

UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

FACULTAD DE PSICOLOGÍA

NIVELES COGNITIVOS Y CREENCIAS EN
JUEGOS

T E S I S

QUE PARA OBTENER EL GRADO DE:

LICENCIATURA EN PSICOLOGÍA

PRESENTA:

JAIME OSVALDO ISLAS FARIAS

DIRECTOR DE TESIS:

DR. ARTURO BOUZAS RIAÑO

REVISOR:

DR. ÓSCAR ZAMORA ARÉVALO

SINODALES:

MTRO. MIGUEL HERRERA ORTÍZ

LIC. LIGIA COLMENARES

LIC. JOSÉ LUIS BAROJA MANZANO

Con el apoyo de:

Proyecto PAPIIT IN307214



CIUDAD DE MÉXICO

15 DE ABRIL DE 2018

Niveles Cognitivos y Creencias en Juegos

por

Jaime Osvaldo Islas Farias

Tesis presentada para obtener la

Licenciatura en Psicología

en la

Facultad de Psicología

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad de México, 15 de abril de 2018

Resumen

Niveles Cognitivos y Creencias en Juegos

por Jaime Osvaldo Islas Farias

Un concepto clave en el estudio de Teoría de Juegos es el razonamiento iterado, proceso en el que los jugadores incorporan las creencias que tienen sobre la conducta de los otros jugadores en su toma de decisiones. P-Beauty contest es un juego utilizado para estudiar este proceso, en el que los participantes eligen simultáneamente un número en el rango $[0 - 100]$ y el jugador que haya elegido el número más cercano a $2/3$ del promedio de todos los números es el ganador. Personas reales no se comportan de acuerdo al equilibrio de Nash del juego (elegir 0), pero se acercan a este paulatinamente en juegos repetidos. Se han propuesto modelos que dan cuenta de la conducta real de las personas, capturando la noción de que la elección es un reflejo del número de pasos de razonamiento iterado que las personas son capaces de realizar (nivel cognitivo) y de las creencias que tienen sobre el nivel cognitivo de los demás. Esta tesis estudia la consistencia entre estas elecciones y creencias, y la forma en que esta relación cambia cuando se adquiere experiencia en el juego, al jugarlo de forma repetida. Se utilizó un diseño experimental de juegos repetidos de p-beauty contest con dos subjuegos de 4 periodos. Un jugador participó en los dos subjuegos, mientras que los demás jugadores fueron sustituidos entre subjuegos. Esta manipulación garantiza que un participante cuenta con más experiencia que los demás en el segundo subjuego, y revierte parcialmente la tendencia a elegir números más cercanos a 0 en juegos repetidos (efecto de reset). Además de registrar su número, los participantes indicaron qué número creían que iban a elegir los demás (explicitación de creencias). Se encontró que la entrada de nuevos participantes al juego incrementa el número elegido de los jugadores con experiencia, pero la diferencia entre sus creencias y elecciones no incrementa junto con este, lo que aporta evidencia de que la consistencia entre creencias y elecciones incrementa con la experiencia.

Abstract

Cognitive Levels and Beliefs in games

by Jaime Osvaldo Islas Farias

A key concept in game theory is iterated reasoning, the process in which players take into account the beliefs they have about the behavior of other players in their decision making. P-beauty contest is a game that is used to study this process, in which participants choose simultaneously a number in the range $[0-100]$, and the player that selected the number closest to $2/3$ of the mean of all numbers is the winner. Real people don't behave according to the Nash equilibrium of the game (choosing 0), but get closer to it in repeated games. Models have been proposed that explain the actual behavior of people, capturing the notion that choices are an indication of the number of steps of iterated reasoning that people are capable of taking (cognitive level) and the beliefs they have about the cognitive level of others. This thesis studies the consistency between choices and beliefs, and the way this relationship is affected by acquiring experience in the game, product of repeated play. An experimental design of repeated p-beauty contest games with two subgames of 4 periods each was used. One participant played in both subgames, while the other players were changed between subgames. This manipulation guaranteed that one player had more experience than the others in the second subgame, and also partially reverted the trend of choosing numbers closer to 0 seen in repeated games (reset effect). Besides recording the chosen number, participants had to state the number they believe the other players would choose (elicited beliefs). It was found that the introduction of new players in the game indeed increased the chosen number of experienced players, but the difference between their beliefs and choices did

not increased. This supports the idea that consistency between beliefs and choices increases with experience.

Agradecimientos

A Jaime Islas Curiel, mi papá, por siempre apoyar mis decisiones, incluso si estas no siempre son las más acertadas. Es gracias al esfuerzo con el has salido a trabajar todos los días durante todos estos años que he podido llegar hasta este punto en mi vida académica, profesional y personal. Gracias por enseñarme (con el ejemplo, por si fuera poco) a ser una persona con principios y valores.

Póstumamente a Beatriz Farias Orozco, mi mamá, sin quien, literalmente, yo no estaría aquí. Es también gracias a tu ejemplo (y tus regaños) que me he convertido en el hombre que soy hoy. Quiero pensar que estarías muy feliz por este logro, y orgullosa de mí.

A Adriana Felisa Chávez De la Peña, mi princesa, el amor de mi vida. Gracias a tu ayuda y apoyo (académico y emocional) esta tesis es una realidad. El amor que te tengo y la expectativa de hacer mi vida contigo me hacen mejorar continuamente como persona. Gracias por ser mi mejor amiga, compañera y el amor de mi vida.

Si los amigos son la familia que uno escoge, entonces la familia son los amigos que no escogemos, pero que están ahí incondicionalmente y durante toda nuestra vida. Gracias a toda mi familia por su cariño y apoyo permanente; a mi abuelita y a mis tíos por ser como mis segundos papás, y a mis primos por ser como mis hermanos.

A mis amigos, a quienes no enlisto, porque ellos saben quiénes son (y así evito momentos incómodos si me olvido de mencionar a alguno). La vida está llena de giros inesperados y nunca sabemos dónde terminaremos el día de mañana. Pero aunque no siempre somos una presencia constante en nuestras vidas, y coincidimos solo de forma intermitente, sabemos que nuestra amistad se mantiene. Gracias por todos los momentos divertidos y entrañables.

A todos los masoquistas del Lab 25, que pudiendo hacer cosas más sencillas, siempre eligen ir más lejos y aceptar los mayores desafíos. Gracias por poner la vara tan alta en la calidad de todos sus proyectos; estar rodeado de estudiantes

tan brillantes y trabajadores inevitablemente me impulsó a esforzarme más para intentar alcanzar los estándares establecidos por todos ustedes. Espero que la consideren esta tesis una entrada digna al catálogo de investigaciones realizadas en el Lab 25.

A todos los profesores que participaron en mi formación; para bien o para mal, esta tesis y lo que venga después son un producto parcial de sus enseñanzas. Particularmente agradezco a los miembros de mi comité: A Ligia Colmenares, por sus excelentes comentarios y observaciones al borrador de esta tesis. A Miguel Herrera, porque fueron sus clases en los primeros semestres de la carrera lo que me hizo acercarme al área de la psicología experimental, y en los últimos semestres a teoría de juegos. A José Luis Baroja, con quien he tratado como profesor así como compañero en el Lab 25, y en ambos casos he aprendido muchísimo; considero que José Luis es el modelo ejemplar de cómo debe ser un investigador. A Oscar Zamora, cuyas excelentes clases fueron también un pilar en mi formación y contribuyeron a mi amor por el área. Y finalmente, de manera particular, al “Doc” Arturo Bouzas. Gracias por sus consejos siempre pertinentes y concretos, que permitieron volver una realidad esta tesis, desde que comenzó como solo una idea vaga y difusa. Admiro muchísimo su brillantez y sabiduría, producto de una vida de experiencia académica. Gracias a todos por compartirme su conocimiento y experiencia en mi formación como investigador. A mis compañeras de trabajo en el CEPE. Por abrirme las puertas al atemorizante mundo de los adultos de forma relativamente indolora, y enseñarme que algunas de las habilidades que nos enseñan en la escuela son, de hecho, utilizadas en la vida real. Son todas un ejemplo de responsabilidad, esfuerzo y pasión por el trabajo, y mi punto de referencia sobre cómo hay que desenvolverse en el mundo profesional.

A ti, misterioso lector. Que estés leyendo mi tesis justifica su redacción más allá de su función como requisito para titularme.

Índice general

Resumen	IV
Agradecimientos	VII
1. Introducción	1
2. Marco Teórico	7
2.1. Modelo de nivel-k	7
2.2. Relación entre creencias y elecciones	8
2.3. El efecto de la experiencia	11
2.4. Objetivo	12
3. Método	15
3.1. Participantes	15
3.2. Diseño Experimental	15
3.3. Procedimiento	16
4. Resultados	19
4.1. Análisis Bayesiano	19
5. Discusión	39
6. Conclusión	43

Índice de figuras

Índice de cuadros

Listado de Abreviaturas

Capítulo 1

Introducción

Teoría de juegos es una rama de las matemáticas que estudia la toma de decisiones en situaciones de interacción. Dentro de teoría de juegos, se estudian situaciones como la que representa el juego del ciempiés (estudiada por primera vez por Rosenthal, en 1982). Se trata de un juego donde dos jugadores, A y B, toman turnos para elegir si quedarse con una ganancia acumulada o pasar la ganancia al otro jugador. Si algún jugador elige quedarse con la ganancia acumulada, el otro jugador no recibe nada y el juego termina. En cambio, si el jugador elige pasar, el tamaño de la ganancia acumulada aumenta. El juego dura 10 turnos, y la ganancia acumulada aumenta su valor en cada uno. Si en el último turno, que le corresponde al jugador B, este elige pasar, el jugador A se queda con todas las ganancias.

De acuerdo con las reglas del juego, si los jugadores quisieran maximizar sus ganancias, deberían elegir ‘pasar’ para que la ganancia acumulada siga aumentando, pero al mismo tiempo, dado que ambos jugadores quieren quedarse con la ganancia, deben elegir quedarse con ella antes de que lo haga el otro jugador. ¿En qué turno del juego sería una buena elección (para cada jugador) quedarse con la ganancia acumulada?

El juego establece que en el último turno, cuando la ganancia acumulada tiene su valor más alto, le corresponde elegir al jugador B. Si este jugador elige pasar, todas las ganancias serán para el jugador A, mientras que si decide quedarse con la ganancia acumulada, B obtendrá la mayor ganancia posible. Por lo tanto, el jugador B debería elegir quedarse con la ganancia en el turno 10. Sin embargo, este razonamiento no es un secreto para el jugador A. Si el jugador A sabe que al

jugador B le conviene quedarse con la ganancia en el turno 10 (en cuyo caso A se quedaría sin nada), entonces la mejor estrategia para el jugador A consiste en quedarse con la ganancia acumulada en el turno 9. Pero así como el jugador A es capaz de anticipar que el jugador B elegirá quedarse con la ganancia en el turno 10, el jugador B puede deducir que el jugador A elegirá quedarse con la ganancia en el turno 9. Si este es el caso, el jugador B debería elegir quedarse con la ganancia en el turno 8. Si ambos jugadores repiten este razonamiento para todos los turnos de juego, eventualmente se llega a la conclusión de que la mejor estrategia para cualquiera de los jugadores en cualquier turno es elegir quedarse con la ganancia, sin importar el turno en que se encuentren. A este proceso de repetir un razonamiento para llegar a la mejor estrategia disponible tomando en cuenta que el otro jugador también puede hacerlo, se conoce como razonamiento iterado, y es uno de los conceptos básicos en teoría de juegos.

A las situaciones donde los jugadores eligen la estrategia que maximice su ganancia esperada, tomando en cuenta las estrategias de los otros jugadores, se les conoce como el equilibrio de Nash, otro concepto básico en teoría de juegos. En el juego del ciempiés, el equilibrio de Nash supondría que el jugador A se quede con la ganancia acumulada desde el primer turno del juego, incluso si en este punto su valor es muy pequeño.

En muchas situaciones de interacción, el razonamiento requerido para llegar al equilibrio es demasiado complejo para ser plausible en términos conductuales, por lo que cuando personas reales se encuentran en dicho tipo de situaciones, sus elecciones suelen ser distintas de lo que supondría el equilibrio. Sin embargo, existe evidencia de que el aprendizaje (jugar un juego de forma repetida) genera una tendencia a converger al equilibrio y solo en situaciones donde no se permite que haya aprendizaje, se puede atribuir el equilibrio al pensamiento estratégico por sí sólo (Crawford, Costa-Gomes Iriberri, 2013). El equilibrio se ve reflejado tanto en las elecciones de los participantes del juego, como en sus creencias sobre los otros participantes: jugadores que son racionales (en términos de teoría de decisión) tienen creencias correctas sobre los otros jugadores, si estos también son racionales (Crawford, Costa-Gomes Iriberri, 2013).

Durante el proceso de razonamiento iterado, los jugadores incorporan las creencias que tienen sobre la conducta de los otros jugadores en su toma de decisiones. Keynes (1936) ilustró el proceso de razonamiento iterado con una analogía que se conoce como Beauty contest: Un concurso en el que los participantes deben elegir de entre cien fotografías de rostros, cuáles piensan que los demás participantes considerarán que son los más atractivos. Tomando en cuenta que todos los participantes se enfrentan al mismo problema, para ganar no basta con elegir solamente aquellos rostros que piensen que son los más atractivos, o cuáles piensan que los demás participantes piensan que son más atractivos, sino aquellos que piensen que los demás participantes pensarán que los demás participantes piensan que son los más atractivos. Esto implica tres pasos de razonamiento iterado.

Un agente totalmente racional debería realizar tantos pasos de razonamiento iterado como fueran necesarios para llegar a la solución por dominancia del juego (el equilibrio de Nash). En la realidad, las personas no se comportan de forma perfectamente racional, y la cantidad de pasos de razonamiento iterado que realizan es limitada (Stahl Wilson, 1995, Ho, Caremer Weigelt, 1998). Experimentalmente el juego p-Beauty contest, llamado así a partir de la analogía de Keynes, ha sido utilizado para estudiar el razonamiento iterado. En este juego, en el que participan varios participantes (por lo menos 3), cada uno debe elegir un número entero en el rango $[0 - 100]$, de manera simultánea y sin revelarlo a los otros jugadores. Posteriormente, se calcula la media de todos los números elegidos y este valor se multiplica por un parámetro p que es un número positivo y diferente de 1, conocido de antemano por todos los jugadores, (generalmente se utiliza $p = 2/3$). Al valor resultante de este cálculo se le llama el número objetivo, y el ganador del juego será el participante que haya elegido el número más cercano a este número.

Si los jugadores fueran perfectamente racionales, creyeran que los demás jugadores también lo son y además pudieran realizar una cantidad infinita de pasos de razonamiento iterado, llegarían a la solución por dominancia del juego mediante el siguiente razonamiento: El valor más grande que puede alcanzar el número objetivo (con $p = 2/3$) es $100 \cdot p = 66.66$, así que cualquier número arriba

de este valor es dominado por 66.66. Jugadores racionales obedecerán la dominancia y creerán que los demás jugadores lo harán también, por lo que todos elegirán un número menor a 66.66. Por lo tanto, el valor máximo del número objetivo será $100p^2 = 44.44$ y elegirá cualquier número por arriba de este ser una estrategia dominada. Aplicando

Empíricamente, esto no suele ocurrir, pero cuando el mismo grupo de participantes juegan repetidamente (más de un periodo), se ha reportado consistentemente que sus elecciones se acercan paulatinamente al equilibrio (elegir 0) con cada periodo (Nagel, 1995, Ho, Camerer Weigelt, 1998). También se ha observado que dicha tendencia se interrumpe cuando se agregan nuevos participantes al juego, siendo que los jugadores con experiencia incrementan el número elegido al enfrentarse a estos jugadores novatos, lo que se conoce como efecto de reset (Slonim, 2005).

Se han propuesto varios modelos que dan cuenta de la forma en la que las personas eligen sus números en el juego. Estos modelos capturan la noción de que la elección de las personas es un reflejo del número de pasos de razonamiento iterado que son capaces de realizar, que en la literatura de juegos se conoce como “nivel cognitivo”, así como de las creencias o expectativas que tienen sobre el nivel cognitivo de los demás jugadores (Crawford, Costa-Gomes Iriberrí, 2013).

Algunos estudios (Agranov et al., 2012 y Slonim, 2005) han explorado el efecto que tienen las creencias sobre el desempeño de los otros jugadores en las elecciones de cada participante, evaluando esta relación de forma indirecta y encontrando evidencia a favor de una relación positiva. En contraste, en otros estudios que han intentado tener un acercamiento más directo, se han encontrado inconsistencias entre las creencias sobre lo que harán los otros jugadores y las elecciones reales; Lahav (2015) utilizó un método en el que le pidió a los jugadores hacer una estimación indirecta de las elecciones de los otros jugadores y comparó directamente estas creencias con sus elecciones reales, encontrando una falta de consistencia entre las dos.

El presente trabajo de investigación pretende estudiar de manera directa la

relación entre las elecciones de las personas y sus creencias sobre las elecciones de los demás, y adicionalmente explora si la experiencia que se adquiere jugando el juego de forma repetida influye en la relación de estas dos variables.

El diseño experimental consiste en juegos repetidos de p-beauty contest (estructurados en 2 subjuegos compuestos por 4 periodos cada uno). En cada sesión experimental, un solo jugador participa en el juego durante los dos subjuegos, siendo que al término del Subjuego 1 los demás jugadores son reemplazados por nuevos participantes que no han jugado previamente. Además de registrar los números elegidos por cada participante en cada periodo, se les solicita a los jugadores que reporten directamente los números que creen que los demás elegirán. Este método de explicitar creencias permite comparar directamente las elecciones y las creencias de los jugadores, de forma más evidente a la utilizada por Lahav (2015), si bien igual que en ese estudio, acercarse a los números que eligen los otros jugadores es recompensado con ganancias en el juego para agregar motivación a la expresión de las creencias. Se encontró que la entrada de nuevos jugadores en el segundo subjuego incrementa el número elegido de los jugadores con experiencia en el primer periodo de este nuevo subjuego, pero la diferencia entre sus creencias y elecciones no incrementa junto con esta, lo que aporta evidencia de que la adquisición de experiencia incrementa la consistencia.

El resto de la tesis está dividida en cinco apartados: En el primero se presenta el marco teórico, que describe el modelo de nivel-k usado para explicar la conducta de las personas en el juego. La sección también revisa la relación empírica entre las elecciones de las personas y sus creencias, y la forma en que la experiencia en juegos repetidos influye en las elecciones de los jugadores. También se detallan los objetivos concretos del trabajo de investigación y las estrategias para alcanzarlos. En el segundo apartado se describe el método utilizado, incluyendo información sobre los participantes, el procedimiento y el diseño experimental. En el tercer apartado se presentan los resultados del experimento. Se reporta el grado de consistencia que existe entre creencias y elecciones de los jugadores en el primer subjuego, el efecto de introducir a participantes sin experiencia en el segundo subjuego, y las diferencias en consistencia entre creencias y elecciones

que hay entre los dos subjuegos. Por último, en el cuarto apartado se elabora la discusión a partir de los resultados, y las conclusiones se presentan en el quinto apartado.

Capítulo 2

Marco Teórico

2.1. Modelo de nivel-k

Este modelo fue propuesto por Nagel (1995) para dar cuenta de la conducta de las personas en juegos con solución por dominancia, como ocurre en p-beauty contest. El modelo define niveles cognitivos que describen el número de pasos de razonamiento iterado que realiza una persona en el juego y difiere respecto de los modelos de equilibrio clásicos en que las creencias que se tienen sobre las elecciones de los otros jugadores no están basadas en la definición del equilibrio (Crawford, Costa-Gomes Iriberry, 2013), es decir, se admite la posibilidad de que los otros jugadores no son perfectamente racionales. Según este modelo, los jugadores con un nivel cognitivo 0 serían aquellos que no realizan ningún paso de razonamiento iterado, es decir, que no toman en consideración que las elecciones de los otros participantes influyen en el cálculo del número objetivo. Estos jugadores eligen un número con base en alguna regla arbitraria, (por ejemplo, su número de la suerte o favorito), por lo que podrían elegir cualquier número dentro del rango establecido con una probabilidad similar.

Un jugador de nivel 1 es aquél que se sí considera que las elecciones de los otros jugadores influyen en el cálculo del número objetivo, pero supone que los otros jugadores no han tomado esto en consideración; el jugador de nivel 1 asume que los demás jugadores son de nivel 0 y elige el número que es la respuesta óptima contra este tipo de jugadores. El jugador asume que la media de números elegidos por todos los jugadores estará cerca de 50 (el mejor predictor de la media de un conjunto de números aleatorios en el rango [0 - 100]) y multiplicará este número por p para acercarse lo más posible al número objetivo.

Por su parte, un jugador de nivel 2 no solo considera que las elecciones de otros jugadores influyen en el número objetivo, sino que también asume que los otros jugadores saben esto; el jugador de nivel 2 elegirá el número que es la respuesta óptima contra una población de oponentes de nivel 1. Como estos eligen números cercanos a $50 \cdot p$, el jugador de nivel 2 debe multiplicar por p nuevamente para acercarse al que piensa que será el número objetivo, esto es $50p^2$.

En general, un jugador de nivel k elegirá la respuesta óptima contra una población de jugadores de nivel $k-1$, esto es $50 p^k$. Con base en esta regla, el modelo computa el nivel cognitivo de cada jugador en el juego. Este modelo ha sido utilizado por Camerer, Ho, y Chong (2004). El modelo propone que un jugador de nivel k elegirá la respuesta óptima contra una población de jugadores de nivel $k-1$, sino a una combinación de todos los tipos de jugadores desde el nivel 0 hasta $k-1$, a partir de una distribución Poisson.

En general, de acuerdo con la evidencia obtenida experimentalmente, los modelos de nivel- k explican mejor el comportamiento de las personas en juegos con solución por razonamiento iterado que los modelos de equilibrio (Crawford, Costa-Gomes Iriberry, 2013).

En ambos modelos la elección de los jugadores depende de tres elementos: 1) sus creencias sobre cómo juegan los participantes de nivel 0, 2) sus expectativas sobre el nivel cognitivo de los oponentes, y 3) el número de pasos de razonamiento que son capaces de hacer en el juego (Agranov et al., 2012). En la siguiente sección se ahonda sobre el segundo elemento: las expectativas (i. e. creencias) sobre el nivel cognitivo de los oponentes, y la evidencia que se ha encontrado sobre su relación con la elección.

2.2. Relación entre creencias y elecciones

Para aportar evidencia empírica de la influencia de las creencias acerca de la sofisticación de los otros jugadores sobre las elecciones de los jugadores en p-beauty contest, Agranov et al. (2012) manipularon las creencias que los participantes tenían sobre sus oponentes en un juego de p-beauty contest, informando a cada participante que jugaría contra 7 estudiantes graduados de economía

con conocimiento sobre este tipo de juegos, o bien, contra 7 computadoras programadas para seleccionar con la misma probabilidad números en el rango [0 – 100]. En este estudio se encontró que los números registrados por los participantes corresponden con un nivel cognitivo significativamente mayor en la condición en la que se enfrentaban a estudiantes graduados que en la condición de las computadoras. Este resultado parece sugerir que el nivel cognitivo que muestran las personas en juegos de p-beauty contest depende no únicamente de su sofisticación cognitiva, sino también de sus creencias sobre la sofisticación de los otros jugadores.

Para estudiar de forma mucho más directa la relación entre creencias y elecciones, Lahav (2015) utilizó un método de elicited beliefs (la mejor forma de traducirlo sería “hacer explícitas las creencias”) en sesiones experimentales compuestas por 5 períodos de p-beauty contest con hasta 20 participantes. En cada periodo, además de elegir su propio número, se les pidió a los participantes que estimaran cuántos de los otros participantes elegirían un número dentro de cada uno de 10 intervalos dentro del el rango [0-100] (0-10, 11-20, 21-30, ..., 91-100) y con dichas estimaciones, se calcularon las creencias de los participantes sobre el número promedio en cada periodo del juego. Al implementar este método, en contraste con investigaciones previas, Lahav concluye que las elecciones no son un reflejo preciso de las creencias de los participantes, pues encuentra diferencias significativas entre el número objetivo computado de acuerdo a las creencias de los participantes acerca de las tiradas de los demás jugadores y el número que de hecho eligen en el juego. Otra investigación con resultados similares fue realizada por Costa-Gomes y Weizsäcker en el 2008, utilizando un método para explicitar creencias en juegos sencillos de 3x3 (donde dos jugadores tienen que elegir entre tres estrategias posibles) en los que el razonamiento iterado permite llegar al equilibrio del juego. En dicha investigación, se encontró que en la mayoría de los casos, de acuerdo con el modelo de nivel-k, las creencias registradas por los jugadores acerca de sus oponentes los situaban en el nivel 2, mientras que sus elecciones correspondían al nivel 1. Ya que no parece que las elecciones observadas habrían sido una respuesta óptima ante las creencias explicitadas que se registraron, este resultado pone en duda que estas últimas sean la base a partir de la cual los participantes emiten sus respuestas. A la luz

de estos hallazgos, los autores concluyen que los jugadores basan sus decisiones en reglas de elección que influyen en ambas, creencias y decisiones.

Debido a que el presente trabajo de tesis incorpora parte del método de Lahav (2015) para recopilar las creencias de los jugadores acerca de las tiradas de sus oponentes en juegos de p-beauty contest, se enfatizan los siguientes puntos respecto a este estudio:

1) No se conoce con exactitud la creencia de los participantes sobre el número objetivo. Esta se calcula de manera aproximada, a partir del número de jugadores que se cree que elegirán un número en cada intervalo y tomando la media de cada intervalo como el valor más representativo de los mismos.

2) Con alrededor de 20 personas participando en el juego, parece inverosímil, dada la demanda cognitiva, que los jugadores puedan calcular con precisión el número objetivo derivado de sus creencias para emitir su respuesta.

3) Se contó con dos grupos control. En uno, no se realizó la explicitación de creencias, y en el otro, dicho procedimiento se realizó después de que los participantes eligieran su número. Comparado la elección de los participantes en los grupos control con el grupo experimental, los resultados indican que solicitar a los participantes que registraran sus creencias no cambia significativamente su número elegido.

4) El último periodo del juego fue el único en el que no se encontraron diferencias significativas entre creencias y elecciones, lo que podría sugerir que dicha discrepancia disminuye con la experiencia.

5) Dado que se sabe que los participantes tienden a elegir números cada vez más pequeños en cada periodo y que esto reduce invariablemente la magnitud de cualquier diferencia entre elecciones y creencias (ya que las diferencias entre números más pequeños son, por definición, más pequeñas), Lahav implementó un método de normalización con el que ponderó las diferencias entre las creencias y elecciones de cada participante en cada periodo por el promedio de los

números elegidos por todos los participantes en dicho periodo. Sin embargo, como la medida de normalización depende de la tirada de todos los jugadores, la magnitud de la diferencia normalizada es influida por el nivel cognitivo promedio, y “castiga” (incrementa la magnitud) las diferencias entre creencias y elecciones cuando estas no son tan sofisticadas como las del promedio. Debido a esto, podría no ser la mejor forma de compensar la tendencia al equilibrio. El resultado mencionado en el punto 4 permite cuestionar si la discrepancia entre creencias y elecciones se ve afectada por la experiencia que tienen los participantes en el juego. En la siguiente sección se revisa el efecto de la experiencia en juegos repetidos de p-beauty contest.

2.3. El efecto de la experiencia

Para estudiar el efecto de la experiencia en juegos repetidos de p-beauty contest, Slonim (2005) realizó sesiones experimentales de 12 periodos, distribuidos equitativamente en tres subjuegos, y con tres jugadores. En una primera condición, al terminar cada Subjuego se reemplazaba a dos de los tres participantes por jugadores nuevos, siendo que sólo un jugador permaneció en el experimento durante los 12 periodos completos. En una segunda condición, los dos participantes retirados al término de cada subjuego eran sustituidos por jugadores con la misma experiencia que el jugador que se mantenía en el juego (es decir, que habían jugado la misma cantidad de periodos). En ambas condiciones, los participantes tenían información acerca del nivel de experiencia de los demás jugadores (el número de periodos jugados).

Slonim (2005) reportó que en el primer periodo de los subjuegos 2 y 3, los jugadores con más experiencia mostraron un mayor nivel cognitivo (es decir, eligieron números más cercanos a 0) cuando sabían que jugaban contra oponentes que contaban con la misma experiencia que ellos, que cuando jugaban contra oponentes que no habían jugado previamente. Este resultado aporta evidencia a favor de que las creencias sobre el nivel cognitivo de los otros jugadores influyen en las elecciones. Por su parte, los jugadores sin experiencia no mostraron diferencias significativas en sus elecciones cuando jugaron con oponentes expe-

rimentados o no experimentados.

En cuanto al efecto de la experiencia en el desempeño de los jugadores, también se observó que los jugadores experimentados ganan el juego con mayor frecuencia cuando juegan con jugadores sin experiencia, ventaja que se reduce periodo a periodo, conforme los otros jugadores adquieren experiencia.

Un último resultado reportado por Slonim (2005), y probablemente el de mayor relevancia para efectos de la presente tesis, corresponde a un efecto de reset en la tendencia a ir reduciendo el número elegido en cada periodo, al iniciar un nuevo subjuego. Es decir, que los jugadores con experiencia presentan una reversión en la tendencia a elegir números cada vez más cercanos al equilibrio cuando nuevos jugadores entran en el juego.

Con base en estos hallazgos, y los reportados en las secciones anteriores, se procede a plantear formalmente el objetivo de este trabajo de investigación, así como las estrategias metodológicas empleadas para llevarlo a cabo.

2.4. Objetivo

El presente trabajo de investigación busca evaluar si la reducción en las diferencias entre las creencias y elecciones, reportada con un método de explicitación de creencias similar al propuesto por Lahav (2015), es efecto principalmente de la experiencia obtenida al participar repetidas veces en un juego.

Para responder a esta pregunta, se propone una versión modificada del método de Lahav (2015) que contempla la participación de un grupo más pequeño de jugadores, lo que facilita preguntar a los participantes directamente por las creencias sobre los números específicos que elegirán los demás participantes, permitiendo una estimación de las creencias más precisa y reduciendo la demanda cognitiva para los jugadores, de manera que resulta más verosímil esperar una correspondencia directa entre las elecciones de los participantes y el cómputo del número objetivo de acuerdo a sus creencias.

Para atenuar el peso que tiene la tendencia a elegir números cada vez más pequeños entre cada periodo, se busca promover el efecto de reset reportado por Slonim (2005) en la elección de los jugadores con experiencia, al tener jugadores que participarán durante más de un subjuego y sustituyendo al resto por nuevos jugadores en un p-beauty contest repetido. Se espera que el efecto de reset opere no sólo en las elecciones, sino también en las creencias del jugador experimentado, lo que permitiría evaluar si la diferencia entre estas sigue reduciéndose, como ocurre en lo reportado por Lahav (2015) en el último periodo registrado, donde parece ser que las elecciones y las creencias se vuelven “consistentes”. Si, mediante el efecto de reset, el jugador experimentado elige un número más grande en el primer periodo del segundo subjuego pero la diferencia entre su elección y su creencia en dicho periodo disminuye, esto aportaría evidencia a favor de que la experiencia reduce las inconsistencias entre creencias y elecciones.

En conjunto, el diseño experimental aquí propuesto permite determinar si las creencias de los jugadores que participan en los dos subjuegos se acercan más a las elecciones reales de los otros jugadores en el segundo subjuego gracias a que cuentan con mayor experiencia.

Capítulo 3

Método

3.1. Participantes

Se contó con la participación de 50 estudiantes de los primeros semestres de la carrera de Psicología en la Facultad de Psicología de la Universidad Nacional Autónoma de México. A cambio de su participación todos los estudiantes recibieron medio punto extra sobre la calificación de uno de sus exámenes, y los participantes que obtuvieron mayores ganancias al final del juego recibieron adicionalmente otro medio punto extra. Su participación en el experimento fue voluntario, podían abandonarlo si lo deseaban, y se mantuvo su anonimato entre ellos y para el análisis de datos, mediante el uso de una clave de identificación única asignada a cada participante.

3.2. Diseño Experimental

Se realizaron 10 sesiones experimentales con 5 participantes diferentes cada una, en un aula sin distracciones externas. Al inicio de cada sesión se asignó aleatoriamente un rol a cada participante (A, B, C, D y E).

Las sesiones estuvieron compuestas de ocho periodos de p-beauty contest, divididos en dos subjuegos de cuatro periodos. En cada subjuego participaron solo tres personas: En el primer subjuego, los participantes A, B y C jugaron por cuatro periodos, mientras los participantes D y E esperaban en un aula diferente sin poder hablar entre ellos. Terminado el subjuego 1, los participantes B y C

se retiraban y eran reemplazados para el subjuego 2 por los participantes D y E. En el segundo Subjuego, los participantes A, D y E jugaron por cuatro periodos, al final de los cuales se dio por terminado el experimento. Además de elegir su número en el juego, se solicitó a cada participante que en cada periodo escribieran dos números que creyeran que estarían lo más cerca posible de los números elegidos por los otros dos jugadores.

Se decidió utilizar únicamente dos subjuegos ya que los efectos de la experiencia reportados por Slonim (2005) no mostraron ser significativamente diferentes entre el subjuego 2 y 3.

Es importante mencionar que el subjuego 1 es una replicación parcial del experimento de Lahav (2015), dado que es un juego repetido con explicitación de creencias.

3.3. Procedimiento

En cada sesión se citó a los 5 participantes en un aula donde se les leyeron las instrucciones del experimento: Se les explicó cómo jugar p-beauty contest y se les dio la instrucción adicional de intentar adivinar el número a elegir por los otros jugadores y registrar dos creencias por periodo. También se les explicó que la sesión se dividiría en dos subjuegos con tres participantes en cada uno. Finalmente se hicieron de su conocimiento las ganancias que podrían acumular en caso de dar el número más cercano al número objetivo (6 puntos para el jugador ganador; en caso de empates, los 6 puntos se dividían equitativamente entre todos los ganadores), o bien, de acertar en la predicción de las tiradas de sus compañeros (1 punto por cada número que estuviera dentro de un rango de ± 5 de un número elegido por otro jugador). Se les pidió a los participantes que repitieran las reglas del juego para garantizar que habían entendido las indicaciones. Las instrucciones se anexan como Apéndice al final del presente documento.

Los participantes asignados mediante sorteo a los roles D y E esperaron en

un aula separada, en la que se cuidó que no hablaran entre ellos. Mientras tanto, a los otros tres participantes (roles A, B y C) se les entregaron cuatro formatos de respuesta para que en cada periodo escribieran su número elegido y sus creencias. Los formatos incluían una clave de identificación única para cada participante, el número de periodo y los espacios para que los participantes escribieran su número elegido y su estimación de los números a elegir por los otros dos participantes. Un ejemplo de estos formatos se puede consultar en el apéndice.

Los participantes A, B y C jugaron cuatro periodos de p-beauty contest, durante los cuales registraron tanto su propia elección como sus predicciones acerca de las elecciones de sus oponentes. Al terminar cada periodo, luego de que los participantes llenaran sus formatos, estos se recogían y se anotaban en el pizarrón del aula todos los números elegidos, sin indicar qué participante había elegido cada uno. Inmediatamente y de manera explícita, se calculaban el promedio de estos números y el número objetivo, y se anotaban en el pizarrón. Una vez hecho esto, se anunciaba qué jugador o jugadores habían elegido el número más cercano al número objetivo, y se les anotaba una ganancia de 6 puntos de juego distribuidos entre todos los ganadores. También se estableció qué jugadores habían estimado acertadamente los números elegidos por los otros jugadores con un margen de error de 5 números de diferencia y se les anotó una ganancia de 1 punto por cada acierto. Una vez realizado este proceso, se devolvía su formato de respuesta a cada candidato, incluyendo en el reverso de los mismos el total de puntos de juego obtenidos en ese periodo (que podía oscilar entre 0 y 8). Luego de que los participantes conocieran los puntos que obtuvieron en cada periodo, se borraba el pizarrón y se daba paso al siguiente periodo.

Al terminar los cuatro primeros periodos de juego (el subjuego 1), se contabilizaron los puntos obtenidos por cada participante y se anunciaba al participante ganador de la mayor cantidad de puntos. Si el jugador con más puntos era el participante B o C, se le anunciaba que obtenía otro medio punto extra adicional en su examen, además del medio punto que ya había obtenido por participar en el experimento. Si la persona con más puntos era el participante A, se le decía que para obtener el medio punto extra adicional debía ganar también en el siguiente subjuego.

Tras anunciar el ganador del Subjuego 1, se despedía a los participantes B y C, y los participantes D y E ingresaban al aula. Antes de comenzar con el Subjuego 2, se dio un breve recordatorio de las reglas del juego, haciendo hincapié en que el participante A contaba con más experiencia que los jugadores D y E, al haber jugado cuatro periodos.

El subjuego 2 se llevó a cabo de la misma forma que el subjuego 1. Al final de cuatro periodos se contabilizaron los puntos de juego adquiridos por cada participante, y el que obtuvo la mayor cantidad recibió el medio punto extra sobre su examen, además del medio punto extra otorgado por haber participado en el experimento.

Con el término del subjuego 2 y el anuncio del participante ganador, se despedía a los participantes y se daba por terminada la sesión.

Capítulo 4

Resultados

Para el análisis de los resultados obtenidos, se utilizaron tanto pruebas estadísticas bayesianas como pruebas clásicas o frecuentistas. En el presente apartado se presentan únicamente los resultados de los análisis bayesianos, mientras que los resultados de los análisis clásicos se incluyen como apéndice; en general, ambas aproximaciones apuntan hacia las mismas conclusiones.

4.1. Análisis Bayesiano

Se decidió presentar los resultados del análisis bayesiano porque se considera que el factor de Bayes, que funciona como una alternativa al p-value como medida de significancia estadística, refleja de forma mucho más clara la relación entre los datos obtenidos y las hipótesis nula y alternativa, al señalar cuántas veces es más probable que la evidencia haya sido producida bajo la hipótesis alterna respecto a la hipótesis nula (BF10), o viceversa (BF01).

Consistencia entre creencias y elecciones en el primer subjuego El primer subjuego del presente experimento presenta una replicación parcial del experimento de Lahav (2015). En ambos casos se juega p-beauty contest por varios periodos y se utiliza un método para explicitar creencias. Las diferencias radican en el número de jugadores que participan en un periodo (3 en el presente experimento, y hasta 20 en el de Lahav), la cantidad de periodos jugados (4 en el presente experimento y 5 en el de Lahav), y el método utilizado para explicitar las creencias (en el presente experimento se preguntó directamente por el número que se creía elegirían los otros jugadores, mientras que en el experimento

de Lahav se estimaba el número de jugadores que elegirían un número dentro de ciertos rangos).

En la Figura 1 se contrasta la elección de cada participante con su creencia promedio para cada uno de los periodos del subjuego 1, cuando se incluye la multiplicación por p y cuando esta se omite. Se observa que en ambos casos la diferencia entre creencias y elecciones se reduce periodo a periodo, pues en periodos posteriores los puntos se acercan más a la línea de identidad. Como ya se ha mencionado, esto puede deberse únicamente a que los valores elegidos se acercan más al límite inferior del espacio de elección.

Figura 1. Contraste entre elecciones y creencias promedio. Los puntos más cercanos a la línea de identidad indican un mayor grado de consistencia. En la columna de la izquierda se incluye la multiplicación por p , mientras que en la derecha se omite.

Para evaluar la consistencia entre las creencias de los participantes y sus números elegidos tomando en cuenta el efecto de suelo, se emplearon dos métodos: el primero de ellos, computa la diferencia normalizada entre las creencias y las elecciones de los participantes de acuerdo con las elecciones promedio observadas en cada periodo (Lahav, 2015); el segundo, calcula la diferencia relativa entre creencias y elecciones a partir del punto medio entre ambos valores (usado por Slonim, 2005, para calcular el cambio relativo de los números elegidos por los jugadores de un periodo al siguiente). A su vez, igual que Lahav, para evaluar el cómputo realizado por los participantes, se utilizaron estos métodos de dos formas, incluyendo u omitiendo la multiplicación por el parámetro p .

El procedimiento sugerido por Lahav (2015) para calcular las diferencias normalizadas entre las creencias y las elecciones de cada participante en cada periodo, fue incorporado a partir de la siguiente ecuación:

$$DN_i^t = ((2/3B_i^t - C_i^t))/C_t$$

Donde DN_i^t es la Diferencia Normalizada entre creencias y elecciones para cada participante en el periodo t .

Una vez computadas las diferencias por cada participante y periodo, se cal-

culó el promedio de las mismas para poder someterlas a un análisis estadístico que permitiera evaluar si estas fueron significativamente diferentes de 0. Para ello, se realizaron pruebas-T bayesianas de una sola muestra para cada periodo de juego. En la Tabla 1 se reporta el factor de Bayes obtenido en dicho análisis, que permite estimar qué tantas veces es más probable que la evidencia corresponda con la hipótesis alterna (hay diferencia entre creencias y elecciones) respecto a la hipótesis nula (no hay diferencia entre creencia y elecciones). Como puede verse, solamente se encontraron diferencias significativas entre creencias y elecciones en los primeros dos periodos del juego.

Tabla 1. Subjuego 1. Diferencias normalizadas. Bayesian One Sample T-Test
 BF error Periodo 1 19.300 1.823e-6 -0.366 Periodo 2 34.545 3.137e-4 -0.342 Perio-
 do 3 0.281 2.840e-5 -0.097 Periodo 4 0.652 0.015 -0.147

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias normalizadas promedio entre las creencias y las elecciones de los jugadores en cada periodo del primer subjuego.

En la Figura 2 se presentan las distribuciones prior y posterior computadas en cada periodo. Las distribuciones prior señalan la hipótesis nula (las diferencias estarán cercanas a 0) y la distribución posterior presenta el estimado realizado a la luz de los datos. La forma más sencilla de interpretar estas gráficas es como una razón de probabilidades: si la densidad de probabilidad es mayor en la distribución prior que en la distribución posterior para el punto que señala un tamaño del efecto 0 ($=0$), quiere decir que la evidencia favorece la hipótesis alterna, ya que que a la luz de la evidencia es “muy poco probable” (menos de lo que se esperaba de acuerdo a la distribución prior) que el tamaño de efecto tenga un valor cercano a 0.

Periodo 1 Periodo 2

Periodo 3 Periodo 4

Figura 2. Se presenta la comparación entre las distribuciones prior y las distribuciones posteriores computadas en las pruebas-T bayesianas de una sola muestra, por cada uno de los cuatro periodos contenidos en el primer subjuego. Estos resultados son consistentes con lo que reporta Lahav (2015): en los prime-

ros periodos no hay consistencia entre creencias y elecciones, pero esta parece adquirirse conforme avanzan los periodos. Así mismo, en todos los periodos se encontraron diferencias negativas entre creencias y elecciones, lo que indica que en promedio las creencias de los participantes estuvieron por debajo de sus elecciones reales.

En el estudio presentado por Lahav (2015), el cómputo de la diferencia entre las creencias y las elecciones se repitió pero omitiendo la multiplicación por p en las creencias, en un intento por evaluar si las inconsistencias halladas entre las creencias y las elecciones se debían a que los participantes no habían considerado dicha operación. El presente trabajo también incorporó dicha variación del análisis, que se llevó a cabo de acuerdo a la siguiente ecuación, en la que ya no se incluye la multiplicación por $2/3$ en B_i^t : $DN_i^t = ((B_i^t - C_i^t))/C_t$ Nuevamente, las diferencias promedio

Tbayesianas de una sola muestra. Este análisis arroja resultados inversos a los encontrados cuando se multiplican los primeros periodos no muestrales y las diferencias significativas y los últimos, s. Aunado a ello, las diferencias se
Test B Error Periodo 10,2070,010 - 0,049 Periodo 20,1960,013 - 0,012 Periodo 33,8113,017e -
60,355 Periodo 41,8614,032e - 60,280

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias normalizadas promedio entre las creencias y las elecciones de los jugadores en cada periodo del primer subjuego.

De acuerdo con el factor de Bayes, aunque la hipótesis alterna es más probable en los periodos 3 y 4 (es decir, parece ser que sí hay diferencias entre las creencias y las elecciones), la evidencia a favor de esta conclusión es relativamente pequeña, (particularmente en el periodo 4, donde podría considerarse anecdótica). En la Figura 3 se incluyen las distribuciones prior y posteriores de cada periodo.

Periodo 1 Periodo 2

Periodo 3 Periodo 4

Figura 3. Se presenta la comparación entre las distribuciones prior y las distribuciones posteriores computadas en las pruebas-T bayesianas de una sola muestra, de acuerdo con las diferencias normalizadas computadas por participante, en cada uno de los cuatro periodos contenidos en el primer subjuego, entre sus elecciones y creencias, asumiendo que estas no se multiplican por p .

Este resultado difiere considerablemente de los hallazgos reportados por Lahav

(2015), donde al excluir la multiplicación por $2/3$, las diferencias en los cuatro periodos se volvieron positivas, significativamente diferentes de 0, y en general, más grandes que cuando sí se consideraba la multiplicación por $2/3$. Estos resultados fueron interpretados como un indicador de que los participantes sí tomaron en cuenta multiplicación por p . En el análisis de los datos obtenidos en el presente estudio, cuando se incluye la multiplicación por p en el cálculo de las diferencias entre creencias y elecciones, se encuentra que éstas fueron significativas en los primeros periodo, en tanto que al excluir dicha multiplicación, las diferencias significativas se reportan sólo en los últimos periodos, aunque la evidencia parece ser débil. En el caso del presente estudio, estos resultados pueden apuntar a que los participantes comienzan el juego sin considerar la multiplicación por $2/3$, pero la incorporan en sus decisiones al avanzar los periodos (o por lo menos, aprenden que el número objetivo siempre está por debajo del número promedio). En general, el presente estudio presenta hallazgos similares a los reportados por Lahav (2015) al emplear el método propuesto para calcular las diferencias normalizadas entre las creencias y las elecciones de los participantes, aún a pesar de las diferencias en el método de explicitación de creencias: Existen discrepancias entre las creencias y las elecciones de los participantes en los primeros periodos de juego, pero no en los últimos.

La elección real de los participantes se encuentra entre su percepción de cuál será el número promedio y cuál será el número objetivo.

Siendo que la diferencia más importante entre lo hallado en el presente estudio y lo reportado por Lahav (2015), es que, en promedio, los participantes no parecen incorporar la multiplicación por $2/3$ en su elección, al menos en los primeros periodos. Además de replicar el método de diferencias normalizadas utilizado por Lahav (2015), las diferencias entre creencias y elecciones fueron evaluadas con un segundo método que no dependía de la elección promedio de todos los participantes en cada periodo para ponderarlas. La medida utilizada fue la diferencia relativa entre creencias y elecciones de cada participante i en cada periodo t , calculada de la siguiente manera: $DR_i^t = ((2/3B_i^t - C_i^t))/(0,5(2/3B_i^t + C_i^t))$

Las diferencias relativas computadas en promedio por cada periodo fueron evaluadas en términos de qué tanto se alejaron de 0, mediante la realización de pruebas-T bayesianas de una sola muestra, que se presentan en la Tabla 3.

Se encontró que en tres de los cuatro periodos hubo diferencias significativas entre creencias y elecciones. En la Figura 4 se incluyen las distribuciones prior y posterior de cada periodo.

Tabla 3. Subjuego 1. Diferencias relativas. Bayesian One Sample T-Test BF error Periodo 1 72.283 7.775e-5 -0.457 Periodo 2 12.797 2.047e-6 -0.328 Periodo 3 0.214 0.008 -0.052 Periodo 4 1.871 4.022e-6 -0.212

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias relativas promedio entre las creencias y las elecciones de los jugadores en cada periodo del primer subjuego.

Periodo 1 Periodo 2

Periodo 3 Periodo 4

Figura 4. Se presenta la comparación entre las distribuciones prior y las distribuciones posteriores computadas en las pruebas-T bayesianas de una sola muestra, de acuerdo con las diferencias relativas computadas en cada uno de los cuatro periodos contenidos en el primer subjuego, entre sus elecciones y creencias.

Tal y como se encontró con el método de Diferencias Normalizadas, todas las diferencias son negativas, indicando que consistentemente las creencias estuvieron por debajo de las elecciones reales. Sin embargo, contrario a lo que se esperaría con base en los resultados hallados con el método de Diferencias Normalizadas donde los participantes reducen la inconsistencia entre sus elecciones y las creencias explicitadas conforme adquieren experiencia, con el método de Diferencias Relativas se observan inconsistencias (diferencias estadísticamente significativas) entre creencias y elecciones en el periodo 4, aunque no en el periodo 3. De cualquier forma, de acuerdo con el factor de Bayes, la evidencia a favor de la hipótesis alterna en el periodo 4 es anecdótica. Posteriormente, se computaron las Diferencias Relativas omitiendo la multiplicación por p , de la siguiente forma: $DR_i^t = ((B_i^t - C_i^t)) / (0,5(B_i^t + C_i^t))$

Se realizaron pruebas-T de una sola muestra para determinar si las Diferencias Relativas promedio en cada periodo son significativamente diferentes de 0. En la Tabla 4 se presentan los resultados obtenidos por la prueba-T bayesiana. A su vez, la Figura 5 presenta la relación entre las distribuciones prior y las posteriores de cada periodo. Tabla 4. Subjuego 1. Diferencias relativas. Se omite la multiplicación por p . Bayesian One Sample T-Test BF error Periodo 1 0.269

5.246e-4 -0.101 Periodo 2 0.211 0.009 0.044 Periodo 3 8.393 2.307e-6 0.315 Perio-
do 4 0.810 5.669e-6 0.167

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias relativas promedio entre las creencias y las elecciones de los jugadores en cada periodo del primer subjuego.

Periodo 1 Periodo 2

Periodo 3 Periodo 4 Figura 5. Se presenta la comparación entre las distribuciones prior y las distribuciones posteriores computadas en las pruebas-T bayesianas de una sola muestra, de acuerdo con las diferencias relativas computadas en cada uno de los cuatro periodos contenidos en el primer subjuego, entre sus elecciones y creencias, sin tomar en cuenta la multiplicación por p .

Similar a lo observado cuando se omitió la multiplicación por p en el método de Diferencias Normalizadas propuesto por Lahav (2015), se encontró una reversión en la significancia reportada en todos los periodos, aunque nuevamente, la evidencia en el periodo 4 es anecdótica. En tres de los cuatro periodos se observaron diferencias positivas. Esto indica que con el uso del método de Diferencias Relativas, en promedio, las creencias están más cercanas y ligeramente por arriba de las elecciones reales de los participantes, cuando no se toma en cuenta la multiplicación por p . Aunque cada uno de los métodos empleados para la evaluación de la consistencia entre las creencias y las elecciones de los participantes compensa la tendencia hacia el equilibrio de forma diferente (y el problema de suelo resultante), los dos mostraron resultados muy similares: en los primeros periodos las diferencias entre creencias y elecciones son grandes, pero se reducen en los periodos posteriores, sugiriendo que los jugadores se vuelven consistentes conforme adquieren experiencia y aprenden también que para acercarse al número objetivo necesitan elegir números por debajo del número promedio. Efecto de reset De acuerdo con los resultados reportados por Slonim (2005), los jugadores presentan un efecto de “reset” en la tendencia a elegir números cada vez más pequeños cuando los otros jugadores son reemplazados por nuevos jugadores que no tienen experiencia en el juego. Tomando estos hallazgos en cuenta, el presente estudio incorporó un Subjuego 2, donde sólo uno de los jugadores del Subjuego 1 permaneció jugando por otros cuatro periodos mientras el resto fue reemplazado por jugadores nuevos. Esta manipulación se hizo para evaluar la tendencia que en el estudio de Lahav (2015) lleva a

asumir que los jugadores se vuelven más consistentes conforme adquieren experiencia. En otras palabras, agregar un segundo Subjuego permite evaluar, con base en las respuestas del participante con experiencia, si la consistencia entre las elecciones y las creencias es algo que se adquiere con la experiencia o si es sólo el resultado del efecto de suelo asociado a la tendencia típicamente reportada en cualquier serie de juegos p-beauty contest repetido a elegir números cada vez más pequeños. Para poder comparar los resultados del subjuego 1 con el subjuego 2, fue necesario determinar si la incorporación de nuevos jugadores al inicio del subjuego 2 interrumpió la tendencia a elegir números cada vez más pequeños en el jugador que se mantuvo en el juego (participante A). Es decir, para corroborar que el diseño experimental propuesto permite responder a la cuestión de si las diferencias entre creencias y elecciones se reducen como reflejo de una consistencia adquirida o como producto del efecto de suelo, es necesario evaluar la presencia del efecto de Reset reportado por Slonim (2005). En la Figura 6 se muestran los cambios en las elecciones de los participantes A entre periodos consecutivos. Los primeros tres cuadros muestran los cambios dentro del subjuego 1. El cuarto cuadro evalúa directamente el Efecto de Reset al presentar el cambio entre el último periodo del subjuego 1 y el primer periodo del subjuego 2. En estos cuadros, los puntos que caen por debajo de la línea de identidad indican que se eligieron números más pequeños de un periodo a otro, con lo que se observa que más participantes A eligen números más pequeños entre los periodos del primer subjuego (por ejemplo, el 80

Periodo 1 a Periodo 2 Periodo 2 a Periodo 3

Periodo 3 a Periodo 4 Periodo 4 a Periodo 5 (subjuego 2)

Figura 6. Cambio en el número elegido periodo a periodo.

De acuerdo con una prueba binomial bayesiana de una cola, los datos presentan evidencia anecdótica en favor de la hipótesis alterna que establece que los participantes eligen números más pequeños entre cada periodo del subjuego 1 y números más grandes al iniciar el subjuego 2 (ver Tabla 5). La falta de robustez en la evidencia puede deberse a que solamente se realizaron 10 sesiones experimentales. Tabla 9. Cambio en el número elegido periodo a periodo. Periodo 5 pertenece al subjuego 2. Bayesian Binomial Test Level Counts Total Proportion BF 1 vs 2 mayor 3 10 0.300 0.176 menor 7 10 0.700 1.376 2 vs 3 mayor 4 10 0.400 0.243 menor 6 10 0.600 0.643 3 vs 4 mayor 2 10 0.200 0.135 menor 8 10

0.800 4.002 4 vs 5 mayor 7 10 0.700 1.376 menor 3 10 0.300 0.176

Prueba binomial bayesiana de los casos en los que los jugadores eligen un número más grande que en el periodo anterior. Para determinar si, en promedio, el número elegido por el participante A en el último periodo del subjuego 1 es menor que su número elegido en el primer periodo del subjuego 2, se realizó una prueba-T de una cola, donde se encontró que si bien el promedio de los números elegidos en el primer periodo del subjuego 2 es mayor que en el subjuego 1, la diferencia parece ser pequeña y no significativa. No obstante, al revisar los datos de cada sesión experimental se encontraron anomalías en la ejecución del participante A de la sesión 3, cuyo comportamiento difiere considerablemente de lo que se reporta en la literatura y de lo observado en el resto de las sesiones. En la Figura 7 se presentan las elecciones de todos los participantes en cada periodo de los dos subjuegos, por cada sesión experimental. Se puede apreciar que el participante A de la sesión 3 no presenta la tendencia a elegir números más pequeños periodo a periodo y por el contrario, elige números muy grandes en el último periodo de cada subjuego, una estrategia que no aporta ningún tipo de ventaja en el juego. Este patrón de respuesta sugiere que el participante A de la sesión 3 pudo no haber entendido la dinámica del juego.

Figura 7. Números elegidos por cada uno de los cinco participantes incluidos en cada sesión experimental, a lo largo de los cuatro periodos que conformaron cada subjuego. El participante A, que jugó tanto en el Subjuego 1 como en el Subjuego 2 contra participantes sin experiencia, se señala en rojo. Para evaluar qué tan saliente es la tirada del jugador A de la sesión 3 en el primer periodo del segundo subjuego, en la Figura 8 se presentan dos diagramas. El primero, de caja y bigotes, muestra que la elección realizada por el jugador A de la sesión 3 en el primer periodo se encuentra por arriba del rango intercuadrático multiplicado por 1.5, lo que permite clasificarla como un valor atípico. El segundo diagrama, de violín, complementa la información del diagrama de caja y bigotes con un diagrama de densidad a los lados; la forma externa del diagrama representa todos los resultados posibles, y el grosor indica que tan comunes son.

Figura 8. Periodo 4. Número elegido de los participantes A.

Tomando en cuenta dicho resultado, se repitieron las pruebas-T de una cola para determinar si el número elegido por el participante A en el último periodo del subjuego 1 es menor que el número elegido en el primer periodo del

subjuego 2, omitiendo las tiradas del participante A de la sesión 3. Al realizar dicho análisis, el número elegido en promedio por los participantes A en el último periodo del subjuego 1 fue significativamente menor que la elección promedio registrada en el primer periodo del subjuego 2, encontrando evidencia del efecto de reset (prueba clásica $t = -2.317$, $p = .025 < 0.5$, prueba bayesiana $BF_{10} = 3.57$, error = $1.377e-4$). En la Figura 9 se presentan las elecciones promedio de los participantes A y de los participantes no-A (participantes B al E) en cada periodo jugado. La figura presenta en gráficos diferentes las elecciones promedio computadas cuando se consideran los datos de la sesión 3 (panel izquierdo) y cuando no (panel derecho).

Con la sesión 3 Sin la sesión 3

Figura 9. Número elegido promedio. Con y sin la sesión 3.

Con base en estos resultados, se puede afirmar que se logró replicar el efecto de reset reportado por Slonim (2005). Con ello, se comprueba la validez en la comparación de la consistencia entre las creencias y elecciones de los participantes A en cada periodo, a lo largo de los dos subjuegos. Consistencia entre creencias y elecciones entre subjuegos Las creencias y elecciones de los participantes en el subjuego 2 fueron sometidos a los mismos análisis reportados para el subjuego 1 en la sección 3.1: se computaron las diferencias normalizadas y las diferencias relativas, incluyendo y excluyendo la multiplicación por p . Sin embargo, al analizar las respuestas registradas en el subjuego 2 se hizo una distinción entre la ejecución de los jugadores con experiencia (participantes A) y los jugadores que participaban en el juego por primera vez (participantes D y E). En la figura 10 se contrasta la elección de cada participante con su creencia promedio para cada uno de los periodos del subjuego 2, cuando se incluye la multiplicación por p y cuando esta se omite. Los participantes A están marcados en rojo. En general, se observa una menor inconsistencia entre las creencias y elecciones de estos participantes comparados con los D y E.

Figura 10. Contraste entre elecciones y creencias promedio. Puntos en rojo señalan al participante A. Los puntos más cercanos a la línea de identidad indican un mayor grado de consistencia. En la columna de la izquierda se incluye la multiplicación por p , mientras que en la derecha se omite. Se calculó la diferencia normalizada entre las creencias y elecciones en cada periodo del subjuego 2, para el participante A y los participantes D y E. Posteriormente, se realizaron

pruebas-T bayesianas para determinar si estas diferencias fueron estadísticamente diferentes de 0. En la Tabla 6 y 7 se presentan los resultados obtenidos en las pruebas-T bayesianas de una sola muestra que evalúan el promedio de las diferencias normalizadas entre creencias y elecciones contra 0. De acuerdo con la Tabla 6, los participantes A no mostraron diferencias significativas (el factor de Bayes en este caso indica qué tantas veces es más probable la hipótesis nula de que no haya diferencias), lo que indica que hubo una mayor consistencia entre sus creencias y elecciones en todos los periodos del subjuego 2. En contraste, en la Tabla 7 se puede observar que los participantes D y E mostraron diferencias significativas en los primeros tres periodos del subjuego, y sólo parecieron adquirir consistencia entre sus creencias y elecciones en el último periodo. Este último resultado es muy similar a lo reportado en todos los participantes en el subjuego 1, donde todos los jugadores compartían el mismo nivel de experiencia. Además, en todos los casos, las creencias estuvieron por debajo de las elecciones reales. En las Figura 11 y 12 se incluyen las distribuciones prior y posterior de las diferencias normalizadas por periodo para los participantes A y los participantes D y E, respectivamente. Tabla 6 . Subjuego 2. Diferencias normalizadas. Participante A. Bayesian One Sample T-Test BF error Periodo 5 3.010 0.006 -0.049 Periodo 6 3.212 0.007 -0.011 Periodo 7 2.012 0.003 -0.133 Periodo 8 3.205 0.007 -0.028

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias normalizadas promedio entre las creencias y las elecciones del participante A en cada periodo del segundo subjuego.

Tabla 7. Subjuego 2. Diferencias normalizadas. Participantes D y E. Bayesian One Sample T-Test BF error Periodo 5 0.315 0.003 -0.349 Periodo 6 0.102 6.001e-4 -0.243 Periodo 7 0.063 3.519e-4 -0.307 Periodo 8 4.002 0.022 -0.047

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias normalizadas promedio entre las creencias y las elecciones de los participantes D y E en cada periodo del segundo subjuego.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 11. Se presenta la relación entre las densidades de las distribuciones prior y posteriores en el punto de “no tamaño del efecto” computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias normalizadas de los participantes A en cada periodo

contenidos en el subjuego 2.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 12. Se presenta la relación entre las densidades de las distribuciones prior y posteriores en el punto de “no tamaño del efecto” computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias normalizadas de los participantes D y E en cada periodo contenidos en el subjuego 2. Se repitió el cálculo de las diferencias normalizadas en los cuatro periodos del subjuego 2 omitiendo la multiplicación por p , con el propósito de evaluar si los jugadores tomaron en cuenta este cálculo para la elección de su número. Tras la realización de las pruebas-T bayesianas de una sola muestra, sólo se encontraron diferencias significativas entre las elecciones y las creencias de los participantes A en los dos primeros periodos, (ver Tabla 8). En general, la magnitud de las diferencias parece ser mayor en todos los periodos cuando se excluye la multiplicación por p que cuando esta sí es incluida. Este resultado sugiere que los jugadores con experiencia previa sí tomaron en cuenta la multiplicación por p , o por lo menos aprendieron desde el primer subjuego que el número objetivo siempre está por debajo del número promedio. Por otra parte, en el caso de los participantes sin experiencia (D y E) solo se encontraron diferencias significativas respecto de 0 en el último periodo (ver Tabla 9), resultado que coincide con lo reportado en el subjuego 1, sugiriendo nuevamente que los participantes incorporaron la multiplicación por p (o comprendieron la tendencia del juego hacia el equilibrio) solo después de varias repeticiones. Así mismo, en las Figuras 13 y 14 se presentan las distribuciones prior y posterior computadas en las pruebas-T realizadas por cada periodo para los participantes con y sin experiencia. Tabla 8. Subjuego 2. Diferencias normalizadas. Se omite la multiplicación por p . Participante A. Bayesian One Sample T-Test BF error Periodo 5 0.537 0.001 0.333 Periodo 6 0.046 9.917e-5 0.414 Periodo 7 1.274 0.004 0.238 Periodo 8 0.643 0.003 0.438

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias normalizadas promedio entre las creencias y las elecciones del participante A en cada periodo del segundo subjuego.

Tabla 9. Subjuego 2. Diferencias normalizadas. Se omite la multiplicación por p . Participantes D y E. Bayesian One Sample T-Test BF error Periodo 5 4.265 0.022 0.024 Periodo 6 1.417 0.004 0.170 Periodo 7 3.541 0.021 0.071 Periodo 8 0.115

7.221e-4 0.440

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias normalizadas promedio entre las creencias y las elecciones de los participantes D y E en cada periodo del segundo subjuego.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 13. Se presenta la razón entre las densidades de las distribuciones prior y posteriores computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias normalizadas de los participantes A en cada periodo contenidos en el subjuego 2, sin tomar en cuenta la multiplicación por p.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 14. Se presenta la razón entre las densidades de las distribuciones prior y posteriores computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias normalizadas de los participantes D y E en cada periodo contenidos en el subjuego 2, sin tomar en cuenta la multiplicación por p. Posteriormente, se computaron las diferencias relativas entre creencias y elecciones en los cuatro periodos del subjuego 2 y se realizaron pruebas-T bayesianas para determinar si estas eran significativamente diferentes de 0. Los resultados de este análisis se presentan en la Tabla 10 para el participante A y en la Tabla 11 para los participantes D y E. Tabla 10. Subjuego 2. Diferencias relativas. Participante A. Bayesian One Sample T-Test BF error Periodo 5 3.085 0.006 -0.042 Periodo 6 3.231 0.007 -0.007 Periodo 7 1.832 0.002 -0.162 Periodo 8 3.165 0.007 0.038

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias relativas promedio entre las creencias y las elecciones del participante A en cada periodo del segundo subjuego.

Tabla 11. Subjuego 2. Diferencias relativas. Participantes D y E. Bayesian One Sample T-Test BF error Periodo 5 0.272 0.002 -0.407 Periodo 6 0.044 3.044e-4 -0.283 Periodo 7 0.080 4.394e-4 -0.341 Periodo 8 4.295 0.022 -0.007

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias relativas promedio entre las creencias y las elecciones de los participantes D y E en cada periodo del segundo subjuego.

Los resultados que se obtienen son muy similares a lo que se observa al utilizar el método de diferencias normalizadas: el participante A muestra ser con-

sistente en los cuatro periodos y los participantes D y E son inconsistencias en los primeros tres. En las Figuras 15 y 16 se incluyen las distribuciones prior y posterior de cada prueba-T bayesiana realizada por periodo, por cada tipo de participante.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 15. Se presenta la razón entre las densidades de las distribuciones prior y posteriores computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias relativas de los participantes A en cada periodo contenidos en el subjuego 2.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 16. Se presenta la razón entre las densidades de las distribuciones prior y posteriores computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias relativas de los participantes D y E en cada periodo contenidos en el subjuego 2. Finalmente, se repitió el cálculo de las diferencias relativas entre las creencias y las elecciones de los participantes, omitiendo la multiplicación por p y se realizaron pruebas-T bayesianas para comparar las diferencias computadas en cada periodo contra 0. En la Tabla 12 se presentan los resultados de las pruebas realizadas para evaluar las diferencias calculadas para los participante A en cada periodo, y en la Tabla 13 para los participantes D y E. De acuerdo a estos análisis, se observó en los participantes A una reversión de las significancias estadísticas reportadas en todos los periodos y diferencias en promedio más grandes, lo que sugiere que las creencias de estos jugadores son más consistentes con sus elecciones cuando se asume que tomaron en cuenta la multiplicación por p . Por su parte, los participantes D y E también presentan una reversión en la significancia de las pruebas realizadas por periodo, pero en su caso, las diferencias promedio parecen ser más pequeñas cuando sí se considera la multiplicación por p , excepto en el último periodo. Este último hallazgo es consistente con la idea sugerida por lo reportado al evaluar las diferencias considerando la multiplicación por p : los participantes aprenden a multiplicar por p en los últimos periodos del subjuego. En las Figuras 17 y 18 se presentan las distribuciones prior y posterior computadas en cada prueba-T realizada por periodo para los participantes A y para los participantes D y E, respectivamente.

Tabla 12. Subjuego 2. Diferencias relativas. Se omite la multiplicación por p .

Participante A. Bayesian One Sample T-Test BF error Periodo 5 3.883 4.510e-4 0.346 Periodo 6 25.537 1.528e-4 0.386 Periodo 7 0.896 0.008 0.225 Periodo 8 2.753 1.237e-4 0.413

Prueba-T bayesiana de una sola muestra que compara contra 0 las diferencias relativas promedio entre las creencias y las elecciones del participante A en cada periodo del segundo subjuego.

Tabla 13. Subjuego 2. Diferencias relativas. Se omite la multiplicación por p. Participantes D y E. Bayesian One Sample T-Test BF error Periodo 5 4.076 0.022 -0.056 Periodo 6 1.932 0.010 0.108 Periodo 7 4.066 0.022 0.039 Periodo 8 0.048 3.096e-4 0.372

Prueba-T de una sola muestra que compara contra 0 las diferencias relativas promedio entre las creencias y las elecciones de los participantes D y E en cada periodo del segundo subjuego.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 17. Se presenta la razón entre las densidades de las distribuciones prior y posteriores computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias relativas de los participantes A en cada periodo contenidos en el subjuego 2, cuando se asume que estos no multiplican sus creencias por p.

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 18. Se presenta la razón entre las densidades de las distribuciones prior y posteriores computadas con las pruebas-T bayesianas de una sola muestra realizadas para evaluar las diferencias relativas de los participantes D y E en cada periodo contenidos en el subjuego 2, al asumir que estos no multiplican sus creencias por p.

En general, los resultados obtenidos en términos de la evaluación de la consistencia con que los jugadores con experiencia responden en el segundo subjuego, (en comparación con los otros jugadores sin experiencia), apuntan a que la reducción reportada por Lahav (2015) en las diferencias entre las elecciones y las creencias de los participantes a lo largo una serie de periodos de p-beauty contest, ocurre como producto de la experiencia adquirida por los participantes y no como co-producto del efecto de suelo asociado a la tendencia identificada en este tipo de juegos a elegir números cada vez más pequeños.

¿Las creencias se vuelven más precisas con la experiencia? Finalmente, dado

el efecto que la experiencia demostró tener sobre la ejecución de los participantes en juegos repetidos de p-beauty contest, tanto en la elección de sus números de acuerdo a la experiencia de otros jugadores, como a la adquisición de consistencia entre sus creencias y elecciones, se valoró una última pregunta de investigación derivada del tipo de datos obtenidos con el diseño propuesto. Dicha pregunta estuvo orientada a evaluar el efecto que pudo haber tenido la experiencia sobre la precisión en la predicción de las elecciones que los jugadores experimentados hacen sobre las elecciones de los otros jugadores. En la Figura 19 se presenta la comparación de las creencias de cada jugador en cada periodo contra los números que de hecho eligieron los otros jugadores. Se observa una alta falta de precisión en los primeros periodos de cada subjuego (periodos 1 y 5). La precisión del participante A (marcado en rojo) parece ser similar a la de los otros participantes.

Figura 19. Contraste entre elecciones promedio de los otros jugadores y las creencias promedio de cada jugador. Puntos en rojo señalan al participante A. Los puntos más cercanos a la línea de identidad indican un mayor grado de consistencia. Las líneas punteadas indican el rango de ± 5 en el que se recompensaban las creencias cercanas a las elecciones. En la columna de la izquierda se incluye la multiplicación por p, mientras que en la derecha se omite. Se realizó un último conjunto de pruebas-T de dos muestras para comparar el número de veces que los participantes en los subjuegos 1 y 2 lograron acercarse (dentro de un margen de error de ± 5) a la predicción de los números a elegir por sus oponentes en cada periodo. Para evaluar el efecto de la experiencia sobre la precisión de las predicciones hechas acerca de las elecciones de sus oponentes, se comparó el número de predicciones acertadas hechas por los participantes A en el Subjuego 1 contra el número de aciertos que obtuvieron una vez adquiriendo experiencia, en el subjuego 2 y se encontró que hubo un aumento en la cantidad de aciertos obtenidos en el segundo subjuego (41.25). En el caso del resto de los participantes tampoco se observaron diferencias significativas en su capacidad para predecir las elecciones de los otros jugadores. Los participantes B y C que participaron en el subjuego 1 obtuvieron el 33.125. Se consideró la posibilidad de que aunque la proporción de aciertos no aumentara significativamente entre subjuegos, las diferencias entre las creencias sobre los otros jugadores y sus elecciones sí disminuyeran sistemáticamente periodo a periodo, por lo que

estas se computaron, haciendo la distinción entre el participante A y los demás participantes. Mediante el uso de pruebas-T bayesianas de muestras independientes, se comparó el desempeño del participante A y los demás participantes para acercarse con sus creencias a la elección real de los otros jugadores. En la Tabla 14 se reportan los resultados de este análisis. La evidencia indica que no hay diferencias en el desempeño del participante A con respecto a los demás, en ningún periodo del juego, y las diferencias entre creencias y elecciones de los otros jugadores no cambian de forma sistemática entre periodos. En las Figura 20 se presentan las distribuciones prior y posterior computadas en cada prueba-T realizada por periodo. Tabla 14. Diferencias entre creencias sobre la tirada de los otros jugadores y la elección real de estos. Bayesian Independent Samples T-Test BF error Periodo 1 0.468 1.064e-5 -3.500 -9.400 Periodo 2 0.408 1.065e-5 -1.900 -0.150 Periodo 3 0.428 6.613e-5 4.350 7.525 Periodo 4 0.526 0.001 -2.400 4.050 Periodo 5 0.424 5.433e-5 -1.800 2.475 Periodo 6 0.403 3.276e-6 5.700 6.525 Periodo 7 0.434 7.814e-5 -1.950 1.000 Periodo 8 0.409 1.241e-5 6.150 5.025

Prueba-T bayesiana de muestras independientes que compara la diferencia entre creencias y elecciones reales de los otros jugadores del participante A y los demás participantes en todos los periodos.

Periodo 1 Periodo 2

Periodo 3 Periodo 4

Periodo 5 Periodo 6

Periodo 7 Periodo 8 Figura 20. Se presenta la razón entre las densidades de las distribuciones prior y posteriores computadas con las pruebas-T bayesianas de muestras independientes realizadas para evaluar las diferencias entre creencias y elecciones reales de los otros jugadores del participante A y los demás participantes.

Con base en los análisis realizados para la evaluación de los posibles cambios en la precisión de las predicciones hechas por cada jugador acerca de las elecciones de sus oponentes, se concluye que la experiencia no parece tener un efecto significativo sobre la habilidad de los participantes de anticipar las tiradas de sus contrincantes. Este resultado hace sentido con la manipulación experimental propuesta en el presente estudio: los participantes D y E que juegan con los participantes A en el segundo subjuego, son independientes de los participantes con los que jugaba en el primero (B y C), y no habría razón para

esperar que sigan las mismas estrategias en la elección de sus tiradas. *Discusión*

El experimento realizado aporta evidencia acerca de la relación que existe entre la elección de un número elegido en el juego de p-beauty contest y el cómputo de las creencias que se tienen sobre las elecciones de los demás jugadores, mediante la incorporación de una variación del método para explicitar creencias propuesto por Lahav (2015) en un procedimiento que aprovecha los efectos de Reset reportados por Slonim (2005) para evaluar el efecto que tiene la experiencia sobre la consistencia con que las elecciones de los jugadores reflejan sus creencias sobre los demás participantes en el juego. Para evaluar la diferencia entre las elecciones de los jugadores y sus creencias sobre las tiradas de sus oponentes, se utilizaron dos métodos diferentes. Primero, se tomó en cuenta la diferencia normalizada entre las creencias y las elecciones al ponderar esta por el promedio de los números elegidos por todos los jugadores en cada periodo; y después, se computó la diferencia relativa que toma como factor de ponderación el valor intermedio entre las creencias y las elecciones. Ambos métodos buscan compensar la tendencia que presentan las elecciones de los jugadores a converger en un equilibrio cercano a 0 cuando el juego se repite a lo largo de varios periodos. La diferencia sustancial entre ambos, es que la diferencia normalizada depende de la elección promedio registrada por todos los jugadores en el periodo a evaluar y la diferencia relativa se calcula únicamente a partir de la creencia y elección del jugador en cuestión. Si bien estas dos métodos llevaron al cálculo de distintos valores por cada jugador en cada periodo, la relación entre estos se presentó de la misma forma: Los jugadores presentan inconsistencias entre sus creencias y sus elecciones cuando no tienen experiencia, tal y como se observó en los primeros periodos jugados por los participantes sin experiencia en el subjuego 1 (participantes A, B y C) y en el subjuego 2 (participantes D y E). Dichas inconsistencias se reducen conforme los participantes adquieren experiencia, hacia el final del primer subjuego, y se mantienen a lo largo de los cuatro periodos que conforman el segundo subjuego para los participantes A, que juegan con los participantes sin experiencia D y E. En promedio, las elecciones reales de los jugadores se situaron por encima del número objetivo hipotético, asociado con las creencias registradas en cada periodo, y en cambio, se mantuvieron por debajo del promedio de sus creencias. Para evaluar la posibilidad de que las inconsistencias observadas se debieran a que los jugadores no estuvie-

ran tomando en cuenta que el promedio de sus creencias debía multiplicarse por p , al momento de elegir su número, se incluyeron variaciones en el cálculo de las diferencias normalizadas y relativas que omitían la multiplicación por p . Con ello se observó que la elección de los participantes era más consistente con el promedio de sus creencias (sin incluir la multiplicación por p) en los primeros periodos, pero conforme adquirieron experiencia entre periodos, sus elecciones se fueron acercando más a la del número objetivo estimado de acuerdo a sus creencias (en los últimos periodos del primer subjuego jugado por cada participante y durante todo el segundo subjuego, en el caso de los participante A). Este resultado indica que los participantes aprenden a incluir la multiplicación por p conforme adquieren experiencia en el juego. De cualquier forma, no es posible determinar si los participantes incorporan el cálculo explícitamente, o simplemente aprenden de forma intuitiva a elegir números cada vez más pequeños, por debajo del promedio de sus creencias. El inicio del segundo subjuego estuvo marcado por la introducción de dos nuevos jugadores (D y E) que reemplazaron a dos de los jugadores participantes en el subjuego 1 (B y E), siendo que uno de los jugadores de dicho subjuego permaneció durante cuatro periodos más. Con esta manipulación experimental se replicó exitosamente el efecto de Reset reportado por Slonim (2005), permitiendo evaluar la consistencia entre las elecciones y las creencias de los participantes A como una función de su experiencia, sin la influencia del efecto de suelo. Los resultados obtenidos a este respecto confirman la importancia que tiene la experiencia de los participantes sobre su desempeño en el juego de p -beauty contest, en términos de la consistencia entre los números elegidos y los números que se estimaba que tirarían los demás jugadores. El experimento que se realizó para conducir este estudio se llevó a cabo en 10 sesiones experimentales compuestas de 8 periodos de juego. En cada periodo 3 jugadores elegían números e intentaban adivinar los números de los otros jugadores. Considerando la cantidad de datos generados con este diseño, así como teniendo en cuenta diseños utilizados en la literatura (Kocher et al., 2007, Slonim, 2005), se considera que 10 sesiones fueron suficientes para obtener resultados válidos, pero que replicar el experimento con una muestra más grande podría incrementar la robustez de los hallazgos reportados.

Capítulo 5

Discusión

El experimento realizado aporta evidencia acerca de la relación que existe entre la elección de un número elegido en el juego de p-beauty contest y el cómputo de las creencias que se tienen sobre las elecciones de los demás jugadores, mediante la incorporación de una variación del método para explicitar creencias propuesto por Lahav (2015) en un procedimiento que aprovecha los efectos de Reset reportados por Slonim (2005) para evaluar el efecto que tiene la experiencia sobre la consistencia con que las elecciones de los jugadores reflejan sus creencias sobre los demás participantes en el juego. Para evaluar la diferencia entre las elecciones de los jugadores y sus creencias sobre las tiradas de sus oponentes, se utilizaron dos métodos diferentes. Primero, se tomó en cuenta la diferencia normalizada entre las creencias y las elecciones al ponderar esta por el promedio de los números elegidos por todos los jugadores en cada periodo; y después, se computó la diferencia relativa que toma como factor de ponderación el valor intermedio entre las creencias y las elecciones. Ambos métodos buscan compensar la tendencia que presentan las elecciones de los jugadores a converger en un equilibrio cercano a 0 cuando el juego se repite a lo largo de varios periodos. La diferencia sustancial entre ambos, es que la diferencia normalizada depende de la elección promedio registrada por todos los jugadores en el periodo a evaluar y la diferencia relativa se calcula únicamente a partir de la creencia y elección del jugador en cuestión.

Si bien estas dos métodos llevaron al cálculo de distintos valores por cada jugador en cada periodo, la relación entre estos se presentó de la misma forma: Los jugadores presentan inconsistencias entre sus creencias y sus elecciones cuando no tienen experiencia, tal y como se observó en los primeros periodos jugados

por los participantes sin experiencia en el subjuego 1 (participantes A, B y C) y en el subjuego 2 (participantes D y E). Dichas inconsistencias se reducen conforme los participantes adquieren experiencia, hacia el final del primer subjuego, y se mantienen a lo largo de los cuatro periodos que conforman el segundo subjuego para los participantes A, que juegan con los participantes sin experiencia D y E.

En promedio, las elecciones reales de los jugadores se situaron por encima del número objetivo hipotético, asociado con las creencias registradas en cada periodo, y en cambio, se mantuvieron por debajo del promedio de sus creencias. Para evaluar la posibilidad de que las inconsistencias observadas se debieran a que los jugadores no estuvieran tomando en cuenta que el promedio de sus creencias debía multiplicarse por p , al momento de elegir su número, se incluyeron variaciones en el cálculo de las diferencias normalizadas y relativas que omitían la multiplicación por p . Con ello se observó que la elección de los participantes era más consistente con el promedio de sus creencias (sin incluir la multiplicación por p) en los primeros periodos, pero conforme adquirieron experiencia entre periodos, sus elecciones se fueron acercando más a la del número objetivo estimado de acuerdo a sus creencias (en los últimos periodos del primer subjuego jugado por cada participante y durante todo el segundo subjuego, en el caso de los participante A). Este resultado indica que los participantes aprenden a incluir la multiplicación por p conforme adquieren experiencia en el juego. De cualquier forma, no es posible determinar si los participantes incorporan el cálculo explícitamente, o simplemente aprenden de forma intuitiva a elegir números cada vez más pequeños, por debajo del promedio de sus creencias.

El inicio del segundo subjuego estuvo marcado por la introducción de dos nuevos jugadores (D y E) que reemplazaron a dos de los jugadores participantes en el subjuego 1 (B y E), siendo que uno de los jugadores de dicho subjuego permaneció durante cuatro periodos más. Con esta manipulación experimental se replicó exitosamente el efecto de Reset reportado por Slonim (2005), permitiendo evaluar la consistencia entre las elecciones y las creencias de los participantes A como una función de su experiencia, sin la influencia del efecto de suelo. Los resultados obtenidos a este respecto confirman la importancia que

tiene la experiencia de los participantes sobre su desempeño en el juego de p-beauty contest, en términos de la consistencia entre los números elegidos y los números que se estimaba que tirarían los demás jugadores. El experimento que se realizó para conducir este estudio se llevó a cabo en 10 sesiones experimentales compuestas de 8 periodos de juego. En cada periodo 3 jugadores elegían números e intentaban adivinar los números de los otros jugadores. Considerando la cantidad de datos generados con este diseño, así como teniendo en cuenta diseños utilizados en la literatura (Kocher et al., 2007, Slonim, 2005), se considera que 10 sesiones fueron suficientes para obtener resultados válidos, pero que replicar el experimento con una muestra más grande podría incrementar la robustez de los hallazgos reportados.

Capítulo 6

Conclusión

Se realizó un experimento de p-beauty contest repetido con una variación del método de explicitación de creencias presentado por Lahav (2015). Los resultados encontrados en el presente estudio confirman el hallazgo principal reportado por Lahav acerca de la consistencia con que las elecciones de los participantes en cada periodo reflejan sus creencias sobre las tiradas de sus oponentes: en un comienzo los participantes eligen números poco consistentes con las creencias reportadas, pero conforme van adquiriendo experiencia al participar en más periodos, sus elecciones y creencias se vuelven consistentes. El presente trabajo aporta evidencia a favor de la relación experiencia-consistencia, descartando al efecto de suelo como la causa principal de la reducción de las diferencias registradas entre elecciones y creencias en cada periodo, mediante la incorporación de un segundo subjuego donde participantes que adquirieron experiencia durante el primer subjuego fueron enfrentados a nuevos oponentes, generando un efecto de reset que llevara a los participantes con experiencia a elegir números más grandes pero consistentes con sus creencias.

Además, los resultados encontrados muestran que los participantes no sólo se vuelven más consistentes conforme adquieren experiencia, sino que también comienzan a elegir números que caen por debajo del promedio de sus creencias (lo cual podría sugerir que aprenden a incorporar la multiplicación por p al elegir el número con el que competirán en cada periodo). Todos los juegos de p-beauty contest que se llevaron a cabo en este trabajo de tesis se realizaron con grupos de únicamente tres participantes. Esta decisión obedeció principalmente a mantener la misma cantidad de jugadores con y sin experiencia que reporta Slonim (2005), además de facilitar la explicitación de creencias especí-

ficas al contar con grupos pequeños. Se realizaron 10 sesiones experimentales de acuerdo con diseños experimentales similares presentes en la literatura (Ho, Camerer Weigelt, 1998, Kocher, Sutter Wakolbinger, 2007, Slonim, 2005). Adicionalmente, la repetición del juego por 4 periodos en cada subjuego permitió detectar aquellos participantes que no tienen una estructura discernible en sus elecciones.

La principal fortaleza del presente diseño experimental radica en la inclusión de un segundo subjuego donde nuevos jugadores entran en el juego, permitiendo con ello:

- Un análisis de medidas repetidas de las creencias y elecciones del participante A en dos escenarios diferentes (un Subjuego sin experiencia y un Subjuego con esta).
- Un mayor tamaño de la muestra de jugadores sin experiencia.
- Poner a prueba, a partir del efecto de reset reportado en las elecciones de jugadores con experiencia al enfrentarse a jugadores sin experiencia, la posibilidad de que los resultados reportados por Lahav (2015), seann un co-producto de la tendencia hacia el equilibrio que se ha observado de manera consistente en juegos repetidos de p-beauty contest.
- Al enfrentar al participante A con distintos grupos de jugadores, se garantiza que el incremento observado en la consistencia entre sus creencias y elecciones no es resultado de la retroalimentación específica que obtiene sobre la conducta de sus oponentes.

En conjunto, estos elementos permiten atribuir los resultados obtenidos a la experiencia, lo que tiene implicaciones directas sobre la pregunta de investigación planteada en un inicio, y en general, la literatura en teoría de juegos y

razonamiento iterado.

Lahav (2015) concluye que parece haber poca evidencia en favor de la consistencia entre creencias y elecciones, y que el nivel de sofisticación de las personas no puede estimarse a partir de sus creencias pues estas no parecen verse reflejadas en sus elecciones. El autor propone como posibles explicaciones una “anomalía conductual” o el malentendido de las instrucciones, reconociendo la necesidad de más investigaciones. Por su parte, Costa-Gomes y Weizsäcker (2008), que estudiaron juegos con creencias hechas explícitas en los que los jugadores no recibían retroalimentación y quienes también encontraron que las elecciones de los participantes no eran consistentes con sus creencias, concluyen que no es posible asumir que las elecciones de las personas son dirigidas por las creencias que tienen sobre los otros jugadores, por lo menos en juegos complejos con ausencia de retroalimentación, y que la construcción de expectativas sobre la conducta de otros jugadores implica un proceso de aprendizaje a partir de la retroalimentación recibida en interacciones repetidas.

Debido al diseño del experimento de Lahav (2015), resulta difícil asumir que los jugadores tienen expectativas individuales sobre las elecciones de cada jugador, al tener que enfrentarse a un amplio número de ellos (veinte jugadores por sesión experimental). Por su parte, aunque en el experimento de Costa-Gomes y Weizsäcker (2008), los juegos eran de solo dos participantes, se evitó la retroalimentación, lo que no permitió que se crearan expectativas sobre el otro jugador. En contraste, en el presente experimento se trabajó con grupos pequeños de jugadores y se permitió la retroalimentación al final de cada periodo de juego, dando oportunidad a que se crearan expectativas sobre los otros jugadores. Sin embargo, la posibilidad de formarse expectativas acerca del comportamiento de los otros jugadores no es suficiente para explicar el aumento en la consistencia entre creencias y elecciones por parte de los participantes, pues en el segundo subjuego donde se introducen nuevos jugadores el incremento observado en la consistencia de las elecciones se mantiene. Que el aumento en la consistencia entre las creencias y elecciones del participante A se mantenga a pesar de que sus oponentes son sustituidos en el segundo subjuego, sólo puede atribuirse a su experiencia, ya que de acuerdo con lo previamente sugerido por Slonim (2005),

adquirir experiencia acelera la tasa de aprendizaje en un juego, o en juegos similares, tal y como se vio reflejado en sus resultados, donde luego del primer subjuego, la convergencia al equilibrio fue más rápida y los jugadores con experiencia mantuvieron una mayor ventaja en el juego, aun jugando con personas nuevas. En palabras de Ho, Camerer Weigelt (1998), los jugadores “aprenden a aprender”.

El efecto de la experiencia se distingue del efecto de la formación de expectativas sobre los otros jugadores y sus estrategias, porque los cambios en la consistencia se mantienen aun cuando los demás jugadores cambian. De esta forma, aún si las expectativas que se tiene sobre los nuevos jugadores fueran similares a las que se tenían sobre los jugadores anteriores (en el periodo equivalente), y resultaran ser erróneas respecto de las elecciones observadas, esto no compromete a la consistencia que se registra entre sus propias creencias y elecciones, porque se ha adquirido experiencia sobre el juego en sí mismo, sus reglas y la forma en que las personas las aprenden. Esto se alinea con las conclusiones planteadas por Costa-Gomes Weizsäcker (2008), acerca de la influencia que tiene la situación (el contexto, el juego) sobre la formación de creencias y elecciones simultáneamente, que remplace el supuesto de que la situación influye primordialmente en las creencias, que posteriormente permean las elecciones.

En su revisión de modelos de no-Equilibrio en Teoría de Juegos, Crawford, Costa-Gomes, Iriberri (2013) concluyen que aunque existe evidencia experimental de que las personas se desvían del equilibrio de Nash, lo hacen de forma sistemática, a partir de un componente estructural que se puede modelar; los modelos de nivel-k son los mejores para predecir la conducta en muchos juegos con este componente, pues permiten predecir dichas desviaciones, detectando sus causas y anticipando su frecuencia, proporcionando respuestas a distintos problemas empíricos. El presente experimento replicó el efecto de reset observado por Slonim (2005), y se considera que dicho resultado es importante como evidencia a favor de los modelos de nivel-k (que establecen que las personas anclan sus creencias de manera intuitiva de acuerdo con el juego, ajustándolas mediante respuestas óptimas iteradas), en tanto que tras la introducción de nuevos jugadores (sin experiencia; en niveles cognitivos 0-1), los jugadores ex-

perimentados se alejan del equilibrio de Nash para responder de forma óptima (a la luz de sus nuevos contrincantes; eligiendo números que corresponderían con niveles cognitivos 1-2). En otras palabras, este tipo de resultados confirman la idea de que los jugadores experimentados son racionales, pero pueden tener creencias sobre los otros jugadores que impliquen asumir que estos no lo son.

Referencias