin embargo, afirma que las emociones “pueden ser evaluadas de la manera en que las creencias son evaluadas: como racionales o irracionales (respecto a la manera en que fueron adquiridas), como beneficiosas y nocivas, incluso como verdaderas o falsas”. Aquí defenderé la postura de Nussbaum. Dado que ninguno de los puntos puede sostener con claridad la posibilidad de la existencia de emociones racionales, parece no haber suficientes razones para aceptar que las hay; no obstante, sí que pueden ser evaluadas. Existen



casos en los que las emociones pueden entorpecer las decisiones tomadas, pero también pueden aportar información relevante para las decisiones de un individuo; en otras palabras, si bien existe una falibilidad de las emociones, de tal forma que puedan llevar a conclusiones irracionales, es posible que, mediante una educación de las mismas, las emociones pueden evaluarse críticamente de manera que pueden apoyar a tomar mejores decisiones.

### *3.3. Emociones adecuadas*

Uno de los puntos mencionados anteriormente desarrollaba la idea de plantear las emociones racionales a partir de la adecuación. Tal y como se dijo, esto no parecía satisfacer el concepto de una emoción racional; no obstante, es posible afirmar que, en palabras de Sousa, “el papel de la emoción es el de suplementar la insuficiencia de la razón. [...] Durante un tiempo variable, pero siempre limitado, una emoción limita la cantidad de información que el organismo va a tener en cuenta, las interferencias [...] y el conjunto de opciones relevantes de entre las que va a escoger” (citado en Modzelewski, 2017, p. 264); es decir que, cuando una emoción se introduce al momento de llevar a cabo una elección, ayuda a delimitar los aspectos relevantes para el individuo, permitiendo reaccionar adecuadamente, según sea la situación (un ejemplo de esto puede darse a través del miedo, que puede impulsar al individuo a reaccionar de manera rápida y eficaz ante una situación de peligro, en vez de detenerse a pensar en cuál sería la mejor línea de acción posible). Es por ello por lo que, pese a que la adecuación de las emociones no sea el elemento que pueda afirmar la posibilidad de hablar de emociones racionales, sí que parece ser el elemento que permite desarrollar la tesis de que las emociones son un elemento que, junto con la racionalidad, pueden ayudar a tomar mejores decisiones.

La adecuación se planteará a través de lo expuesto en 3.2.1. Dada la relación que hay entre las emociones y las creencias, una emoción será adecuada dependiendo de la creencia con la que se relaciona. Así mismo, la creencia será o no racional, tal y como se explicó en el capítulo uno, según la información recabada y disponible. De esto puede seguirse, por un lado, que una emoción pueda ser adecuada aún si la creencia es irracional, ya que esta última estaría en conflicto con la evidencia disponible, y, por otro lado, que una emoción pueda ser inapropiada, aunque se base en una creencia racional, es decir, la creencia concordaría con

la evidencia. Para una mayor comprensión de este punto analizaré entonces qué es lo que puede considerarse como una emoción irracional o inadecuada.

Lo dicho hasta ahora puede llevar a la idea de que aquello que hace irracional a una emoción son precisamente los antecedentes cognitivos del individuo, es decir, si hay una falla en la información recabada o bien en las creencias. Existen casos en los que se puede dar una emoción al recabar cierta información, limitando o afectando la formación de creencias adecuadas al hecho en cuestión. Otro caso es en el que resulta de originar una emoción de segundo orden para justificar una emoción que ha generado un cambio en las creencias. Los primeros casos se pueden ejemplificar a través de las obras de Shakespeare; Otelo, por ejemplo, sufre de una emoción de celos al adquirir cierta información sobre su amada Desdémona, lo que conlleva a una formación de creencias errónea y haciendo que él actúe en consecuencia. El segundo caso puede verse a través de la ira; crear una creencia que justifique la ira para hacer que el individuo se sienta bien; en la obra El señor de los anillos de J.R.R Tolkien se presenta un caso peculiar de este tipo. Gandalf, al despedirse de Bilbo Bolsón, le habla acerca del anillo, argumentando que tiene que dejarlo pues esa ha sido su promesa. Bilbo, cegado por el poder del anillo, se enfada, y justifica su negativa de hacerlo, afirmando que él nunca fue un ladrón y que es suyo, pues lo ganó justamente.

A través de los ejemplos anteriores se puede notar que las emociones no son irracionales por sí mismas; en otras palabras, no parecen, ni tienen por qué encontrarse desligadas de la razón, sino que pueden ser consideradas inadecuadas “en relación con las circunstancias en las que tienen lugar, sus objetos y las creencias que los acompañan” (Modzelewski, 2017, p. 269); si bien, el segundo ejemplo podría considerarse como un caso límite a esta cuestión, no es un caso conflicto de la razón con las emociones, sino que, por el contrario, dicha emoción se está viendo afectada por algo externo a esta relación. La adecuación es, por lo tanto, la forma en la que podemos hablar acerca de cómo se relacionan las emociones con la racionalidad.

### *3.3.1. Emociones adecuadas en el Dilema del prisionero*

En el capítulo anterior se desarrolló el análisis clásico del dilema del prisionero. El dilema dice lo siguiente:

Si confiesas haber participado en el crimen, y tu compañero no lo hace, entonces como premio te dejo libre, y tu compañero irá 20 años preso. En cambio, si ambos confiesan irán los dos a la cárcel por 5 años. Si ninguno de los dos decide confesar, entonces no tendré pruebas suficientes de que ustedes cometieron el crimen, pero los enviaré a la cárcel por un año por portación de armas. A tu compañero le hago la misma oferta.

La matriz, por lo tanto, queda de la siguiente manera:

	C	$\neg C$
C	(-5, -5)	(0, -20)
$\neg C$	(-20, 0)	(-1, -1)

Fig. 3.3

Los requisitos del dilema del prisionero excluían mecanismos como las amenazas para que el jugador pudiera tomar en cuenta todas las estrategias posibles y no verse obligado a escoger una en específico; la información disponible es limitada, pues sólo está cargada de las interacciones previas con el otro jugador, por lo que no hay un medio posible por el que se pueda conocer cuál será la estrategia tomada por el otro; dado que son dos los jugadores, ambos tienen que tomar en cuenta cuál será la estrategia del otro, es decir que no se puede asumir una respuesta suprimiendo al otro jugador, ni se puede evitar la interacción: las únicas posibilidades existentes es la cooperación o la defección; así mismo, no se puede controlar la respuesta del otro jugador.

Las respuestas que se dieron para el dilema del prisionero se centraron en aquéllas que podían considerarse como los equilibrios de Nash; en principio, parecía que la mejor respuesta era la defección, dado que sin importar cuál fuere la respuesta del otro, esa era la que nos evitaba el peor resultado. No obstante, aunque esa respuesta conlleva lo anterior, y a su vez mantiene la posibilidad de obtener el mejor resultado, al momento de comparar las consecuencias, se puede ver que lo conseguido resulta ser peor de lo que hubiera sido si se hubiera llevado a cabo la acción contraria, es decir, cooperar. Plantear la idea de la defección parece tener los mismos resultados que podrían derivar de un razonamiento meramente instrumental; en otras palabras, el individuo viendo sus intereses individuales trata de conseguir el mejor resultado para sí. Tal y como se desarrolló en el capítulo anterior, un razonamiento estratégico no sólo toma en cuenta las decisiones de un solo individuo, sino de

todo lo que se encuentra determinando la elección. Es decir, por un lado, en el caso del dilema, se toman en cuenta las decisiones tomadas por ambos individuos y, por otro lado, se reconocen que las reglas del juego limitan a responder el dilema según lo establecido. Si bien la cuestión estratégica abre la posibilidad de considerar enormemente la idea de cooperar, pues conlleva a un mejor resultado, el mero razonamiento estratégico no parece ser suficiente para llevar a cabo la decisión, pues, pese a que es el mejor resultado que podría surgir, ¿qué garantiza la cooperación, si la defección permitiría aprovecharse de la confianza de la otra persona y sacar el mejor resultado posible?

La adecuación de las emociones con la racionalidad es, por ende, la idea que aquí se plantea para la resolución del dilema; tal y como se dijo anteriormente, las emociones permiten delimitar los aspectos relevantes al momento de tomar una decisión. Las emociones, no obstante, no darán una respuesta concreta, es decir, el resultado variará dependiendo el caso. Tómese, por ejemplo, el caso pendiente de los juegos evolutivos planteados en el capítulo dos; una forma de interacción con otras personas puede llevarse a cabo a través de vínculos familiares, por lo que, si la otra persona que va a tomar lugar en el dilema tiene una relación de este tipo, la decisión que se tome se verá afectada por esta información añadida. Asimismo, esta información no afecta directamente a la cuestión de la racionalidad estratégica, sino que influye en las creencias debido a las emociones que se derivan de ello. La confianza de elegir concretamente la cooperación como respuesta al dilema se derivará, por tanto, de las emociones provocadas por esa conexión que se tiene con la otra persona. Por el contrario, si la otra persona es un completo desconocido se generará cierta desconfianza y, por ende, miedo al tomar la decisión de cooperar o defecionar.

Anteriormente se afirmó también que las reglas mediante las que se desarrolla el dilema pueden afectar los resultados; tanto que incluso se han realizado campeonatos mundiales del dilema del prisionero. Recuérdese que dos tipos importantes de juegos eran los finitos, que se dividían en juegos de una partida única o de una cantidad determinada de partidas, y los infinitos, o de una cantidad indeterminada de partidas. Analizados desde la relación existente entre la racionalidad y las emociones, los resultados de estos no sólo dependerían, ni se verían afectados únicamente, por las condiciones bajo las cuales el juego se efectúe, sino que también tendrían que ser evaluados a partir de esta información recabada por el tipo de

emoción que se produjera en cada caso. Del mismo modo, en ocasiones será la defección y en otras la cooperación, la respuesta que se determine como la adecuada para llevar a cabo. Tal y como se mencionó anteriormente, si bien las emociones no darán una respuesta consistente acerca de cuál será la mejor solución para el dilema, añadirán la suficiente información para que las creencias del individuo sean consistentes tanto con esta última, como con sus deseos, sus creencias ligadas al hecho en cuestión, y a la emoción misma, y, por ende, la respuesta dada al dilema sea la adecuada.

## Conclusiones generales

El problema abordado a lo largo de todo este trabajo fue el de la toma de decisiones. Se planteó la posibilidad de que las emociones pueden ser un elemento que apoye a la racionalidad para elegir la opción más adecuada. Para ello se hizo uso de tres cuestiones teóricas, a saber, la racionalidad paramétrica, la racionalidad estratégica y las emociones. A través de esto se llegó a las siguientes conclusiones generales:

- (1) No es posible responder cuál sería la mejor evaluación de una toma de decisiones con la racionalidad paramétrica, pues no existe tal cuestión como la perfección racional (la cual abría la posibilidad de considerar como racional incluso algunas cuestiones ideológicas, por ejemplo).
- (2) La racionalidad estratégica permite elaborar un mejor análisis para la toma de decisiones, no obstante, la racionalidad parece ser insuficiente para responder casos complejos, como se detalló a través del dilema del prisionero. Del mismo modo, se rechaza la tesis de la irracionalidad pues, aunque se pueda demostrar que no siempre se plantea un análisis racional, no resulta ningún contraejemplo ni parece ser el método más útil para llevar a cabo una toma de decisión.
- (3) Las emociones si bien no pueden considerarse racionales, pueden plantearse como adecuadas dependiendo de la situación discutida; en otras palabras, las emociones otorgan información adicional que, junto con la racionalidad, permitiría desarrollar una mejor evaluación en una toma de decisión.

A través de la teoría de la elección racional se examinó la idea de que la racionalidad por sí misma podía llevar a cabo la mejor evaluación de una toma de decisiones. Para ello se analizó la racionalidad paramétrica, en la que se postula al individuo como racional; esto abre la posibilidad de una evaluación crítica y, al mismo tiempo, posibilita el autoescrutinio, pues para que algo se considere racional se debe tener coherencia en el conjunto de creencias, deseos, preferencias, e información disponible.

Se tienen que tener en cuenta también los aspectos dejados de lado por la racionalidad paramétrica, a saber, los individuos con los cuales uno se relaciona y el medio cambiante que

puede ser un factor que modifique o determine una evaluación de elección. Estos elementos son tomados de la teoría de juegos, la cual podría considerarse una extensión de la teoría de la elección racional. De estos elementos se derivan, por un lado, las normas, o las reglas del juego, y, por otro lado, las estrategias, es decir, los planes de acción que describen las reacciones posibles de un individuo ante alguna circunstancia que se le presente, como sería el caso de que las preferencias o la toma de decisiones de dos o más individuos se vean involucradas entre sí.

Se planteó un caso específico en el que las estrategias de dos individuos pueden verse enfrentadas y la decisión de cada uno puede afectar o contribuir al otro, a saber, el dilema del prisionero. Es un caso de cooperación o defección. Desde una postura racional se puede considerar el caso de la defección como la respuesta racional, ya que evita el peor de los resultados y, a su vez, mantiene abierta la posibilidad de obtener el mejor de los resultados posibles. No obstante, al considerar que el otro puede tener la misma estrategia, el resultado final no parece ser el mejor resultado que podría obtenerse, es decir, la cooperación puede ser la mejor opción. Sin embargo, aunque la racionalidad estratégica también toma en cuenta esta posibilidad, no la toma como segura. En otras palabras, parece que la racionalidad no es suficiente para que se lleve a cabo una buena evaluación acerca de cuál sería la mejor estrategia, o elección.

Las emociones se postulan, por lo tanto, como el elemento que puede llevar, junto con la racionalidad, a considerar cuál podría ser la mejor elección posible. Para ello se realizó una breve evaluación acerca de qué se entiende cuando se habla de emociones. Las emociones tienen algunas características que pueden distinguirlas, a saber: Su sensación cualitativa, la aparición súbita, la imprevisibilidad, la corta duración, el desencadenamiento cognitivo, el estar dirigidas a un objeto intencional, la expresión fisiológica, la tendencia para realizar acciones y la valencia.

Dado que la importancia de esta tesis se centra en la relación que hay entre la racionalidad y las emociones, se da mayor énfasis al desencadenamiento cognitivo. En este punto, por lo tanto, se añaden las emociones al esquema presentado en el capítulo uno sobre la coherencia existente entre las creencias, los deseos y la información disponible. Por ende, las emociones

así integradas tienen que mantener una coherencia directa tanto con las creencias y con los deseos y, de manera secundaria, también deberán ser consistentes con la información.

Con relación a la elección racional, las emociones pueden controlarse de tal manera que puedan considerarse parte de una estrategia que maximice el resultado posible. Este control puede llevarse a cabo a través de una educabilidad de las emociones, o bien mediante una evaluación sobre ella, a través de una metaemoción, que permita ver la coherencia que tiene esta con la decisión que se llevará a cabo. Esto conlleva a preguntarse si las emociones por sí mismas podrían considerarse como racionales. No obstante, dado la inestabilidad que a veces puede encontrarse en ellas se concluirá que las emociones son, más bien, adecuadas.

Las emociones adecuadas son aquellas que pueden limitar la información que el individuo tomará en cuenta, así como las posibles interferencias que pudieran surgir por diferentes factores, así como las opciones que tienen mayor peso para el individuo de entre todas las disponibles. Concluyendo así, la afirmación a la tesis planteada, a saber: Las emociones adecuadas, junto la racionalidad, permiten llevar a cabo una mejor evaluación acerca de cuál sería la estrategia y la elección que podría considerarse la más óptima.

Mucho de lo discutido aquí sigue vigente, por lo que constantemente se elaboran análisis, objeciones y modificaciones para lo planteado. Una de las grandes desventajas de esta tesis puede ser lo amplio que puede ser cada una de las propuestas utilizadas para hablar únicamente de algunos elementos. Tal y como se dijo anteriormente, la teoría de la elección racional ha tenido varias modificaciones, tales que incluso podría salvar alguna de las objeciones que aquí se le presentan. Los otros dos grandes temas aquí expuestos es la teoría de juegos y las emociones. No obstante, presentarlo de esta manera me permitió mostrar de manera más clara cada uno de los puntos principales rescatados de estos temas. Del mismo modo, otra de las desventajas que puede presentar mi tesis se desarrolla justo en el planteamiento de las emociones. Éste es un concepto bastante contemporáneo y, cada vez que se lleva a cabo un estudio sobre cómo se ha discutido con anterioridad este tema se encuentra con diversas clasificaciones y conceptualizaciones que pueden generar problemas de especificar de manera concisa qué se está entendiendo por emoción; existe una amplia gama de conceptos con los cuáles se relaciona y que muchas veces se tienden a confundir, tales como las pasiones o los sentimientos; además, la discusión sigue vigente con respecto



a su naturaleza. No obstante, creo que una de las ventajas de esta tesis es que, si bien no se da un esbozo que especifique claramente qué se entiende por una emoción, puede seguirse la tesis principal sin importar qué postura tenga uno al respecto, o dicho de un modo más fuerte, cuál sea la que tenga razón dentro de la discusión.

Estas desventajas podrían traer consigo la posibilidad de dejar algún caso importante por desarrollar dentro del tema aquí planteado. A continuación, presentaré cuatro temas que quizá podrían desarrollarse un poco más en investigaciones posteriores o que, por falta de evidencia empírica, no hayan podido desarrollarse del todo en esta tesis; trataré de dar una posible respuesta o un camino posible de adónde podría dirigirse la investigación.

- ¿La irracionalidad no puede dar cuenta de igual manera que la racionalidad una toma de decisión, tal y como algunos teóricos de los juegos señalan?

Esta postura señala a través de unos juegos que, a veces, para ganar, no se necesita pensar racionalmente cuál sería la mejor jugada posible (porque tal vez la cantidad de posibilidades es impensable o bien porque no lo necesita). Demuestra incluso que en la naturaleza se pueden dar estrategias similares, confirmando que no se necesita un ser que pueda pensar claramente cuál es, de hecho, la mejor estrategia disponible.

Concedo un poco que en ocasiones es posible llevar a cabo una elección mínima, en la que no se necesita elaborar una evaluación exhaustiva acerca de qué preferencia tenga un individuo entre  $x$  y  $y$ . En otras palabras, la idea de tomar decisiones podría considerarse a través de niveles, que irían desde cuestiones mínimas tales como preferencias sobre qué alimento se consumirá, hasta casos más concretos, en los que sea complicado determinar cuál de las opciones escoger. Evidentemente, la cuestión de la irracionalidad se dejaría para las cuestiones de elección mínima e incluso nula, como el caso de los reflejos. Aún cediendo este punto, cabría preguntarse si, de hecho, este tipo de elecciones podrían considerarse irracionales, pues incluso en éstas hay una cuestión cognitiva y emocional previa que permite llevar a cabo estas acciones aun incluso aunque se hubieran hecho “sin pensarse”. Un ejemplo de esto podría ser el de colocar la mano en el fuego, pues se sabe que no debe hacerse porque puede hacer daño. Saber esto normalmente se acompaña de un miedo moderado de hacerlo; el reflejo que se da cuando esta situación se presenta espontáneamente, aunque en principio es eso, un acto inmediato de quitar lo que esté al alcance del fuego, viene acompañado de

algunas creencias, deseos, información y carga emocional que hacen que el individuo actúe de esa manera. Incluso se pueden pensar casos en los que no exista este conocimiento previo y, por ende, el reflejo no se dé naturalmente; por ejemplo, un bebé cayendo porque no pudo mantener su equilibrio, debe enseñársele que debe tratar de caer con las manos para evitar sufrir un impacto más fuerte; si bien a veces se puede dar este movimiento naturalmente, no siempre es el caso.

- Los valores y las normas sociales, al ser reguladores de conducta, de emociones, entre otras cosas, ¿no pueden dar cuenta por sí mismas sobre el cómo se debe actuar y, por ende, sobre qué elección se debe tomar?

Si bien sirven como reguladores de conducta y, por ende, limitantes para realizar ciertas acciones posibles, es posible rechazar esto porque, aunque éstas afirmen muchas cosas que son permitidas o rechazadas socialmente, no alteran directamente las decisiones. En otros términos, considérense las reglas de juego del ajedrez. Por un lado, las reglas generales limitan los objetivos del juego: Pueden ganar las blancas, pueden ganar las negras o puede haber un empate, y nada más; por otro lado, hay reglas específicas al respecto de cómo debe moverse cada pieza. Con estas reglas puestas previamente al juego, éste puede presentar una cantidad indeterminada de movimientos posibles mientras se efectúa. Las reglas serían unos reguladores de conducta que indican, por ejemplo, los espacios determinados en los que se puede mover una pieza, y no la elección misma de dónde colocar la pieza. Dicha analogía podría considerarse para las normas y los valores. Restringen qué líneas de acción pueden llevarse a cabo, pero no eliminan la libertad de elegir alguna de las opciones que se presenten en una toma de decisiones; además de que no deciden cuál de las estrategias sería la más adecuada para cada hecho.

- Las emociones durante mucho tiempo se han excluido y se han contrapuesto a la racionalidad, dado que éstas pueden aparecer de una manera descontrolada nublando la capacidad de tomar la mejor decisión.

Una posible investigación podría ahondar más en este tema, que es fundamental para los resultados examinados al final de esta tesis, a través, quizá, de evidencia empírica. No obstante, es posible responder el punto en cuestión a través de esta tesis; para ello concedo un poco para rechazar este planteamiento. Es cierto que una emoción puede presentarse de

una manera descontrolada, haciendo que el individuo que la presenta actúe de una manera desmedida e incluso contraria a la que podría esperarse si elaborara una evaluación del asunto. Una persona enojada puede actuar de manera violenta aún cuando la razón de que haya actuado así sea, por decirlo de algún modo, insignificante. En esta tesis se presentan dos alternativas que, con respecto a la toma de decisión, pueden ser un apoyo para que la emoción sea adecuada. Por un lado, es importante que se lleve a cabo una educabilidad de las emociones para saber en qué momentos se han presentado de manera irracional y bajo qué hechos se podría producir esto para evitarlo; por otro lado, analizar las emociones a través de las metaemociones, pues al evaluar la emoción uno puede desarrollar un autoescrutinio sobre por qué se actuó de esa manera. Esta tesis, si bien no se centra en desarrollar cómo puede darse esta educabilidad emocional, pues eso sería algo fuera del alcance de la misma, es cierto que abre la posibilidad de evitar que se presenten estos casos. En otras palabras, si bien es cierto que en ocasiones las emociones impiden actuar racionalmente, evaluar las emociones tanto las consideradas positivas como negativas, permite saber por qué algo le provoca esa sensación de enojo o de alegría y, por ende, le permite también actuar adecuadamente.

- A lo largo de toda la tesis se habló implícitamente acerca de una cuestión moral, a saber, qué acciones pueden llegar a considerarse como correctas o incorrectas. Algo que faltó desarrollar a lo largo de la tesis es cómo se introduce esta idea. En el caso del dilema del prisionero, si bien por sí mismo se puede comprender por qué sería adecuada la cooperación en vez de la defección, cabría preguntarse por separado por qué se considera, de hecho, como una mejor opción la cooperación y no la defección en otros ámbitos.

Este tema es un punto muy importante por rescatar pues, si bien no es algo que se aborde detalladamente dentro de la tesis, es un asunto recurrente, ya que incluso el introducir alguna ejemplificación ya conllevaría un juicio previo y, en muchas ocasiones, un juicio moral. Esto, por ende, es inevitable; sí, se introducen ciertos juicios morales, por mor a la comprensión del tema. No obstante, y como bien se señala, este es un tema que no se profundiza pues no es el objetivo de esta tesis. Por ello, prefiero dejar este tema abierto, esperando que en un trabajo posterior pueda desarrollarse de manera más detallada.

Señalaré ahora qué puede derivarse de esta tesis. En primer lugar, dado que queda abierta esta cuestión moral, podría considerarse cómo afectarían los resultados de esta tesis en ámbitos como la responsabilidad moral. Al aceptar la relación de las emociones adecuadas y la racionalidad para tomar decisiones podría desprenderse el cómo afectarían dichas discusiones con respecto a la responsabilidad. Por poner un ejemplo claro, actualmente los juicios que se llevan a cabo en el ámbito del derecho, al evaluar si una persona es culpable o no de haber cometido un delito, se tiene como principio el hacerlo meramente a través de la razón, impidiendo la apelación a las emociones. En relación al mismo tema de la responsabilidad, se podría abordar el tema desde ámbitos de filosofía de la mente, ya que incluso se discute la posibilidad de si un robot podría ser responsable del asesinato de una persona, ya que actúa a través de algoritmos y no de manera racional ni, mucho menos, con una carga emocional; dicho ejemplo podría considerarse análogo con respecto a personas que tienen algún problema cerebral que le impide sentir emociones y, por ende, no tener una idea clara sobre si asesinar a una persona es o no algo correcto moralmente y, por ende, se duda de su responsabilidad. En segundo lugar, queda abierta la cuestión acerca de la clasificación de las emociones, pues, como se sabe, es algo que continúa discutiéndose. Y de este punto podría seguirse una cuestión pedagógica, e incluso epistemológica sobre la educabilidad de las emociones. Tal y como se afirmó anteriormente, el hablar de una educabilidad emocional permitiría tanto el reafirmar la tesis planteada aquí, como el continuar en un ámbito social. Desde un ámbito pedagógico podría plantearse la idea de modificar las políticas que se encarguen al respecto y se permita enseñar, a nivel escolar, educación emocional, y, de esto, derivar la cuestión epistemológica, a saber, modificar la idea socialmente aceptada de que las emociones se encuentran peleadas con la razón.

## Bibliografía

- Abitbol, P. y F. Botero (2006), “Teoría de elección racional: estructura conceptual y evolución reciente”, en *Colombia Internacional*, n. 62, Colombia.
- Ainslie, G. (2001), *Breakdown of Will*, Cambridge University Press, Cambridge.
- Amster, P. y P. Pinasco (2014), *Teoría de juegos, una introducción matemática a la toma de decisiones*, FCE, México.
- Anónimo (2016), *Las mil y una noches*, E-book:  
[https://books.google.com.mx/books?id=aaLwCwAAQBAJ&printsec=frontcover&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](https://books.google.com.mx/books?id=aaLwCwAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false)
- Arnold, M.B. (1960), *Emotion and Personality*, Columbia University Press, Nueva York.
- Ashworth, T. (1980), *Trench Warfare, 1914-1918: The Live and Let Live System*, Holmes & Meier, Nueva York.
- Axelrod, R. (1986), *La evolución de la cooperación*, trad. Luis Bou, Alianza, Madrid.
- Bernheim, D. (1984), “Rationalisable strategic behaviour”, *Econometrica*, vol. 52.
- Broad, C. (1971), *Critical Essays in Moral Theory*, Allen and Unwin, Londres.
- Budd, M. (1995), *Values of Art*, Allen Lane, Londres.
- Camps, V. (2011), *El gobierno de las emociones*, Herder, Barcelona.
- Daboni, L., Montesano, A.M. y Lines, M. (Comps.) (1986), *Recent Developments in the Foundation of Utility and Risk Theory*, D. Reidel Publishing Company, Boston.
- Damasio, A. (1996), *El error de Descartes*, Mondadori, Barcelona.
- (2010), *Y el cerebro creó al hombre: ¿cómo pudo el cerebro generar emociones, sentimientos, ideas y el yo?*, Destino, Barcelona.
- Dancy, J. (1985), *Introducción a la epistemología contemporánea*, Tecnos, Madrid.
- Davidson, D. (1963), “Actions, reasons and causes”. *Journal of Philosophy*, vol. 60, no. 23,  
[<http://www.jstor.org/stable/2023177>]

- (1994), “Davidson, Donald”, en Guttenplan Samuel (ed.), *A companion to the Philosophy of Mind*, Blackwell, Oxford.
- (1995), *Ensayos sobre acciones y sucesos*, Crítica - IIF/UNAM, México.
- Davis, M. (1986), *Introducción a la teoría de juegos*, trad. José Carlos Gómez Borrero, Alianza, Madrid.
- Denton, D. (2009), *El despertar de la consciencia: la neurociencia de las emociones primarias*, trad. Joandomènec Ros, Paidós, Barcelona.
- Di Castro, E. (2009), *La razón desencantada: Un acercamiento a la teoría de la elección racional*, IIF-UNAM, México.
- Edgeworth, F. (1881), *Mathematical Psychics: An essay on the application of Mathematics to the moral sciences*, University of California Libraries, Berkeley.
- Ekeland, I. (1992), *Al azar: La suerte, la ciencia y el mundo*, Gedisa, Barcelona.
- Ekman, P. (1980), “Biological and cultural contributions to body and facial movement in the expression of the emotions”, en A. Rorty (comp.), *Explaining the Emotions*, University of California Press, Berkeley.
- Ekman, P., R. Levenson y W. Friesen (1983), “Autonomic Nervous System Activity Distinguishes among Emotions”, *Science*, vol. 221, no. 4616.
- Elster, J. (2002), *Alquimias de la mente: La racionalidad y las emociones*, Paidós, Barcelona.
- (1997), *Economics: Análisis de la interacción entre racionalidad, emoción, preferencias y normas sociales en la economía de la acción individual y sus desviaciones*, Trad. Irene Cudich, Gedisa, Barcelona.
- (2010), “Emotional choice and rational choice”, en Goldie (ed.), *The Oxford Handbook of Philosophy of Emotion*.
- (ed.) (1986), *Rational Choice*, New York University Press, Nueva York.
- (2008), *Reason and Rationality*, Princeton University Press, Nueva Jersey.
- (2001), *Sobre las pasiones*, trad. J.F. Álvarez y A. Kiczkowski, Paidós, Barcelona
- (1989), *Solomonic Judgements: Studies in the Limitations of Rationality*, Cambridge University Press, Nueva York.

- (2014), *Ulises y las sirenas: Estudios sobre racionalidad e irracionalidad*, trad. Juan José Utrilla, FCE, México.
- (1985), *Uvas amargas: sobre la subversión de la racionalidad*, trad. Enrique Lynch, Península, Barcelona.
- Frijda, N. (1995), “Las leyes de la emoción”, en M.D. Avia y M.L. Sánchez (comps.), *La personalidad: Aspectos cognitivos y sociales*, Pirámide, Madrid.
- (1986), *The Emotions*, Cambridge University Press, Cambridge.
- Goldie, P. (ed.) (2010), *The Oxford Handbook of Philosophy of Emotion*, Oxford University Press, Nueva York.
- Hamilton, W. (1963), “The evolution of altruistic behaviour”, *American Naturalist*, vol. 97.
- (1971), “Selection of selfish and altruistic behavior in some extreme models”, *Man and Beast: Comparative Social Behaviour*, Eisenberg F. y W. Dillon (comps.), Smithsonian Press, Washington.
- Hargreaves, S. y Y. Varoufakis (1995), *Game Theory. A Critical Introduction*, Routledge, Londres y Nueva York.
- Hochschild, A. (1983), *The Managed Heart*, University of California Press, Berkeley.
- Hume, D. (1994), *Investigación sobre el conocimiento humano*, Alianza, Madrid.
- Jeffrey, R. (1983), *The Logic of Decision*, McGraw-Hill, Nueva York.
- Kahneman, D. y A. Tversky (comps.) (1982), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge.
- Kant, E. (2012), *Crítica de la razón pura*, Taurus, Madrid.
- Kreps, D. (1990), *Game Theory and Economic Modeling*, Oxford University Press, Nueva York.
- Levi, I. (1986), *Hard Choices: Decision Making Under Unresolved Conflict*, Cambridge University Press, Cambridge.
- Levin, J. y P. Milgrom (2004), “Introduction to Choice Theory”, en línea:  
<https://web.stanford.edu/~jdlevin/Econ%20202/Choice%20Theory.pdf>
- McFadden, D. (1999), “Rationality for economists”, *Journal of Risk and Uncertainty*, no. 19.

- Modzelewski, H. (2017), *Emociones, educación y democracia: una proyección de la teoría de las emociones de Martha Nussbaum*, IIF-UNAM, México.
- Mora Burgos, G. (2003), “Racionalidad y tipos de racionalidad”, *Revista de Filosofía de la Universidad de Costa Rica*, vol. 41, no. 103.
- Nash, J. (1951), “Non-cooperative games”, *Annals of Mathematics*, vol. 54.
- Noguera, A. (2013), “¿Qué es el individualismo metodológico?” en *La era de los individuos: Actores, política y teoría en la sociedad actual*, Carlos Andrés Charry y Nicolás Rojas Pedemonte, LOM ediciones, Chile.
- Nozick, R. (1993), *La Naturaleza de la Racionalidad*, Paidós, Barcelona.
- Nussbaum, M. (2006), *El conocimiento del amor: ensayos sobre filosofía y literatura*, trad. Rocío Orsi y Juana María Inarejos, A. Machado, Madrid.
- (2008), *Paisajes del pensamiento: La inteligencia de las emociones*, trad. Araceli Maira, Paidós, Barcelona.
- Pascal, B. (1981), *Pensamientos*, trad. J. Llansó, Alianza, Madrid.
- Pearce, D. (1984), “Rationalisable strategic behavior and the problem of perfection”, *Econometrica*, vol. 52.
- Pears, D. (1985), *Motivated Irrationality*, St. Augustine’s Press, Michigan.
- Rubinstein A. (1998) *Modeling Bounded Rationality*, The MIT Press, Cambridge.
- Sen, A. (1987), *Ética y economía*, Alianza, Madrid.
- (2012), *La idea de la justicia*, Taurus, Madrid.
- (2002), *Rationality and Freedom*, Harvard University Press, Cambridge.
- Schick, F. (1997), *Hacer elecciones: Una reconstrucción de la teoría de la decisión*, Gedisa, Barcelona.
- Simon, H. (1955), “A Behavioral Model of Rational Choice”, *The Quarterly Journal of Economics*, vol. 69. no. 1.
- (1979), *Models of Thought*, Yale University Press, New Haven.



- (1978), "Rationality as a Process and a Product of Thought". *The American Economic Review*, vol. 68, no. 2.
- (1986), "Rationality in Psychology and Economics". En Hogarth, Robin M. y Reder, Melvin W. (eds.), *The Behavioral Foundations of Economic Theory; Journal of Business (supplement)* vol. 5.
- Solomon, R. (1980), "Emotions and choice", en A. Rorty (comp.), *Explaining the Emotions*, University of California Press, Berkeley.
- (2007), *Ética emocional: Una teoría de los sentimientos*, trad. Pablo Hermida, Paidós, Barcelona.
- (1993), *The Passions*, Hackett, Indianapolis.
- Stigum B. y F. Wenstop (eds.) (1983), *Foundations of utility and risk theory with applications*, D. Reidel Publishing Company, Boston.
- Taylor, M. (1976), *Anarchy and Cooperation*, Wiley, Nueva York.
- Thaler R. (1991), *Quasi-rational economics*, Russell Sage Foundation, Nueva York.
- Tuck, R. (2008), *Free Riding*, Harvard University Press, Cambridge.
- Udehn L. (2001), *Methodological Individualism: Background, History and Meaning*, Routledge, Londres.
- Vidal de la Rosa, G. (2008), "La teoría de la elección racional en las ciencias sociales", *Sociológica*, no. 67, México.
- Winter, S. (1964), "Economic 'natural selection' and the theory of the firm", *Yale Economic Essays*, no. 4.