

# Elección en palomas: Un análisis a través del dilema del prisionero<sup>1</sup>

*(Choice in pigeons: Analysis through Prisoner's dilemma)*

**Eduardo Parra-García<sup>1</sup>, Javier Nieto<sup>1</sup> y Livia Sánchez-Carrasco<sup>1,2</sup>**

Universidad Nacional Autónoma de México

<sup>1</sup>Facultad de Psicología

<sup>2</sup> Instituto de Neurobiología, campus Juriquilla  
(México)

## RESUMEN

Para analizar el efecto de diferentes estrategias, simuladas por una computadora, en la adquisición de respuestas cooperativas; se entrenó a siete palomas a jugar el dilema del prisionero reiterado en contra de un oponente, simulado por una computadora, que empleaba una de tres posibles estrategias: (1) seleccionar la misma estrategia que eligió la paloma en el ensayo previo (Tic-for-Tat, TFT), (2) cooperar o traicionar aleatoriamente en cada ensayo (Random, RND) y (3) cooperar cuando la paloma traicionaba o traicionar cuando la paloma cooperaba (estrategia individualizada, I). En cada ensayo las palomas recibieron acceso diferencial al alimento después de picar la tecla asociada a la estrategia de cooperar o traicionar, por tanto, el tiempo de acceso al alimento dependía de la estrategia asignada a la computadora. Durante el experimento se entrenó a los sujetos en las tres posibles condiciones (TFT, RND o I) durante 30 sesiones de manera contrabalanceada. Los resultados muestran un mayor porcentaje de respuestas cooperativas en los sujetos que enfrentaron a un oponente que empleaba la estrategia TFT. Adicionalmente, se observó que el porcentaje de respuestas cooperativas en la condición TFT disminuyó cuando esta condición era antecedida por las condiciones RND e I. Los resultados se discuten en términos de otros estudios sobre cooperación.

*Palabras clave:* Cooperación, efectos de historia, palomas, dilema del prisionero

## ABSTRACT

To analyze the effect of different strategies, simulated by a computer, on acquisition of cooperative responses; seven pigeons played the iterated Prisoner's Dilemma against a simulated opponent pre-programmed to play one of three strategies: (1) doing whatever the pigeon did on the previous trial (Tic-for-Tat, TFT), (2) cooperate or defect randomly on each trial (Random, RND), and (3) cooperate when the pigeon defect

1) La correspondencia relacionada con el presente trabajo deberá enviarse a Eduardo Parra-García (eparrag@ciencias.unam.mx) o Livia Sánchez-Carrasco (livia@unam.mx) al Laboratorio B21 de Mecanismos Neuronales y Cognitivos del Aprendizaje Asociativo, ubicado en Av. Universidad 3004, Edificio B, Col. Copilco-Universidad, Coyoacán, México, D.F. Los autores agradecen a Rosalva Cabrera Castañón, Alejandra Valadez Vergara, Rodrigo Carranza Jasso y Alma Delia Pérez López por sus comentarios a versiones iniciales de este manuscrito.

and defect when it cooperate (Individualist strategy, I). Each pigeon received differential amounts of access to food following choices of either cooperation key or defection key on each trial, thus it was contingent with the computer's adopted strategy. Pigeons played against the three strategies (TFT, RND and I) for 30 sessions each, the order of strategies were counterbalanced among subjects. The results showed a higher percentage of cooperation responses when subjects played against TFT strategy. Additionally, the results showed that cooperative responses in TFT condition decreased after pre-exposing subjects to RND or I conditions. The results are discussed in terms of cooperation studies.

*Key words:* Cooperation, history effects, pigeons, Prisoner's dilemma

La teoría de juegos se emplea en la economía, psicología, ciencias políticas y biología para modelar la conducta cooperativa y de conflicto. Por tanto, se aplica a situaciones en las que están involucrados dos o más individuos (i.e. jugadores), que pueden realizar dos o más acciones (i.e. estrategias), así las consecuencias de sus acciones (i.e. ganancias) dependen de las estrategias ejercidas por los otros jugadores. Fundamentalmente, cuando se estudia la conducta en situaciones de interacción se emplean dos tipos de juegos: secuenciales y simultáneos; en los juegos secuenciales los jugadores deben alternar los movimientos (e.g. una partida de ajedrez), mientras que en los juegos simultáneos los jugadores pueden actuar al mismo tiempo (e.g. piedra, papel, tijeras). En un juego simultáneo es difícil conocer la estrategia que usará cada uno de los jugadores. El juego del dilema del prisionero es uno de los ejemplos más famosos de juegos simultáneos y se ha convertido en una herramienta popular para analizar la dinámica de interacciones sociales, políticas y económicas (Rapaport, 1998), así como, la interacción entre matrimonios o parejas (Dawkins, 1989), políticos (Axelrod, 1984), grupos religiosos (Lumsden, 1973), países (Hamburger, 1979) y chimpancés (Trievers, 1971). Cuando la interacción entre los jugadores es repetida se dice que el juego es iterado y se denomina dilema del prisionero iterado (DPI). El interés de analizar el DPI se centra en la dicotomía entre el conflicto y la cooperación que ocurre como resultado del juego. Por tanto, se espera que la comprensión de los factores que afectan la cooperación y el conflicto permitan resolver conflictos entre grupos de individuos (Axelrod, 1984).

El juego del dilema del prisionero toma su nombre de una historia en que la policía captura a dos sospechosos de un crimen, los interroga en habitaciones separadas y cada uno de ellos sabe que:

1. Si ambos confiesan obtendrán 10 años de cárcel cada uno.
2. Si sólo uno de ellos confiesa, el que confesó irá a prisión un año, mientras que el que no lo hizo será sentenciado a 25 años.
3. Si ninguno de ellos confiesa, no será posible aplicarles todos los cargos y sólo podrán ser sentenciados a tres años de cárcel.

La tabla del juego (ver Tabla 1), conocida como matriz de pagos, indica claramente que si el Prisionero 2 confiesa, el Prisionero 1 obtendrá 10 años de prisión si confiesa (i.e. inocente, S), o 25 años si no lo hace (i.e. castigo, P). Por otro lado, si el Prisionero 2 no confiesa, el Prisionero 1 será sentenciado a un año de prisión si confiesa (i.e. tentación, T) y a tres años de prisión si no lo hace (i.e. recompensa, R). Como puede observarse las ganancias para el Prisionero 2 serán las mismas que para el Prisionero 1, si éste realiza las mismas acciones. En términos de la teoría de juegos, confesar es la mejor estrategia (i.e. estrategia dominante) para ambos jugadores, ya que con ella obtienen una sentencia de menor duración, independientemente de la estrategia que elija el otro prisionero. Asimismo, es importante señalar que los valores de las ganancias, cuyas notaciones corresponden a las recompensas asociadas a cada celda que se muestra en la Tabla 1, deben satisfacer las siguientes desigualdades para que el juego sea considerado un juego del dilema del prisionero verdadero (1)  $T > R > P > S$  y (2)  $R > (T+S)/2$  (Green, Price y Hamburger, 1995). Sin embargo, el dilema aparece si ambos jugadores confiesan, ya que cada uno recibirá una sentencia mayor a la que obtendría si

ambos cooperaran (i.e. no confesaran, 3 años de sentencia para cada uno). Axelrod y Hamilton (1981) sugirieron que en el juego de DPI la cooperación entre los jugadores emerge si las interacciones cooperativas ocurren de manera repetida cuando la acción inicial es cooperativa y los movimientos de ambos jugadores replican las acciones previas. El ejemplo más estudiado de este tipo de estrategias se conoce como Tit-for-Tat (TFT) (Axelrod, 1980), en la cual el Jugador 2 coopera en el primer ensayo sin importar lo que hará el Jugador 1, mientras que para los ensayos siguientes la regla para seleccionar una estrategia es, optar por la misma estrategia que eligió el Jugador 1 en el ensayo anterior. Por tanto, es posible suponer que los sujetos iniciarán un juego cooperando y continuarán haciéndolo mientras el otro jugador lo haga.

Tabla 1. Matriz de pagos del juego del dilema del prisionero.

		<i>Prisionero 1</i>	
		<i>Confesar</i>	<i>No Confesar</i>
<i>Prisionero 2</i>	<i>Confesar</i>	10 años, 10 años (P, P)	25 años, 1 año (S, T)
	<i>No Confesar</i>	1 año, 25 años (T, S)	3 años, 3 años (R, R)

Los valores a la izquierda de la coma corresponden a los pagos del Prisionero 1, mientras que los valores del lado derecho corresponden a los pagos del Prisionero 2. En este caso T= 1 año, R= 3 años, P= 10 años y S= 25 años.

Recientemente, se ha empleado el juego del dilema del prisionero con el propósito de determinar si los animales no humanos pueden aprender a cooperar cuando se enfrentan a estrategias que favorecen la cooperación, tal como TFT (e.g. Green, Price y Hamburger, 1995; Harris y Maden, 2002; Maynard-Smith, 1984). Estos estudios se han realizado empleando diferentes aproximaciones metodológicas y especies. Por un lado, existen trabajos que sugieren la existencia de cooperación en golondrinas (Lombardo, 1985) y peces (Milinski, 1987; Dugatkin, 1988), la cual a su vez se ajusta a las predicciones de la estrategia TFT; es decir, se observa mayor cooperación cuando el oponente emplea también estrategias cooperativas. Sin embargo, Lazarus y Metcalfe (1990; Masters y Waite, 1990) sugieren que antes de concluir que los sujetos descritos en los experimentos previos son sensibles a la cooperación es necesario demostrar que las ganancias obtenidas equivalen a una matriz de pagos que cumple con los requisitos de un juego del dilema del prisionero verdadero, así como determinar que los sujetos cumplen con los criterios secuenciales de dicha estrategia, es decir cooperan o traicionan después de que el oponente ha elegido la estrategia correspondiente en el ensayo previo. El caso es el mismo para hallazgos reportados por Rapoport, Guyer y Gordon (1976), de los cuales es difícil derivar conclusiones sobre las características de las estrategias empleadas por los sujetos experimentales, ya que la estrategia elegida por el sujeto experimental dependía de la acción realizada por otro sujeto, de cuya conducta los experimentadores no tenían control.

En términos generales, la principal desventaja de los estudios que muestran cooperación en animales no humanos es la falta de control experimental sobre la estrategia del oponente y la falta de información sobre la matriz de pagos. Sin embargo, Green, et al. (1995) diseñaron un procedimiento de condicionamiento operante que permite analizar la conducta de palomas en un juego del dilema del prisionero, cuya matriz de pagos cumple los criterios descritos previamente, a la vez que permite determinar las acciones del oponente. En dicho experimento, se expuso a palomas a diferentes juegos, entre ellos el dilema del prisionero, mientras se programaba la estrategia del oponente a través de una computadora. En cada uno de los juegos la

estrategia del oponente podía ser: (1) TFT, y se esperaba que bajo estas condiciones las palomas eligieran la misma estrategia, por lo que se observaría un alto número de respuestas de cooperación; o (2) aleatoria, en la cual la probabilidad de cooperar o traicionar era la misma, por lo que se esperaba observar un gran número de respuestas no cooperativas. Los resultados mostraron que independientemente de la estrategia del oponente las palomas elegían la estrategia cooperativa sólo el 10% de los ensayos en cada sesión. Por tanto, los autores concluyeron que la conducta de las palomas era controlada por el reforzamiento inmediato, es decir los animales respondían en la opción que les otorgaría el mayor pago de forma inmediata, sin importar las acciones del oponente (ver Hall, 2003). En contraposición, estudios realizados empleando procedimientos experimentales similares muestran que las palomas son capaces de responder en forma cooperativa en más del 50% de los ensayos cuando juegan contra un oponente que simula la estrategia TFT (Baker y Rachlin, 2002; Sanabria, Baker y Rachlin, 2003).

El interés en el estudio de la cooperación no sólo se ha centrado en la capacidad de los organismos a preferir estrategias cooperativas cuando enfrentan a un oponente que emplea estrategias TFT, si no también en el efecto del entrenamiento previo en la elección de una estrategia. Por ejemplo, Silverstein, Cross, Brown y Rachlin (1998) diseñaron un experimento cuyo propósito fue evaluar el efecto de la experiencia en el desempeño de estudiantes universitarios en el juego del dilema del prisionero iterado. En la primera fase del experimento, se asignó a los estudiantes a uno de cuatro grupos experimentales de forma aleatoria y se les pidió que jugaran el juego del dilema del prisionero contra un experimentador. En esta fase los participantes de cada grupo se enfrentaron a un oponente que empleaba una de cuatro posibles estrategias: TFT, RND, cooperar siempre sin importar las acciones del participante (i.e. ALLC) o no cooperar independientemente de las acciones del participante (i.e. ALLD). Los resultados mostraron que los participantes entrenados en las condiciones de ALLC y ALLD tuvieron un porcentaje de cooperación alrededor de 10%, mientras que para el grupo RND se observó un 43% y para el grupo TFT fue alrededor de 60%. En la segunda fase del experimento, todos los jugadores se enfrentaron a un oponente que empleaba la estrategia TFT. Los resultados encontrados en esta fase mostraron que tras varios ensayos los jugadores que no habían sido entrenados en la estrategia TFT comenzaron a cooperar, sin embargo, se encontró que los sujetos entrenados en la estrategia RND mostraban bajos niveles de cooperación al inicio de la fase, mientras los sujetos entrenados en la estrategia TFT mostraban mayores niveles de cooperación al inicio de la prueba. Estos hallazgos han sido relacionados con los resultados en estudios sobre historia de reforzamiento (Weiner, 1964; 1969), los cuales han mostrado la influencia de las contingencias previas en la emisión de la respuesta actual (Freeman y Lattal, 1992), en sujetos animales humanos y no humanos (Freeman y Lattal, 1992; Okouchi y Lattal, 2006). Por ejemplo, Freeman y Lattal (1992) diseñaron un experimento en el que se entrenó a los sujetos a responder a una contingencia de reforzamiento (e.g. bajo un programa de Razón Fija (RF) que cambió después de 50 sesiones a un programa de reforzamiento de intervalo fijo (IF), sin que las condiciones estímulo se modificaran. Los resultados mostraron que tras el cambio de contingencias de reforzamiento los sujetos pueden tardar en ajustarse a las nuevas contingencias.

El presente experimento se diseñó con el propósito de aportar evidencia sobre la conducta de cooperación en palomas en una tarea del DPI, así como para determinar el efecto de la historia de reforzamiento en la conducta cooperativa. Como se mencionó anteriormente existen diversos experimentos en palomas que se han diseñado con el propósito de analizar la conducta de cooperación en palomas, ante un oponente que simula diferentes estrategias (e.g. RND, TFT, chicken, etc.). En particular, los resultados son ambiguos con respecto a situaciones en las que las palomas enfrentan una estrategia TFT. Por un lado, Green, Price y Hamburger (1995), así como Hall (2003), muestran bajos niveles de respuestas cooperación, cercanos al 10%. Mientras Baker y Rachlin, 2002, así como Sanabria, Baker y Rachlin, 2003, muestran niveles de cooperación superiores, los cuales se aproximan al 50%. En particular, una característica importante del experimento reportado por Green et al. (1995) es el uso de diversas variantes de cada una de las estrategias

analizadas a lo largo del entrenamiento, por ejemplo, una de las palomas que fue entrenada inicialmente en TFT en la primera condición, durante la segunda condición continuó enfrentando la estrategia TFT pero con los colores de las opciones de respuesta intercambiados, por último en la tercera condición se restablecieron las condiciones de la primera condición. Lo mismo ocurrió cuando los sujetos enfrentaron la estrategia RND. Adicionalmente, se emplearon otras variantes de la estrategia TFT donde se modificó la matriz de pagos. Por tanto, estas manipulaciones no permiten establecer un análisis más detallado del efecto que tiene la condición previa en la conducta de cooperación posterior. Finalmente, es importante señalar que aunque en cada uno de los estudios descritos las palomas se enfrentaron a un oponente que empleaba distintas estrategias, el principal interés de los mismos fue evaluar las respuestas de cooperación emitidas por las palomas ante cada una de ellas. Mientras que el presente estudio nos permitirá determinar también si entrenar a palomas con oponentes que emplean estrategias no cooperativas (i.e. RND e I) afectará el porcentaje de respuestas cooperativas que se registró en fases donde se enfrentan a oponentes que usan estrategias cooperativas (i.e. TFT).

## MÉTODO

### *Sujetos*

Se emplearon siete palomas, experimentalmente ingenuas, de aproximadamente un año de edad. Las cuales se mantuvieron con libre acceso al agua y al  $80\% \pm 10$  g de su peso en libre alimentación, al final de las sesiones experimentales se les administró alimento (i.e. mezcla de granos) suplementario para mantenerlas en su peso.

### *Aparatos*

Se utilizaron cuatro cajas operantes para palomas MED Associates® (30 cm largo, 24 cm ancho y 29 cm alto), las cuales se colocaron dentro de cajas sonoamortiguadas que tenían un ventilador que funcionó como ruido blanco para atenuar los sonidos provenientes del exterior.

El piso de la caja era de barras de acero inoxidable con una separación de 8mm. En el panel frontal de las cajas había tres teclas de respuesta y una abertura rectangular por la cual los sujetos tenían acceso al comedero. Las teclas de respuesta eran de plástico transparente con un diámetro de 2.5 cm y se podían iluminar en diferentes colores (Amarillo, Rojo, Verde), durante el experimento las teclas estuvieron alineadas verticalmente a una altura de 24.5 cm respecto al piso y con una separación de 6 cm entre cada tecla. La fuerza necesaria para accionar las teclas fue de 0.15 N. La abertura de acceso al comedero medía 6 cm de ancho por 6.5 cm de alto y estaba ubicada en la parte central del panel frontal, a una altura de 5.5 cm respecto al piso de la caja. Durante los periodos de entrega del alimento el comedero se iluminaba y daba acceso a los animales a una mezcla de grano. En el panel posterior de cada caja se encontraba la luz general a 1 cm del techo.

Las sesiones experimentales se controlaron a través de una computadora Pentium II que registraba las respuestas y presentaciones de estímulos en tiempo real a través del sistema Med-PC para Windows.

### *Procedimiento*

Antes de iniciar el experimento se entrenó a las palomas en un procedimiento de auto-moldeamiento para establecer la respuesta de picoteo. Durante esta fase, se entrenó a cada sujeto en un programa idéntico al reportado por Brown y Jenkins (1968, Experimento 1). En este programa la tecla izquierda se iluminaba en color

blanco por 8 s, y posteriormente los sujetos tenían acceso al alimento por 4 s. La sesión terminaba después de que se completaban 90 ensayos. El criterio para continuar con la siguiente fase de entrenamiento requería que las palomas emitieran al menos una respuesta en el 85% de los ensayos. En caso de que después de tres sesiones de entrenamiento los sujetos no cumplieran con el criterio de cambio de fase, se entrenaban a través de un procedimiento de reforzamiento por aproximaciones sucesivas, una vez establecido el picoteo se regresaban al programa de automoldeamiento hasta que cumplieran con el criterio de cambio de condición.

En la siguiente fase de entrenamiento se reforzaron las respuestas a las tres teclas disponibles en la cámara de condicionamiento, las cuales se iluminaron aleatoriamente en color blanco, una en cada ensayo, a fin de eliminar los sesgos de lateralidad producidos por el programa anterior. Se reforzó la respuesta de picoteo a la tecla encendida bajo un programa de Razón Fija 1. Cada sesión estuvo vigente por 90 ensayos y se realizaba el cambio de fase una vez que los sujetos concluían la sesión en menos de 60 min.

Por último, a fin de generalizar la respuesta de picoteo a las teclas iluminadas con colores se entrenó a las palomas a responder mientras dichas teclas estaban iluminadas en color rojo o verde. En cada ensayo se seleccionó aleatoriamente el color en que se iluminaría la tecla, así como la posición (izquierda o derecha), mientras se mantuvo inoperativa la tecla central, de forma tal que las respuestas a esta no tenían consecuencias programadas. En total cada una de las dos teclas se encendió en 45 de los ensayos de la sesión, en alguno de los dos colores disponibles. Esta condición se mantuvo vigente hasta que los sujetos completaron 90 ensayos de la sesión en un tiempo menor a 60 min.

Una vez concluido el entrenamiento, se inició el experimento que constó de tres fases, en las cuales los sujetos pasaron por tres diferentes condiciones: Random (RND), Tit-for-Tat (TFT) y Individualizada (I). Se asignó a los sujetos a una de las seis diferentes secuencias. Considerando que uno de los propósitos del presente estudio fue analizar la capacidad de las palomas para aprender a cooperar cuando eran expuestas a estrategias cooperativas (TFT) y no cooperativas (RND y I) se inició el entrenamiento de dos sujetos en la estrategia TFT, tres más a la estrategia RND y dos más en la estrategia I. Adicionalmente, se entrenó a los sujetos en las otras dos condiciones restantes durante las dos fases posteriores, a fin de evaluar el efecto del entrenamiento previo con estrategias no cooperativas en la ejecución de la condición TFT (Ver Tabla 2).

Tabla 2. Diseño del Experimento, se muestran las condiciones a las cuales fueron asignados los sujetos

<i>Sujeto</i>	<b>Condición</b>		
	<i>1</i>	<i>2</i>	<i>3</i>
Paloma 1	RND	TFT	P
Paloma 2	P	RND	TFT
Paloma 3	RND	TFT	P
Paloma 4	TFT	RND	P
Paloma 6	P	TFT	RND
Paloma 7	TFT	P	RND
Paloma 8	RND	P	TFT

Los ensayos en los programas de la fase experimental (i.e. TFT, RND o I) se diseñaron de acuerdo al procedimiento empleado por Green et al. (1995). Cada ensayo tuvo una duración de 25 s, al inicio se iluminaba

la luz general de la caja durante 2 s, después de esto la luz general se apagaba y las teclas laterales de la caja eran activadas una en color rojo y la otra en color verde. La posición en la que aparecía cada uno de estos colores se cambiaba aleatoriamente durante la sesión. Las teclas permanecían activas por un máximo de 10 s si la paloma no respondía, las teclas se apagaban y la caja permanecía oscura el tiempo restante hasta completar los 25 s. Si la paloma respondía en una de las teclas, la tecla donde respondió permanecía encendida por otros 6 s mientras que la otra tecla se apagaba inmediatamente, 3 s después de la respuesta de la paloma se encendía la tecla central en colores rojo o verde y permanecía encendida 3 s más. El color de dicha tecla dependía de la respuesta que hubiera emitido el sujeto en el ensayo anterior y reflejaba la estrategia asignada por la computadora. Posteriormente, se activaba el comedero, dando acceso a los sujetos a la mezcla de granos, el tiempo de acceso estuvo determinado por las combinaciones de colores, presentadas en la matriz de pagos que aparece en la Tabla 3.

*Tabla 3.* Tiempo de acceso al comedero que podían obtener los sujetos en cada ensayo. Rojo y verde se refiere al color de la tecla en que respondió la paloma en un ensayo. En el caso de la computadora se refiere al color de la tecla central que se iluminó en el mismo ensayo. En este caso T= 5s, R= 3s, P=1s y S= 0s

		<i>Computadora</i>	
		<i>Rojo</i>	<i>Verde</i>
Paloma	Rojo	3 s	0 s
	Verde	5 s	1 s

Cuando transcurría el tiempo asignado para la entrega de alimento la caja experimental permanecía apagada el tiempo necesario para completar los 25 s de cada ensayo. Cuando se completaron los 25 s de un ensayo, comenzaba inmediatamente el siguiente ensayo. La sesión terminaba cuando se completaban 55 ensayos y cada fase estuvo vigente durante 30 sesiones.

La diferencia entre cada uno de los programas estuvo marcada por el criterio empleado para indicar la respuesta del otro jugador (i.e. encender la tecla central en color rojo [cooperar] o verde [no cooperar]). En el programa RND la tecla central se encendía en color rojo o verde después de que la paloma respondía en alguna de las teclas, cada una de las opciones con una  $p = 0.5$ , es decir, la respuesta de la computadora no dependió de la respuesta de la paloma. Para la condición TFT la computadora encendió la tecla central en color rojo en el primer ensayo sin importar lo que la paloma hubiera elegido, para los siguientes ensayos la computadora encendía la tecla con el mismo color que la paloma eligió el ensayo anterior. En la condición I la computadora encendía la tecla central en color rojo en el primer ensayo sin importar lo que la paloma hubiera elegido, mientras que para los siguientes ensayos el color de la tecla central se elegía dependiendo lo que la paloma hubiera respondido en el ensayo anterior, el color de la tecla central era opuesto al color que eligió la paloma en el ensayo anterior (e.g. si la paloma elegía verde entonces la computadora elegía rojo). Así, la respuesta cooperativa estaba representada por el color rojo y la no cooperativa por el color verde.

Una vez concluidas las fases del experimento se realizó una prueba adicional con el propósito de determinar si las respuestas de los sujetos eran sensibles al reforzamiento, para ello se continuó el entrenamiento en los programas descritos previamente, pero la emisión de una respuesta durante los 15 s posteriores a que se encendiera la luz producía la pérdida del reforzador. Esta fase estuvo vigente durante siete sesiones.

## RESULTADOS

Durante el experimento los sujetos respondieron en la mayoría de las sesiones en al menos el 70% de los ensayos. Para realizar el análisis se contabilizaron las respuestas cooperativas (i.e. presión de la tecla roja) y se transformaron en porcentajes. Con el propósito de analizar el porcentaje de respuestas cooperativas emitidas por las palomas en cada una de las condiciones (i.e. TFT, I y RND), en ausencia de experiencia con otras estrategias, se muestra en la Figura 1 el porcentaje de respuestas cooperativas promedio por sujeto, registradas en las últimas 15 sesiones de la primera fase, así como el promedio de estos. Dicha figura muestra, que el porcentaje promedio de la condición es mayor para los sujetos del grupo TFT (i.e. cercana al 60%), que para aquellos asignados a los grupos I y RND. Asimismo, los sujetos de grupo I y RND muestran porcentajes similares de respuestas cooperativas, que se aproxima al 40% .

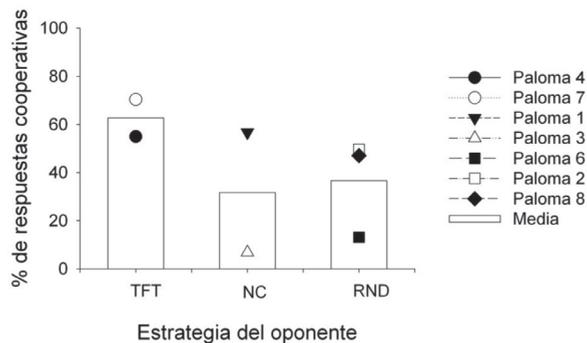
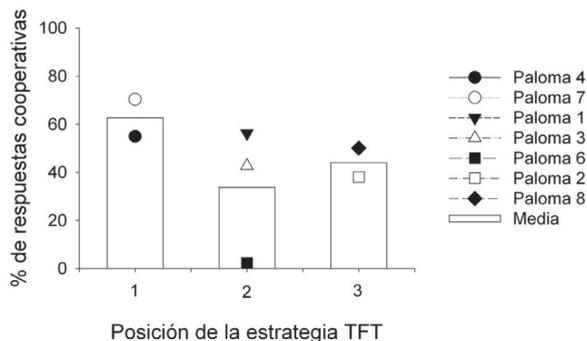


Figura 1. Porcentaje de respuestas cooperativas promedio y por sujeto durante las primera fase del experimento para cada una de las estrategias: TFT, RND y I.

Posteriormente, con el propósito de analizar el efecto del entrenamiento previo, en las condiciones que producen pocas respuestas cooperativas (I y RND), sobre la ejecución en la condición TFT se contrastó la ejecución de los sujetos en esta condición, cuando su posición en el diseño experimental se ubicó en cualquiera de las tres posiciones posibles. Se puede observar un mayor porcentaje de respuestas cooperativas cuando se entrenó a los sujetos en esa condición en primer lugar, asimismo se observa un decremento en el porcentaje de respuestas cooperativas cuando los sujetos fueron entrenados en dicha estrategia, después de ser expuestos a la estrategia I y RND. En conclusión, los resultados parecen sugerir que el entrenamiento inicial con un oponente que emplea la estrategia TFT favorece la cooperación en este mismo tipo de situaciones, mientras que el entrenamiento previo con oponentes que emplean estrategias no cooperativas o aleatorias disminuye la propensión de dichos sujetos a cooperar.

Los resultados de la fase de omisión mostraron una clara reducción en el porcentaje de respuestas para ambas teclas, que fue para la tecla izquierda del 66% al 0.04%, y para la tecla derecha del 38% al 0.03%.



*Figura 2.* Porcentaje de respuestas cooperativas promedio y por sujeto cuando los sujetos enfrentaron la estrategia TFT, sin entrenamiento con oponentes que empleaban estrategias no cooperativas (1), con entrenamiento durante 30 sesiones contra un oponente que empleó una de las dos posibles estrategias no cooperativas [NC, I] (2), y después de ser entrenados en dos diferentes condiciones de estrategias no cooperativas durante 30 sesiones cada una.

## DISCUSIÓN

El presente experimento se diseñó con dos propósitos, el primero de ellos fue aportar evidencia para determinar la capacidad de los animales no humanos de aprender a cooperar cuando enfrentan oponentes que emplean estrategias cooperativas. Los resultados revelaron que cuando los sujetos son entrenados inicialmente contra un oponente que emplea una estrategia TFT, se observa un alto porcentaje de respuestas cooperativas, el cual es cercano al 60%. Estos hallazgos se contraponen a los resultados reportados por Green et al. (1995) y Hall (2003) que muestran un porcentaje de respuestas cooperativas cercanos al 10% en palomas entrenadas contra un oponente que usa la estrategia TFT en el juego del dilema del prisionero y son congruentes con otros resultados realizados con animales no humanos (Baker y Rachlin, 2002; Sanabria, Baker y Rachlin, 2002; Lombardo, 1985; Milinski, 1987; Dugatkin, 1988) y humanos (Silverstein, et al., 1998) que sugieren que en situaciones de cooperación los organismos muestran un mayor número de respuestas cooperativas. Los resultados del presente experimento son relevantes porque a diferencia del experimento reportado por Green et al. (1995) las palomas muestran conductas cooperativas. Es posible atribuir esta diferencia a alguno de los siguientes factores: señalización de la recompensa, el tipo de consecuencia empleado, así como el hecho de que el pre-entrenamiento fue más exhaustivo.

El segundo propósito del presente trabajo fue analizar el efecto del entrenamiento previo en estrategias no cooperativas sobre la propensión a cooperar. Los resultados sugieren que el entrenamiento inicial con oponentes que emplean las estrategias RND e I hace que los sujetos muestren un menor número de repuestas de cooperación (Ver Figura 2). Una posible interpretación de este hallazgo es que los sujetos sean insensibles al valor de la consecuencia y estén respondiendo siempre en una sola opción de respuesta. Por ello, y con el propósito de asegurar que los animales eran sensibles a la matriz de pagos y que sus respuestas eran mantenidas por las contingencias de reforzamiento se realizó una fase de prueba en la cual la emisión de una respuesta producía el retiro del reforzador. Los resultados mostraron un decremento abrupto en el nivel de respuesta, lo cual es consistente con hallazgos que muestran que sólo las respuestas instrumentales y no los hábitos producen una disminución en el nivel de respuesta al introducir el procedimiento de omisión

(Dickinson, y Chanock, 1985; Dickinson, Squire, Varga y Smith, 1998). Así, es difícil atribuir los presentes resultados a la insensibilidad de los sujetos a la recompensa.

En conclusión, a partir de nuestros resultados es posible sugerir que las palomas son capaces de aprender a responder de forma cooperativa cuando enfrentan oponentes que emplean la estrategia TFT, sin embargo, a fin de analizar con mayor detalle el efecto que tiene sobre la cooperación el entrenamiento previo en estrategias no cooperativas será necesario incluir más grupos que permitan determinar los factores que producen un mayor decremento en la propensión a cooperar, que en el presente experimento puede atribuirse al tiempo de entrenamiento en la estrategia no cooperativa o el tipo de estrategia empleada. Adicionalmente, sería de particular interés realizar experimentos empleando roedores como sujetos, ya que algunos autores (Baker y Rachlin, 2002; Green et al. 1995) han sugerido estas especies muestran funciones de descuento temporal menores que las observadas en palomas (Ver, Hall, 2003), por lo que esperaríamos que hubiera una mayor propensión a cooperar.

## REFERENCIAS

- Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(1), 3-25. doi: 10.1177/002200278002400101
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books
- Axelrod, R. y Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211, 1390-1396.
- Baker, F. y Rachlin, H. (2002). Self-control by pigeons in the prisoner's dilemma. *Psychonomic Bulletin and Review*, 9(3), 482-488. doi:10.3758/BF03196303
- Brown P. L. y Jenkins H. M. (1968). Autoshaping of the pigeon's key peck. *Journal of the Experimental Analysis of Behavior*, 11, 1-8. doi: 10.1901/jeab.1968.11-1
- Dawkins, R. (1989). *The Selfish Gene*, Oxford University Press, 2nd edn
- Dickinson, A. y Charnock, D. 1985. Contingency effects with maintained instrumental reinforcement. *Quarterly Journal of Experimental Psychology*, 37B: 397-416.
- Dickinson A, Squire S, Varga Z, Smith JW (1998) Omission learning after instrumental pretraining. *Quarterly Journal of Experimental Psychology*, 51B, 271-286.
- Dugatkin, L.A. (1988). Do guppies (*Poecilia reticulata*) play tit for tat during predator inspection visits? *Behavioral Ecology and Sociobiology*, 23, 395-399.
- Green, L., Myerson, J., Holt, D., Slevin, J., y Estle, S. (2004) Discounting of delayed food reward in pigeons and rats: Is there a magnitude effect?. *Journal of Experimental Analysis of Behavior*, 81, 39-50, doi: 10.1901/jeab.2004.81-39.
- Green, L., Price, P.C. y Hamburger, M. E. (1995). Prisoner's dilemma and the pigeon: Control by immediate consequences. *Journal of the Experimental Analysis of Behavior*, 64, 1-17, doi: 10.1901/jeab.1995.64-1.
- Freeman, T. J. y Lattal, K. A. (1992). Stimulus control of behavioral history. *Journal of the Experimental Analysis of Behavior*, 57, 5-15. doi: 10.1901/jeab.1992.57-5.
- Hall, S. (2003). Transitions between cooperative and non-cooperative responding in the pigeon's dilemma. *Behavioral Processes*, 60(3), 199-208. doi: 10.1016/S0376-6357(02)00123-7
- Harris, A. C. y Madden G. J. (2002). Delay discounting and performance in prisoner's dilemma game. *Psychological Record*, 52, 429-440.
- Hamburger, H. (1979) *Games as Models of Social Phenomena*. W.H.Freeman.
- Lazarus, J. y Metcalfe, N. B. (1990). Tit-for-tat cooperation in sticklebacks: A critique of Milinski. *Animal Behaviour*, 39, 987-988. doi: 10.1016/S0003-3472(05)80965-0

- Lombardo, M. P. (1985). Mutual restraint in tree swallows: A test of the tit for tat model of reciprocity. *Science*, 227, 1363-1365.
- Lumsden, M. (1973) 'The Cyprus conflict as a Prisoner's Dilemma game'. *Journal of Conflict Resolution*, 17:7-31.
- Masters, W. M., y Waite, T. A. (1990). Tit-for-tat during predator inspections, or shoaling? *Animal Behaviour*, 39, 603-604. doi: 10.1016/S0003-3472(05)80431-2
- Maynard-Smith, J. (1984). Game theory and the evolution of behavior. *The Behavioral and Brain Sciences*, 7, 95-125. doi: 10.1017/S0140525X00026327
- Milinski, M.(1987). Tit for tat in sticklebacks and the evolution of cooperation. *Nature*, 325, 433-435.
- Okouchi, H. y Lattal, K. (2006). An analysis of reinforcement history effect. *Journal of the Experimental Analysis of Behavior*, 86, 31-42. doi: 10.1901/jeab.2006.75-05
- Rapaport, A. (1998) *Decision theory and decision behaviour*, Macmillan, Houndmills
- Rapoport, A., Guyer, M. J., y Gordon, D. G. (1976). The 2 X 2 game. Ann Arbor: The University of Michigan Press.
- Sanabria, F., Baker, F. y Rachlin, H. (2003). Learning by pigeons playing against Tit-for-tat in an operant prisoner's dilemma. *Learning & Behavior*, 31(4), 318-331. doi: 10.3758/BF03195994
- Silverstein, A., Cross, D., Brown, J. y Rachlin, H. (1998). Prior experience and patterning in a prisoner's dilemma game. *Journal of Behavioral Decision Making*, 11, 123-138. doi: 10.1002/(SICI)1099-0771(199806)11:2<123::AID-BDM283>3.0.CO;2-5
- Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology* 46 (1): 35-57. doi:10.1086/406755
- Weiner, H. (1964). Conditioning history and human fixed interval performance. *Journal of the Experimental Analysis of Behavior*, 7, 383-385. doi: 10.1901/jeab.1964.7-383
- Weiner, H. (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior*, 12, 349-373. doi: 10.1901/jeab.1969.12-349
- Williams, D.R. y Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *Journal of Experimental Analysis of Behavior*, 12:511-520

Received: October 1, 2013;  
Accepted: February 26, 2014