



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN PSICOLOGÍA
ANÁLISIS EXPERIMENTAL DEL COMPORTAMIENTO**

**ANÁLISIS DEL IMPACTO DEL VALOR INCENTIVO DE LOS
ESTÍMULOS Y EL PAPEL DE LA INHIBICIÓN CONDICIONADA EN
EL PROCEDIMIENTO DE “ELECCIÓN SUBÓPTIMA”.**

T E S I S

**QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN PSICOLOGÍA**

PRESENTA:

JUAN RODRIGO ALBA GONZÁLEZ

TUTOR PRINCIPAL:

**DR. OSCAR VLADIMIR ORDUÑA TRUJILLO
FACULTAD DE PSICOLOGÍA, UNAM**

MIEMBROS DEL COMITÉ

**DR. ARTURO BOUZAS RIAÑO
FACULTAD DE PSICOLOGÍA, UNAM**

**DR. FELIPE CABRERA GONZÁLEZ
UNIVERSIDAD DE GUADALAJARA**

**DR. OSCAR ZAMORA ARÉVALO
FACULTAD DE PSICOLOGÍA, UNAM**

**DR. GUSTAVO BACHÁ MÉNDEZ
FACULTAD DE PSICOLOGÍA, UNAM**



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

A la memoria de Miguel Alva Pineda

Crecí y aprendí muchas cosas con él, incluso más de lo que yo creía. Recuerdo que de niño me gustaba ver a mi abuelo reparar cualquier cosa. Siempre habíamos tenido una relación especial, pero algo que marcó nuestra relación fue sin duda el enseñarme a utilizar cualquier herramienta. Eso, sin darnos cuenta fue el inicio de toda mi inquietud que desde niño tenía por conocer el funcionamiento de las cosas, reparar, crear, responder preguntas que por más pequeñas que sean, siempre han motivado mi deseo por descubrir y seguir aprendiendo. Teníamos una entrañable complicidad y hoy que culmino una de las etapas más hermosas de mi vida, quiero agradecerte por darme las herramientas para descubrir al mundo.

Y aquí estamos 5 años después en la vida de un NO tan joven investigador. Después de todo, ¿Qué es un investigador? Nada más que un niño curioso que mira a través de una ventana pequeña tratando de explicar el porqué de las cosas. Muchas veces ese niño se olvida de cosas importantes y no se da cuenta de ello. El Rodrigo de hace 5 años estaría de acuerdo con esta respuesta, sin embargo, hoy en día creo que para todo debe existir un perfecto equilibrio y sin el apoyo de la familia y amigos no estaría hoy aquí escribiendo estas líneas que representan la culminación del doctorado.

Agradezco a mi familia por apoyarme no solo en esta etapa académica, gracias por siempre ayudarme a cumplir todas mis metas, este doctorado es pequeño relativo a todas las enseñanzas que me han dado ustedes de la vida, son el mejor equipo que puede existir. A veces la vida nos pone el marcador en contra, pero siempre todos juntos ganamos cualquier batalla. Papá y Mamá gracias por su infinito apoyo, ustedes dos son unos grandes luchadores que sacaron adelante a la familia, cada risa, cada esfuerzo, cada adversidad ganada nos tiene hoy en día así de felices y unidos. Miguel gracias por preocuparte por mí y siempre apoyarme, siempre encuentro en ti ese apoyo incondicional ante cualquier situación, sabes que esto es mutuo y entre leones nos cuidamos. Hanna, el hermano gruñón ahora está un poco lejos, pero está siempre cerca para apoyarte en todo lo que quieras, gracias por siempre ser esa hermana cariñosa y divertida o al menos reírte conmigo, sabes, eso es de las cosas que más extraño.

Quiero agradecer a una persona en especial, Montserrat una plática un viernes por la tarde cambió por completo el rumbo de nuestra vidas, gracias por ser mi amiga, compañera, coautora, consejera, cómplice, de todo un poco. Sabes que el doctorado no sería lo mismo sin ti a mi lado, gracias por las risas, las sesiones experimentales, las largas pláticas acompañadas de una coca de vidrio jajaja. La mayoría de este camino lo hice acompañado de la mejor persona. Nos quedan muchas aventuras e historias por vivir.

Gracias a mis amigos de la Facultad, Andrea, Mario y Mariel, cada uno de ustedes enriquecieron mis días con algunas pláticas ñoñas y la mayoría divertidas, espero pronto con cada uno de ustedes celebrar sus doctorados. Gracias a mis compañeros del Laboratorio Enrique, Julio, Axel, Alfredo, Rodrigo, Fernanda y William, por su todo su apoyo. En este laboratorio conocí lo que es trabajar con un verdadero equipo, sé que todos muy pronto estarán terminando sus doctorados. Dany, gracias por ser un excelente amigo, nunca imaginé que encontraría a una persona que me complementaría tan bien en el trabajo, gracias por llegar a las 6:30 am para iniciar las sesiones experimentales y por todo el apoyo personal y académico, amigo créelo si tú quieres llegarás muy lejos, siempre estaré en primera fila aplaudiendo tus éxitos.

El autor reconoce su deuda con su director de tesis Vladimir Orduña, gracias por la ayuda académica y personal durante todo el doctorado, gracias por creer en mí desde el primer día que estuve en tu laboratorio, por siempre apoyar mis propuestas y enseñarme que una sesión de trabajo en un artículo siempre es mejor si hay cocas de vidrio involucradas. Doctor Arturo gracias por siempre de una manera humilde inspirar a sus alumnos a seguir aprendiendo, una simple plática con usted siempre es motivo para generar más investigaciones. Doctor Felipe gracias por todo su apoyo durante mi doctorado, por siempre inspirarme esa motivación por nunca dejar de hacer preguntas y ver nuestras investigaciones con implicaciones más allá de lo aparentemente cubierto. Gracias a mis jurados Dr. Oscar Zamora y Dr. Gustavo Bachá, por su atención al presente trabajo, ustedes me conocieron desde el inicio de mi carrera y es para mí un honor culminar mi doctorado bajo sus asesorías.

Por último, quiero agradecer muy especialmente al Dr. Florente López, desde que me entregó su carta de recomendación para ingresar al posgrado me dijo que ya estaba del otro lado, que este doctorado era mío. Encontrarlo en los pasillos del edificio D siempre fue motivo de pequeñas pláticas y risas, nunca olvidaré su frase de “la cosecha de experimentos nunca se acaba”.

Especialmente agradezco a mi amigo el Dr. Julio Espinosa por su constante e incondicional apoyo a lo largo de toda mi formación, lamentablemente se adelantó, pero sin usted no estaría cumpliendo mi sueño de ser doctor.

Gracias a todos los que en esta ocasión olvidé, pero saben que cada uno de ustedes han marcado mi vida de una manera muy importante.

Gracias a todos

Juan Rodrigo Alba González

Para Mireya:

Los años a tu lado parecen pocos comparados con todo lo que nos falta por vivir, cada día que paso a tu lado me haces una mejor persona, cada día aprendo tantas cosas de ti, eres el mejor equipo de la vida. En la dedicatoria de licenciatura soñaba con hacer una vida junta a ti, hoy que lo estoy viviendo confirmo que eres lo mejor que me puede pasar en la vida, eres lo más valioso que tengo. Gracias por apoyarme en todas mis locuras y soportar a este gruñón. Eres mi familia, junto con nuestros compañeros de este hermoso camino, Happy, Chandler, Nairobi y Joey.

Te amo.

El presente trabajo se realizó con apoyo de la beca Conacyt 450692 al autor de este trabajo y por el apoyo financiero del Proyecto 281548 CONACYT y IN306818 PAPIIT-DGAPA otorgados al director de esta Tesis, Dr. Vladimir Orduña.

ÍNDICE

RESUMEN	i
INTRODUCCIÓN	5
CAPÍTULO I PREFERENCIA POR ALTERNATIVAS CON ESTÍMULOS DISCRIMINATIVOS INDEPENDIENTEMENTE DE LA PROBABILIDAD GLOBAL DE REFORZAMIENTO EN PALOMAS	12
CAPÍTULO II HIPÓTESIS QUE EXPLICAN LA "ELECCIÓN SUBÓPTIMA" EN PALOMAS	21
EFFECTO DE CONTRASTE	22
MODELO ECOLÓGICO.....	24
HIPÓTESIS DE SEÑALES PARA BUENAS NOTICIAS (SIGN)	26
IMPACTO DEL ESTÍMULO PREDICTOR DE LA AUSENCIA DE REFORZAMIENTO	27
CAPÍTULO III GENERALIDAD ENTRE ESPECIES EN EL PROCEDIMIENTO DE "ELECCIÓN SUBÓPTIMA"	31
HUMANOS	32
RATAS	33
CAPÍTULO IV VARIABLES RESPONSABLES DE LAS DIFERENCIAS ENTRE ESPECIES EN EL PROCEDIMIENTO DE "ELECCIÓN SUBÓPTIMA"	36
INHIBICIÓN CONDICIONADA	37
VALOR INCENTIVO	38
DIFERENCIAS INDIVIDUALES CON LA CAPACIDAD DE ATRIBUIR VALOR INCENTIVO	44
CAPÍTULO V GENERALIDAD DE LA CONDUCTA ÓPTIMA EN RATAS EN EL PROCEDIMIENTO DE "ELECCIÓN SUBÓPTIMA"	50
EXPERIMENTO 1	51
MÉTODO	54
SUJETOS.....	54
APARATOS	54
PROCEDIMIENTO.....	56
RESULTADOS	58
DISCUSIÓN	64
EXPERIMENTO 2	66
MÉTODO	66
SUJETOS.....	66
APARATOS	66
PROCEDIMIENTO.....	64
RESULTADOS	69
DISCUSIÓN	71
EXPERIMENTO 3	72

MÉTODO	72
SUJETOS.....	72
APARATOS	72
PROCEDIMIENTO.....	70
RESULTADOS	74
DISCUSIÓN	76
DISCUSIÓN GENERAL CAPÍTULO V.....	77
CAPÍTULO VI GENERALIDAD DE LA CONDUCTA ÓPTIMA EN RATAS EN EL PROCEDIMIENTO DE “ELECCIÓN SUBÓPTIMA” PROPUESTO COMO UN MODELO ANIMAL DE JUEGO PATOLOGICO	81
EXPERIMENTO 4	82
MÉTODO	85
SUJETOS.....	85
APARATOS	85
PROCEDIMIENTO.....	86
RESULTADOS	88
DISCUSIÓN GENERAL.....	95
CAPÍTULO VII IMPACTO DEL VALOR INCENTIVO DE LOS ESTÍMULOS EN LA ALTERNATIVAS DEL PROCEDIMIENTO DE “ELECCIÓN SUBÓPTIMA EN RATAS Y PALOMAS”	98
EXPERIMENTO 5	99
MÉTODO	101
SUJETOS.....	101
APARATOS	102
PROCEDIMIENTO.....	105
RESULTADOS	106
DISCUSIÓN	110
EXPERIMENTO 6 A	116
MÉTODO	118
SUJETOS.....	118
APARATOS	118
PROCEDIMIENTO.....	118
RESULTADOS	121
DISCUSIÓN	122
EXPERIMENTO 6 B	123
MÉTODO	123
SUJETOS.....	123
PROCEDIMIENTO.....	123

RESULTADOS	124
DISCUSIÓN	128
DISCUSIÓN GENERAL CAPITULO VII	129
CAPÍTULO VIII DISCUSIÓN GENERAL.....	133
PROBABILIDAD GLOBAL DE REFORZAMIENTO ASOCIADA CON LA ALTERNATIVA DISCRIMINATIVA Y NO DISCRIMINATIVA	137
PRESENTACIÓN DE UN ESTÍMULO PREDICTOR DE LA AUSENCIA DE REFORZAMIENTO Y EL DESARROLLO DE LA CONDUCTA SUBÓPTIMA EN RATAS	138
GENERALIDAD DE LA CONDUCTA ÓPTIMA DE LAS RATAS EN EL PROCEDIMIENTO DE ELECCIÓN SUBÓPTIMA PROPUESTO COMO UN MODELO ANIMAL DE ADICCIÓN AL JUEGO	141
IMPACTO DEL VALOR INCENTIVO DE LOS ESTÍMULOS Y EL DESARROLLO DE LA CONDUCTA ÓPTIMA EN PALOMAS	143
VALOR INCENTIVO DE LOS ESTÍMULOS EN LAS ALTERNATIVAS DEL PROCEDIMIENTO DE ELECCIÓN SUBÓPTIMA	144
INHIBICIÓN CONDICIONADA Y VALOR INCENTIVO DE LOS ESTÍMULOS COMO VARIABLES QUE EXPLICAN LA DIFERENCIA ENTRE ESPECIES EN EL PROCEDIMIENTO DE ELECCIÓN SUBÓPTIMA	145
CONCLUSIONES.....	151
REFERENCIAS	158

RESUMEN

En el procedimiento de elección subóptima, las palomas, pero no las ratas, eligen sistemáticamente una alternativa con estímulos discriminativos que permiten discriminar la entrega y no entrega del reforzador a pesar de estar asociada a una menor probabilidad de reforzamiento, mientras que las ratas prefieren la alternativa asociada a una mayor probabilidad de reforzamiento, aunque los estímulos no permitan predecir la entrega del reforzador. Para explicar estas preferencias opuestas, se ha propuesto que las ratas y las palomas tienen un efecto diferencial del inhibidor condicionado que surge del estímulo que predice la ausencia de reforzamiento: si bien no tiene impacto en las palomas, influye fuertemente en las preferencias de las ratas. No obstante, recientemente se ha sugerido que no existe una diferencia fundamental en el comportamiento de las ratas y las palomas, pero el procedimiento empleado para evaluar cada una de estas especies ha generado la diferencia; en particular, se propuso que ambas especies prefieren la alternativa Discriminativa cuando los estímulos discriminativos tienen valor incentivo. Esta tesis incluye seis experimentos destinados a investigar la generalidad de esta propuesta debido a que la evidencia respecto al valor incentivo de los estímulos ha encontrado resultados contradictorios. El propósito general de esta investigación es evaluar la generalidad del comportamiento óptimo de las ratas y el comportamiento subóptimo de las palomas en el procedimiento de elección subóptima evaluando la hipótesis del valor incentivo de los estímulos.

Los experimentos 1-3 exploran sistemáticamente la influencia de las diferencias de procedimiento entre los estudios con resultados discrepantes. En particular se presentaron diferentes valores de la probabilidad global de reforzamiento asociada a cada una de las alternativas, se manipuló la presencia o ausencia de un estímulo predictor de no reforzamiento (inhibidor condicionado) y por último se modificaron las características de las alternativas (Discriminativa y No Discriminativa). Los resultados obtenidos en estos tres primeros experimentos mostraron una preferencia por la alternativa asociada a una mayor probabilidad global de reforzamiento, es decir, mostraron conducta óptima a pesar del alto valor incentivo de los estímulos. En el experimento 4 se evaluó la generalidad de la conducta óptima de las ratas al utilizar el procedimiento de elección subóptima que ha sido considerado como un modelo animal que permite aproximarse al estudio de la adicción al

juego. Los resultados de este experimento mostraron que las ratas se comportan de manera óptima en este procedimiento al igual que en el procedimiento clásico de elección subóptima a pesar de emplear estímulos con alto valor incentivo. En el experimento 5 se emplearon tonos (estímulos con bajo valor incentivo) en ambas alternativas usando palomas como sujetos experimentales. Los resultados de este experimento indican que las palomas prefieren la alternativa óptima cuando los estímulos empleados en este procedimiento tienen bajo valor incentivo. Por último, en el experimento 6 se manipuló de manera asimétrica el valor incentivo de los estímulos entre las alternativas Discriminativas y No Discriminativas para evaluar su impacto en la elección subóptima, esta manipulación también se llevó cabo utilizando ratas clasificadas como seguidoras de señales para evaluar si muestran una mayor tendencia a la elección subóptima con esta manipulación. Los resultados de estos experimentos mostraron que el valor incentivo de los estímulos en cada una de las alternativas sesga la preferencia en ratas por las alternativas que tengan estímulos con alto valor incentivo y este efecto fue más fuerte para los sujetos clasificados como seguidores de señales.

Los experimentos de la presente tesis investigaron el efecto del valor incentivo de los estímulos en el procedimiento de elección subóptima y han contribuido con evidencia contundente de la generalidad del comportamiento óptimo de las ratas en este procedimiento y han agregado información valiosa que sugiere que el valor incentivo de los estímulos modula el comportamiento subóptimo de las palomas.

ABSTRACT

In the suboptimal choice procedure, pigeons, but not rats, systematically choose an alternative with discriminative stimuli that allows discriminating the delivery and non-delivery of the reinforcer despite being associated with a lower probability of reinforcement, while rats prefer the alternative associated with a greater probability of reinforcement, although the stimuli that it presents do not allow predicting the delivery of the reinforcer. To explain these opposing preferences, it has been proposed that rats and pigeons have a differential effect of the conditioned inhibitor that arises from the stimulus that predicts the absence of reinforcement: while it has no impact on pigeons, it strongly influences the preferences of rats. However, it has recently been suggested that there is no fundamental difference in the behavior of rats and pigeons, but rather, the procedure used to evaluate each of these species has produced the difference; in particular, it was proposed that both species prefer the discriminative alternative when the discriminative stimuli have incentive salience. This thesis includes six experiments designed to investigate the generality of this proposal because the evidence regarding the incentive salience of the stimuli has found contradictory results. The general purpose of this research is to evaluate the generality of the optimal behavior of rats and the suboptimal behavior of pigeons in the suboptimal choice procedure by evaluating the hypothesis of the incentive salience of the stimuli.

Experiments 1-3 systematically explore the influence of procedural differences between studies with discrepant results. In particular, different values of the global probability of reinforcement associated with each of the alternatives were presented, the presence or absence of a stimulus that predicted non-reinforcement (conditioned inhibitor) was manipulated, and finally the characteristics of the alternatives (Discriminative and Non-discriminative) were manipulated as well. The results obtained in these first three experiments showed a preference for the alternative associated with a higher overall probability of reinforcement, that is, rats showed optimal behavior despite the high incentive salience of the stimuli. In experiment 4 the generality of the optimal behavior of the rats was evaluated using the suboptimal choice procedure that has been considered as an animal model that allows us to approach the study of gambling addiction. The results of

this experiment showed that the rats perform optimally in this procedure as well as in the classical procedure of suboptimal choice despite using stimuli with high incentive salience. In experiment 5, tones (stimuli with low incentive salience) were used in both alternatives using pigeons as experimental subjects. The results of this experiment indicate that pigeons prefer the optimal alternative when the stimuli used in this procedure have low incentive salience. Finally, in experiment 6 the incentive value of the stimuli between the Discriminative and Non-Discriminative alternatives was asymmetrically manipulated to evaluate its impact on the suboptimal choice. This manipulation was also carried out using rats classified as sign-trackers to assess whether they would show a greater tendency to choose suboptimally with this manipulation. The results of these experiments showed that the preference of rats is biased towards those alternatives which have greater incentive salience, and this effect is stronger for the subjects classified as sign-trackers.

The experiments of the present thesis investigated the effect of the incentive salience of stimuli in the suboptimal choice procedure. They have contributed strong evidence of the generality of the optimal behavior of rats in this procedure and have added valuable information that suggests that the incentive salience of the stimuli modulates the suboptimal behavior of the pigeons.

INTRODUCCIÓN

El comportamiento se produce en entornos complejos que brindan al organismo la oportunidad de distribuir su conducta a través de múltiples opciones de respuesta. El análisis experimental de la conducta estudia los principios fundamentales que subyacen en el comportamiento de los organismos. Históricamente, la conducta de elección ha sido uno de los principales temas de investigación en el análisis experimental de la conducta (para una revisión ver, Mazur & Fantino, 2014). Las elecciones de un organismo dependen de distintos factores, por lo que en un entorno natural resulta complicado identificar con precisión el efecto de cada uno de ellos. Es por ello que el estudio de la conducta de elección se ha llevado a ambientes controlados de laboratorio, en los que se han utilizado programas de reforzamiento para simular tanto la disponibilidad de recursos como las condiciones necesarias para obtenerlos.

Una forma de estudiar la preferencia es ofrecer a los organismos opciones que difieran en relación con sus consecuencias. Por ejemplo, un sujeto podría elegir entre un reforzador pequeño y uno más grande. En esta situación de elección medimos la preferencia evaluando la distribución de respuestas (o tiempo) a cada alternativa. Si se asignan más respuestas a una alternativa en relación con la otra, se dice que el sujeto prefiere esa alternativa. Uno esperaría que en este ejemplo el sujeto distribuya la mayoría de sus respuestas a la alternativa que proporciona el reforzador más grande.

Un ejemplo específico de un procedimiento para estudiar la elección es el procedimiento de programas concurrentes. Este procedimiento implica una elección entre dos alternativas que difieren en el reforzador asociado con cada una. A menudo, se utilizan diferentes estímulos para señalar cada alternativa, de modo que, por ejemplo, una tecla de respuesta pueda iluminarse con una luz roja y otra tecla con una luz verde. Se predice que

los sujetos elegirán entre las dos alternativas por igual (es decir, serán indiferentes entre las dos) si las dimensiones del reforzador asociadas con cada una son iguales.

En el análisis cuantitativo del comportamiento, muchas teorías y modelos matemáticos de elección se han desarrollado con el tiempo. Su objetivo común es proporcionar una descripción adecuada del comportamiento que permita predecir el comportamiento futuro. La Ley de Igualación es sin duda el modelo de elección más influyente; se derivó de un experimento realizado por Herrnstein (1961) con palomas; en este experimento cada una de dos teclas de respuesta fue asociada con un programa de reforzamiento intervalo variable (IV) diferente, por ejemplo, IV 15 segundos en la tecla de respuesta izquierda e IV 30 segundos en la tecla de respuesta derecha. En este ejemplo, las palomas obtenían dos terceras partes de los reforzadores totales en la tecla de respuesta izquierda y en esa misma tecla de respuesta realizaban dos terceras partes de sus respuestas. Al encontrar un resultado similar en las diferentes condiciones en que se manipuló la proporción de reforzamiento asociada a cada tecla, surgió la Ley de Igualación, que establece que en un programa concurrente IV-IV, la tasa relativa de respuestas iguala la tasa relativa de reforzamiento.

En su forma más básica, la Ley de Igualación relaciona la proporción de respuestas en una alternativa con la proporción de reforzadores que esa alternativa otorga:

$$\frac{B_1}{B_1+B_2} = \frac{R_1}{R_1+R_2} \quad (1)$$

donde B representa la tasa absoluta de respuestas en las alternativas y R representa la tasa absoluta de reforzamiento en las alternativas, los números 1 y 2 representan a la alternativa izquierda y derecha, respectivamente.

Se ha reportado que lo establecido en la Ley de Igualación ocurre también cuando los reforzadores de las alternativas del programa concurrente difieren en magnitud (Catania, 1963), o cuando difieren en la demora de entrega (Chung & Herrnstein, 1967). La investigación sobre la generalidad de la Ley de Igualación ha estado en curso desde su aparición. Otro ejemplo de esto es el estudio Herrnstein (1964), quien lo aplicó al procedimiento de programas concurrentes encadenados. El procedimiento de programas concurrentes encadenados fue desarrollado por Autor (1960) para evaluar la preferencia por diferentes programas concurrentes de reforzamiento. Por ejemplo, en lugar de elegir entre un programa de reforzamiento IV 30-s y un programa IV 90-s, el organismo responde en dos programas concurrentes (generalmente iguales) que conducen a un programa diferente de reforzamiento, que, a su vez, conducen a la entrega de un reforzador. Por lo general, los dos programas de reforzamiento terminales son diferentes y la preferencia del sujeto se ve reflejada en la proporción de respuestas a cada uno de los programas de reforzamiento en los eslabones iniciales.

Los programas concurrentes encadenados son similares a los programas concurrentes simples donde se presentan dos programas de reforzamiento disponibles al mismo tiempo en dos teclas de respuesta. Sin embargo, el procedimiento de programas concurrentes encadenados se divide en dos partes: el eslabón inicial (fase de elección) y el eslabón terminal (fase de resultado). En el eslabón inicial, las teclas de respuesta a menudo se iluminan del mismo color (blanco, por ejemplo). La respuesta en el eslabón inicial permite el acceso a uno de los dos eslabones terminales mutuamente excluyentes, donde se presenta otro estímulo que se correlaciona con ese eslabón terminal. Una vez que está vigente el eslabón de terminal en una de las alternativas, la otra tecla se vuelve oscura y no está operativa. El reforzador primario es otorgado cuando se cumple el requisito del eslabón terminal, y después de esto el ciclo es

reiniciado. La variable independiente cuando se usa este programa generalmente es alguna diferencia entre los eslabones terminales, y la variable dependiente es la tasa relativa de respuesta en cada eslabón inicial, la cual es tomada como una medida de la preferencia por el eslabón terminal.

Baum y Rachlin (1969) extendieron la lógica de la Ley de Igualación para estudiar el impacto de la tasa de reforzamiento sobre la asignación de tiempo de los sujetos en un lugar determinado de la caja experimental. El tiempo que pasó cada animal en cada lado de la caja experimental se registró mediante fotorreceptores en el piso de la caja y en ambos lados de la caja estaban instalados dos dispensadores de alimentos. Los reforzadores eran entregados de acuerdo a un programa concurrente IV-IV. Los autores reportaron que la proporción de tiempo que pasaba en cada lado de la caja experimental cada sujeto era directamente proporcional a la proporción de reforzadores entregados en cada lado de la caja. Con este resultado los autores aportaron una extensión conceptual importante de la Ley de Igualación, sugiriendo que podría aplicarse no solo a la tasa de reforzamiento, sino también al valor relativo asociado con cada alternativa, de esta manera variables independientes como la tasa de reforzamiento, la cantidad y la inmediatez se integran para dar un solo índice de valor.

Con esta propuesta Baum y Rachlin (1969) combinaron las tres variables, la tasa de reforzamiento, la inmediatez y la magnitud de reforzamiento, mediante una relación multiplicativa:

$$\frac{B1}{B2} = \left(\frac{R1}{R2}\right) \left(\frac{1/D1}{1/D2}\right) \left(\frac{M1}{M2}\right) = \frac{V1}{V2} \quad (2)$$

Donde R representa la tasa, M la magnitud, y $1/D$ la inmediatez de reforzamiento, y V es el valor de cada alternativa resultado de la integración de las anteriores variables.

De acuerdo con la Ecuación 2, conocida como la Ley de Igualación Concatenada, el valor se determina multiplicativamente por las dimensiones del reforzador que difieren entre las alternativas. La ecuación 2 es un avance conceptual significativo porque propone que la relación de igualación es un principio general de elección que es potencialmente aplicable a cualquier variable de reforzamiento.

El comportamiento de elección es uno de los principales temas de investigación en el análisis de la conducta, y parten de la premisa de que los organismos han evolucionado para maximizar su éxito (por ejemplo, para encontrar comida), y gran parte de las teorías del aprendizaje (Skinner, 1938; Thorndike, 1911) se basan en esta premisa. Los animales seleccionan aquellas alternativas que conducen a una mayor probabilidad de reforzamiento sobre aquellas que no lo hacen. Cuando el comportamiento de los animales es consistente con esta teoría, fortalece nuestra creencia en la validez de la teoría. Sin embargo, cuando los animales muestran una preferencia por alternativas que resultan en menos comida sobre aquellas que resultan en más comida, es importante tratar de entender por qué lo hacen.

Si bien se han postulado principios y leyes que controlan el comportamiento de elección, también se han estudiado recientemente ejemplos de elección irracional en los que el comportamiento de elección se desvía de esos principios y leyes generales. La ecuación 2 predice que los organismos mostrarán una preferencia por la alternativa con el valor de reforzamiento más alto. Sin embargo, al contrario de esta predicción, hay investigaciones que muestran que los organismos a menudo eligen una alternativa con estímulos discriminativos que proporciona una menor probabilidad de reforzamiento, sobre otra alternativa sin esos estímulos, pero con una mayor probabilidad de reforzamiento. La preferencia por esta alternativa no conduce a la maximización de los reforzadores y, por lo tanto, es un ejemplo de elección irracional.

Este comportamiento es considerado irracional porque la base evolutiva de la conducta animal sugiere que los organismos son sensibles a las alternativas que ofrecen mayor probabilidad y magnitud de alimento (Fantino & Abarca, 1985). Además, va en contra de la teoría de forrajeo óptimo (Stephens & Krebs, 1986) que sugiere que la evolución ha seleccionado animales para alimentarse de manera eficiente, de modo que, deberían elegir las alternativas que los lleven a maximizar su ingesta de alimentos.

Los siguientes apartados esta divididos en cuatro capítulos; en el primer capítulo se describen los principales resultados que dieron origen a este ejemplo de conducta irracional. En el segundo capítulo se describen las hipótesis que se han dado en la literatura para explicar este ejemplo de conducta irracional. En el tercer capítulo se presentan las investigaciones que buscaron la generalidad de este resultado en otras especies y por último en el capítulo cuatro se presentan los experimentos que surgieron a partir de las diferencias entre especies reportadas en la literatura con este procedimiento y las hipótesis que se han propuesto para explicar esta diferencia entre especies.

CAPÍTULO I

**PREFERENCIA POR ALTERNATIVAS CON ESTÍMULOS
DISCRIMINATIVOS INDEPENDIENTEMENTE DE LA
PROBABILIDAD GLOBAL DE REFORZAMIENTO EN PALOMAS**

El primer trabajo que reporta este ejemplo de elección irracional en palomas fue realizado por Kendall (1974). Específicamente, el autor presentó dos condiciones en las que evaluó los efectos de señalar la presencia o ausencia del reforzador. En la primera condición (Señalada) las palomas eligieron entre una alternativa que resultaba en la presentación de una tecla iluminada de blanco que siempre era seguida por un reforzador y una alternativa que resultaba en la presentación de dos teclas (verde y rojo), ambas con una probabilidad de presentación de .5, que eran seguidas por un reforzador o por un blackout, respectivamente. Bajo estas condiciones Kendall llamó “Baja densidad” a la alternativa asociada con los estímulos discriminativos y una baja tasa global de reforzamiento y “Alta densidad” a la alternativa sin estímulos discriminativos y una tasa global de reforzamiento mayor. En la segunda condición (No señalada), la presentación de las teclas roja y verde no predecían la entrega del reforzador. (ver Figura 1).

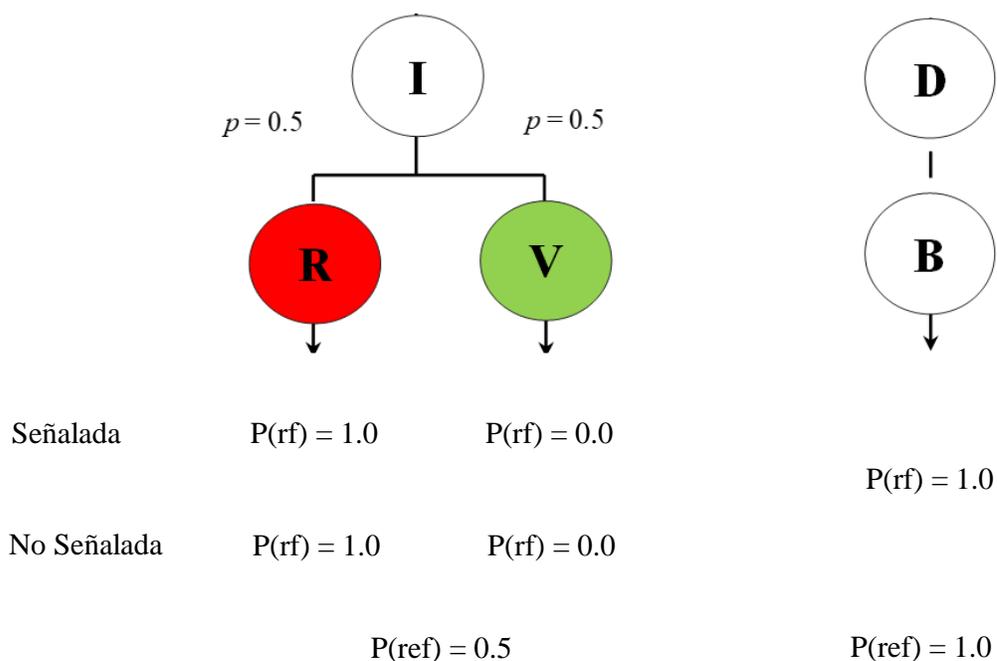


Figura 1 Procedimiento utilizado por Kendall (1974).

El hallazgo más importante de este experimento fue que las palomas eligieron consistentemente la alternativa asociada con estímulos discriminativos (Baja densidad) en la condición Señalada, y en la condición No señalada eligieron con mayor frecuencia la alternativa asociada con la mayor probabilidad global de reforzamiento (Alta densidad). Kendall (1974) explicó estos resultados en términos de reforzamiento condicionado sugiriendo que una interpretación de sus resultados es que un reforzador condicionado (el estímulo asociado con el enlace terminal) puede tener un gran impacto si ocurre en un contexto local en el cual se incluye un estímulo condicionado correlacionado con la falta de reforzador debido a que aumenta la eficacia como reforzador condicionado de los estímulos positivos asociados con la entrega del reforzador.

Fantino, Dunn, y Meck (1979) replicaron el experimento de Kendall (1974) para determinar si las características en su procedimiento desarrollaron la preferencia por la alternativa con estímulos discriminativos (Baja densidad). En específico describieron que la tecla asociada con la alternativa que presentaba los estímulos discriminativos (Baja densidad) se encontraba apagada tanto en el eslabón inicial y durante el eslabón terminal en los ensayos en que el sujeto escogía la alternativa sin estímulos discriminativos (Alta densidad). Por lo tanto, cuando la paloma eligió la alternativa asociada con la probabilidad global de reforzamiento del 50% y se presentó el estímulo asociado con el reforzador las palomas continuaron respondiendo en esa tecla, en cambio cuando se presentó el estímulo asociado con la ausencia del reforzador comenzaron a responder en la tecla inoperativa debido a su similitud con la tecla apagada del enlace inicial. Esto significaba que la tasa global de reforzamiento en la alternativa que presentaba los estímulos discriminativos (50%) estaba realmente más cerca del 100%, porque las palomas comenzaban a responder en la otra alternativa cuando se presenta el estímulo asociado con la ausencia del

reforzador. Por esta razón, Fantino et al., iluminaron de color blanco ambas teclas del enlace inicial, y con esta modificación reportaron que las palomas eligieron la alternativa que no presentaba los estímulos discriminativos. De esta manera atribuyeron el resultado de Kendall a una característica particular en su procedimiento.

Posteriormente, varios grupos de investigación intentaron replicar los hallazgos originales de Kendall y algunos estudios reportan resultados consistentes (Belke & Spetch, 1994; Experimento 3 de Dunn & Spetch, 1990; Kendall, 1974, 1985), mientras que otros muestran resultados contradictorios (Experimento 2 de Dunn & Spetch, 1990; Fantino et al., 1979; McDevitt, Spetch, & Dunn, 1997; Smith & Zentall, 2016; Stagner, Laude, & Zentall, 2012).

A pesar de discrepancias en los resultados que muestran una preferencia de las palomas por las alternativas con estímulos discriminativos, un estudio que dio una evidencia sólida a esta preferencia fue el presentado por Roper y Zentall (1999) quienes presentaron a las palomas dos alternativas que tenían la misma probabilidad global de reforzamiento (.5) pero una alternativa contaba con estímulos que señalaban la presencia o ausencia de un reforzador mientras que la otra alternativa contaba con estímulos que no señalaban la presentación de un reforzador. Los autores reportaron que las palomas mostraron una fuerte preferencia por la alternativa que proporcionaba estímulos discriminativos. Roper y Zentall (Experimento 2) intentaron obtener una medida más sensible de preferencia por los estímulos discriminativos al aumentar el requisito de respuesta para obtener los estímulos discriminativos mientras se mantenía el requisito de respuesta para obtener los estímulos no discriminativos en una sola respuesta. Bajo estas condiciones los autores reportaron que las palomas seguían mostrando una fuerte preferencia por las alternativas con estímulos discriminativos.

En una versión más reciente de esta situación de elección, Gibson, Alessandri, Miller y Zentall (2009) probaron si las palomas prefieren la alternativa asociada con los estímulos discriminativos cuando la diferencia en la probabilidad global de reforzamiento entre las dos alternativas se reduce de 50 a 25 %. Específicamente, si los sujetos elegían la alternativa asociadas con los estímulos discriminativos se presentaba una tecla que siempre era seguida de la entrega de un reforzador, o se presentaba una tecla que estaba asociada a la ausencia del reforzador. Por el contrario, si elegían la otra alternativa, se presentaban dos teclas en cuya presencia se entregaba un reforzador el 75% de las veces. Con estos parámetros, la tasa global de reforzamiento para la alternativa asociada con los estímulos discriminativos era del 50% y para la alternativa no asociada con estos estímulos era de 75%. Los autores reportaron que las palomas preferían sustancialmente la alternativa que tenía los estímulos discriminativos (13 de las 16 palomas prefirieron esta opción).

Posteriormente, Stagner & Zentall (2010) desarrollaron un procedimiento en el que disminuyeron las probabilidades globales de reforzamiento en ambas alternativas. En este procedimiento (ver Figura 2), Al elegir una de las alternativas, con probabilidad .2 se presentaba una tecla iluminada de color rojo, la cual estaba asociada con la entrega de un reforzador después de 10 s o con probabilidad .8 se presentaba una tecla iluminada de color verde, después de la cual nunca se entregaba reforzador. Puesto que los estímulos permiten discriminar la entrega o no entrega del reforzador, a esta alternativa se le considera *Discriminativa*. La otra alternativa consiste en la presentación de una tecla iluminada de color azul con una probabilidad de .2 o la presentación de una tecla amarilla con una probabilidad de .8, ambas con una probabilidad de reforzamiento de .5. Se le considera como una alternativa *No Discriminativa* debido a que durante el periodo de 10 s en presencia de cualquiera de estos estímulos, el sujeto no puede discriminar si recibirá o no reforzador. Los

autores reportaron una fuerte preferencia por la alternativa Discriminativa incluso más extrema que la reportada bajo condiciones de probabilidades .5 vs .75.

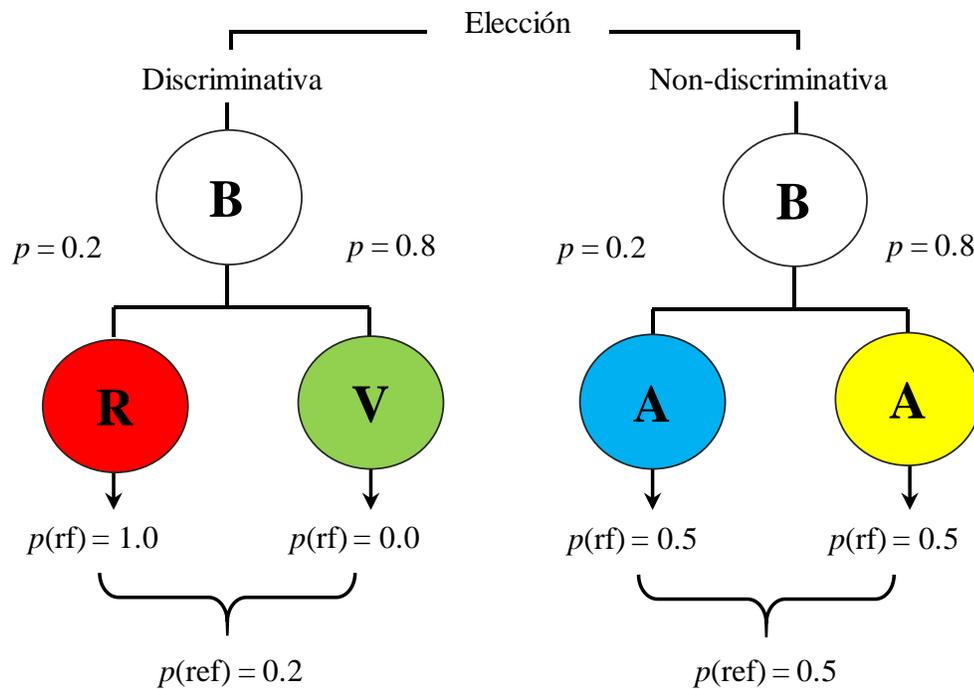


Figura 2 Procedimiento utilizado por Stagner & Zentall (2010).

En la elección entre la alternativa Discriminativa y la No Discriminativa, elegir esta última podría considerarse como conducta óptima en términos de reforzamiento primario, puesto que tiene una probabilidad neta de reforzamiento de .5, mientras que la alternativa Discriminativa tiene una probabilidad neta de reforzamiento de .2. No obstante, utilizando palomas como sujetos, se ha encontrado que muestran una mayor preferencia por la alternativa Discriminativa, es decir, prefieren una menor tasa de reforzamiento a cambio de estímulos que les permitan discriminar entre la entrega y la ausencia de una recompensa. Este resultado ha sido ampliamente replicado por diversos grupos de investigación (McDevitt, Dunn, Spetch, & Ludvig, 2016; Zentall, Smith, & Beckmann, 2019; T. R. Zentall, 2014) y

han demostrado que las palomas tienen una fuerte preferencia por la alternativa subóptima. Debido a la fuerte evidencia mostrada en más de una docena de estudios se le llamó a este procedimiento “Elección Subóptima”.

Aunque los resultados de los experimentos de Gipson et al. (2009) y Stagner y Zentall (2010) demostraron claramente el comportamiento de elección subóptima de las palomas. Es posible que el efecto que mostraron estos estudios con la manipulación de la probabilidad global de reforzamiento ocurre porque las palomas están evitando una alternativa que resulta en estímulos asociados con un resultado incierto (0.75 de probabilidad global de reforzamiento en Gipson et al., 2009, y 0.50 de probabilidad global de reforzamiento en Stagner y Zentall, 2010). Para determinar si las palomas están evitando las alternativas con mayor incertidumbre debería ser posible encontrar un efecto similar al reportado con las palomas manipulando la magnitud del reforzador, en lugar de la probabilidad global de reforzamiento, y eliminando la incertidumbre del resultado asociado con los estímulos en la alternativa No Discriminativa.

Siguiendo con estas ideas, Zentall y Stagner (2011) presentaron una modificación al procedimiento original y presentaron el siguiente procedimiento: Al elegir una de las alternativas, con probabilidad .2 se presentaba una tecla iluminada de color rojo, la cual estaba asociada con la entrega 10 pellets después de 10 s o con probabilidad .8 se presentaba una tecla iluminada de color verde, después de la cual nunca se entregaba reforzador. La otra alternativa consistía en la presentación de una tecla iluminada de color azul con una probabilidad de .2 o la presentación de una tecla amarilla con una probabilidad de .8, en presencia de ambas teclas se entregaban 3 pellets. Así, en promedio la alternativa con valor condicionado diferencial entrega 2 pellets por ensayo, y la alternativa sin valor condicionado diferencial entrega confiablemente 3 pellets por ensayo. Este procedimiento pone énfasis en

manipular la diferencia en magnitudes de reforzamiento entre las alternativas más que las probabilidades, como habían hecho los experimentos anteriores (Figura 3).

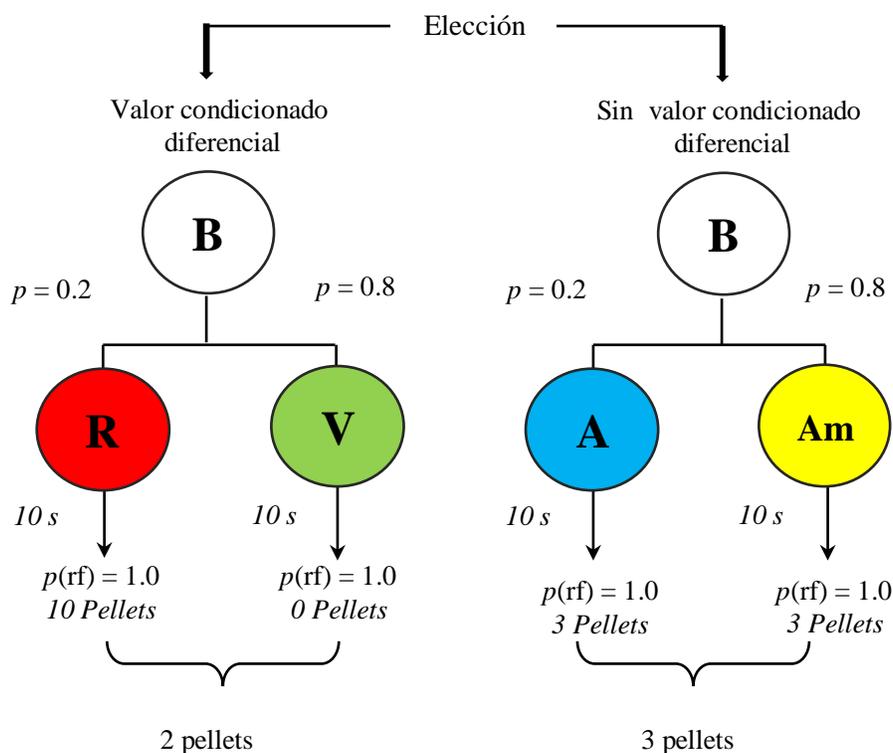


Figura 3 Procedimiento utilizado por Zentall y Stagner (2011).

Si las palomas son sensibles a la magnitud de los reforzadores que obtienen en cada alternativa, se predice que deben elegir la alternativa que en promedio da 3 pellets por ensayo. Sin embargo, contrariamente a esta predicción, los autores reportaron que las palomas mostraron una fuerte preferencia del 87% por la alternativa con valor condicionado diferencial que entregaba en promedio 2 pellets por ensayo. Este resultado muestra que la preferencia subóptima no es el resultado de la aversión a la incertidumbre, y que los estímulos discriminativos en la alternativa Discriminativa desempeñan un papel fundamental (Laude, Beckmann, Daniels, & Zentall, 2014; T. R. Zentall & Stagner, 2011).

Zentall y Stagner (2011) propusieron que este procedimiento es análogo a la conducta de juego patológico debido a que los jugadores patológicos prefieren escoger una alternativa con una expectativa de una victoria bastante pequeña, y persisten en esa alternativa a pesar de perder su dinero. Haciendo el análogo de esta situación con el procedimiento de elección subóptima, como la expectativa de una victoria es bastante pequeña, cuando llegan a ganar generan grandes expectativas, contrario a cuando pierden dinero que genera nulas expectativas de la situación. Más específicamente, los jugadores patológicos tienden a sobrevalorar sus ganancias y subvalorar sus pérdidas considerablemente más frecuentes (Breen & Zuckerman, 1999; Kahneman & Tversky, 1979).

Con estos antecedentes diversos grupos de investigación han propuesto diferentes explicaciones para describir el comportamiento subóptimo presentado por las palomas en este procedimiento. En el siguiente capítulo se presentan los argumentos que dieron origen a estas hipótesis que explican el comportamiento de las palomas en este procedimiento.

CAPÍTULO II

**HIPÓTESIS QUE EXPLICAN LA "ELECCIÓN SUBÓPTIMA" EN
PALOMAS**

EFECTO DE CONTRASTE

Stagner y Zentall (2010) interpretaron los resultados obtenidos con el procedimiento de elección subóptima en términos de contraste. Específicamente, en la alternativa Discriminativa, las palomas al elegir esta alternativa tienen una expectativa global de reforzamiento del 20%, cuando aparece el estímulo predictor de reforzamiento la expectativa aumenta abruptamente al 100% (un aumento del 80%) cuando se entrega el reforzador, en cambio, si aparece el estímulo predictor de la ausencia del reforzador, hay una pequeña disminución en la expectativa de recompensa (del 20 al 0%). Por otro lado, la expectativa en la opción No Discriminativa permanece sin cambios cuando se presentan los estímulos (las palomas inicialmente tienen una expectativa global de reforzamiento en esta alternativa del 50% y cuando se presenta cualquiera de los estímulos, este valor sigue siendo el mismo). A pesar de que el estímulo predictor de la ausencia de reforzamiento es mucho más frecuente que el estímulo predictor de reforzamiento, el alto contraste positivo producido por este último aumenta presumiblemente el valor de la alternativa Discriminativa. Esta interpretación también es consistente con el porcentaje de elección subóptima encontrada en Stagner y Zentall (2010), donde el contraste positivo fue del 80%, en comparación con la preferencia encontrada en Gipson et al. (2009) donde el contraste positivo fue solo del 50%.

Esta propuesta, sin embargo, no pudo explicar resultados posteriores. Por ejemplo, Stagner, Laude y Zentall (2012) dieron a las palomas la posibilidad de elegir entre dos alternativas, ambas con estímulos discriminativos. En una de las opciones, el estímulo predictor de reforzamiento se presentó el 20% de los ensayos y en la otra opción se presentó en el 50% de los ensayos. Si el contraste explica la fuerte preferencia de las palomas en este procedimiento deberían preferir la opción del 20% porque proporciona un mayor contraste

(80% frente al 50%). Contrariamente a esta predicción, los animales se mostraron indiferentes entre las alternativas.

Smith y Zentall (2016) describieron que la indiferencia mostrada por Stagner et al. se debe a que en ambas alternativas proporcionan un contraste positivo al tener presente los estímulos predictores de la entrega un reforzador. Para mostrar este efecto los autores presentaron una alternativa que presentaba el 50% de las veces un estímulo predictor de reforzador y el otro 50% un estímulo predictor de la ausencia del reforzador y una alternativa que presentaba un estímulo que el 100% de las veces predecía la entrega de un reforzador. Los autores reportaron que las palomas se mostraron indiferentes al igual que el estudio de Stagner et al. Los autores propusieron que la elección depende principalmente del valor predictivo de los estímulos asociados con las recompensas en cada alternativa (ambos 100% de reforzamiento) y no la frecuencia de presentación de estos estímulos (50% vs. 100%).

Posteriormente, Case y Zentall (2018), utilizando el procedimiento de Smith y Zentall, y reportaron que las palomas desarrollaron una preferencia por la alternativa Discriminativa después de un entrenamiento extenso (75 sesiones) y los resultados reportados por Stagner et al., y Smith y Zentall se debieron a las pocas sesiones que expusieron a las palomas en este procedimiento (20 y 30 sesiones respectivamente). Case y Zentall concluyeron que hay dos mecanismos responsables de la elección subóptima con este procedimiento. Primero, como propusieron Smith y Zentall, la elección depende del valor predictivo de los estímulos asociados con las recompensas en cada alternativa. En segundo lugar, la preferencia por la alternativa subóptima depende del contraste positivo entre la expectativa de la tasa global de reforzamiento y la expectativa una vez que se presenta el reforzador. Sin embargo, como descubrieron Case y Zentall, este segundo mecanismo se desarrolla más lentamente.

MODELO ECOLÓGICO

Vasconcelos, Montiero y Kacelnik (2015) utilizaron el procedimiento reportado por Stagner y Zentall (2010) presentando una alternativa asociada con dos estímulos discriminativos y una tasa global de reforzamiento del 20% y una alternativa sin esos estímulos discriminativos, pero con una tasa global de reforzamiento del 50% utilizando una especie distinta de ave, los starlings (estorninos) como sujetos y encontraron resultados similares a los previamente reportados con palomas. Además, cuando disminuyeron la frecuencia de aparición del estímulo que predice la presentación de un reforzador a un 5%, los starlings mostraron bajo estas condiciones preferencia por la alternativa Discriminativa, aunque la elección de esta alternativa representaba una reducción del 90% en la cantidad de alimentos que podrían haber obtenido si hubieran Elegido la alternativa No Discriminativa.

Vasconcelos et al (2015), describieron que los animales deben prestar atención a los estímulos asociados con reforzadores e ignorar o incluso evitar los estímulos asociados con la ausencia de reforzadores. Más específicamente, si se ve un estímulo asociado con la ausencia de reforzadores, el animal usa esta información para cambiar su patrón de búsqueda: se aleja de ese estímulo y busca otra fuente de alimento (Stephens, Brown, & Ydenberg, 2007; Stephens & Krebs, 1986). Debido a que el animal se aleja del estímulo, las propiedades de ese estímulo se ignoran y, por lo tanto, no afectan las decisiones futuras. Cuando, en una situación de laboratorio, los animales se ven obligados a permanecer en presencia de un estímulo que predice la ausencia de una recompensa, la información sobre la ausencia del reforzador no se puede utilizar para cambiar la situación actual. Sin embargo, dado que el mecanismo del animal para hacer frente a esta situación es el que se desarrolló en el entorno natural, en el laboratorio también intentan escapar de ese estímulo (alejarse de ese estímulo, por ejemplo).

Basándose en estas ideas Vasconcelos et al., propusieron el modelo ecológico de elección subóptima que formaliza esta noción utilizando como referencia la Regla del Presupuesto Energético (es decir, energía obtenida de los alimentos / tiempo dedicado a la búsqueda de alimentos) de la Teoría de Forrajeo Óptimo (Stephens y Krebs, 1986). En términos generales, cuando la alternativa Discriminativa presenta un estímulo predictor de reforzamiento, todas las demoras asociadas con los ensayos en los que no se entrega un reforzador no se incluyen en los cálculos de la tasa de ingesta de energía para esa alternativa, de manera muy similar a los animales que se alimentan en la naturaleza no incurrirían en los costos de oportunidad de esperar por alimentos que saben que no están disponibles. Eliminar estas demoras de la tasa de consumo de energía para la alternativa Discriminativa aumenta la tasa funcional de consumo de energía para esa alternativa. Debido a que la alternativa No Discriminativa no proporciona estímulos que predicen con certeza la entrega de un reforzador, las demoras asociadas con los ensayos en los que no se entregan reforzadores se incluyen en los cálculos de la tasa de ingesta de energía para esa alternativa, al igual que un animal que busca alimento sin señales discriminativas incurriría en estos costos de oportunidad.

Por lo tanto, dadas estas restricciones este modelo predice la preferencia por la alternativa Discriminativa porque la tasa funcional de consumo de energía para esa alternativa es en realidad más alta que la de la alternativa No Discriminativa una vez que los cálculos de la tasa de consumo de energía se han ajustado en función de la presencia o ausencia de los estímulos para cada alternativa. Por lo tanto, esta perspectiva argumenta que la elección subóptima refleja una preferencia evolucionada por las alternativas que brindan información sobre la presencia o ausencia de reforzadores.

HIPÓTESIS DE SEÑALES PARA BUENAS NOTICIAS (SIGN)

Dunn y Spetch (1990) describieron que la conducta subóptima mostrada por las palomas depende de dos fuentes de reforzamiento: La primer fuente de reforzamiento tiene que ver con los reforzadores primarios (comida) y la segunda fuente está determinada por el valor de los reforzadores condicionados. Basaron esta explicación inspirados en la Teoría de Reducción de la Demora propuesta por Fantino (1969) en la que se propone que la variable responsable de la elección es la reducción del tiempo esperado por el reforzador primario, por lo tanto, un estímulo adquirirá valor en la medida en que indica una reducción en el tiempo promedio de espera por un reforzador.

McDevitt et al., (2016) describieron que en el procedimiento de elección subóptima donde la tasa global de reforzamiento es del 50% para la alternativa Discriminativa y 100% para la alternativa No Discriminativa, al elegir la alternativa Discriminativa los estímulos predictores de presencia o ausencia de reforzador crean un contexto en el que el estímulo predictor de reforzamiento genera una mayor reducción de la demora de los reforzadores que la respuesta de elección en si misma (porque la respuesta de elección es seguida por un reforzador el 50% del tiempo, mientras que la respuesta en el estímulo Discriminativo es seguida por un reforzador el 100% del tiempo). Por otro lado, una respuesta de elección en la alternativa No Discriminativa siempre es seguida por un reforzador, el estímulo para esa alternativa no está asociado con una reducción de la demora del reforzador más allá de la reducción producida por la respuesta de elección en sí. Por lo tanto, el estímulo discriminativo predictor de un reforzador de la alternativa Discriminativa sirve como un reforzador condicionado más fuerte que el estímulo para la alternativa No Discriminativa, lo que resulta en el desarrollo de una preferencia por la alternativa subóptima. Es

importante señalar que esta hipótesis supone que el estímulo predictor de la ausencia de un reforzador no juega otro papel que el de crear un contexto de incertidumbre.

IMPACTO DEL ESTÍMULO PREDICTOR DE LA AUSENCIA DE REFORZAMIENTO

Las dos hipótesis descritas previamente tienen en común que no atribuyen ningún valor al estímulo predictor de la ausencia de reforzamiento en la preferencia de las palomas en este procedimiento. Sin embargo, ninguna de estas hipótesis muestra evidencia de este poco efecto. Aunque es claro que el estímulo asociado con el reforzador en la alternativa Discriminativa se convierte en un fuerte reforzador condicionado, no está claro por qué el estímulo asociado con la ausencia de reforzador no funciona como un inhibidor condicionado suficientemente fuerte para reducir la preferencia por esta alternativa en palomas.

Una posibilidad es que una vez que se identifica el estímulo negativo, las palomas se alejan de él, lo que limita su efectividad. Por ejemplo, Roberts (1972) ha demostrado que la efectividad de un estímulo discriminativo está directamente relacionada con la duración con la que se observa. De acuerdo con esta posibilidad, las palomas en Gipson et al., (2009) y Stagner y Zentall (2010), rara vez respondieron ante el estímulo predictor de no reforzamiento, mientras que en presencia del estímulo predictor de reforzador respondieron contantemente.

Para probar esta hipótesis, Stagner, Laude y Zentall (2011) utilizaron una luz general como estímulo predictor de no reforzamiento para evitar que las palomas se alejaran de este estímulo. Los autores reportaron que las palomas expuestas a la luz general como estímulo condicionado inhibitorio adquirieron de igual manera la preferencia por la opción subóptima, lo que descartó que el poco valor inhibitorio del estímulo condicionado se debiera a la evitación de éste.

Hearst, Besley y Farthing (1970) describieron un procedimiento que implica la presentación de un reforzador condicionado (S +), junto con la presentación de un inhibidor condicionado. Se encuentra evidencia de inhibición condicionada cuando la respuesta al S + disminuye cuando el S- se presenta junto con el S +. Laude et al., (2014) diseñaron un procedimiento en el cual utilizaron una línea vertical como estímulo predictor de no reforzador (S-) en lugar de un color y una tecla iluminada de rojo (S+) como estímulo predictor de reforzador. Posteriormente, presentaron un estímulo compuesto S+/S-, al inicio del experimento y después de la aparición de una elección subóptima. Los autores reportaron que, al inicio del entrenamiento, durante las sesiones en las que los sujetos muestran una preferencia por la alternativa óptima, la inhibición condicionada generada por el estímulo asociado a la no entrega de reforzador tiene una fuerte influencia en la elección del sujeto. Sin embargo, el efecto de la inhibición condicionada disminuye con el paso de las sesiones, correlacionando con un incremento en la preferencia por la alternativa subóptima (Ver Figura 4).

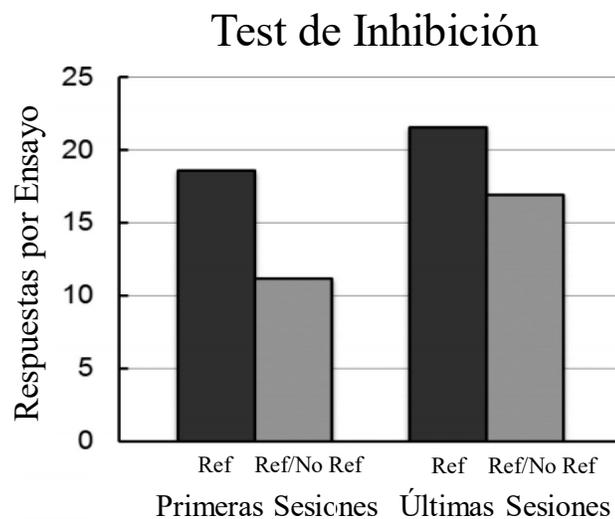


Figura 4 Respuestas al estímulo predictor de reforzamiento en las pruebas de inhibición en el estudio de Laude et al., (2014).

La ausencia de inhibición asociada con el estímulo que predice la ausencia de reforzamiento sugiere que el valor de la alternativa Discriminativa no está determinado por la tasa global de reforzamiento asociada con la elección de esa alternativa, sino que es la probabilidad esperada de reforzamiento asociada con el estímulo predictor de reforzamiento cuando este se presenta. Esto explicaría por qué las palomas prefieren un estímulo predictor de reforzamiento el 100% de las veces que se presenta (incluso si ocurre solo el 20% del tiempo) sobre un estímulo que predice la entrega de un reforzador el 50% de refuerzo (incluso si ocurre el 100% del tiempo).

Fortes, Vasconcelos y Machado (2016) probaron si variables como la probabilidad y la duración del estímulo predictor de no reforzamiento aumentaban el valor de este estímulo. Sin embargo, los autores reportaron que las palomas prefieren la alternativa subóptima incluso cuando se presentó este estímulo con una probabilidad de .95. También reportaron que la duración de este estímulo no es una variable que aumente su valor debido a que las palomas mantuvieron su preferencia a pesar de que la duración de este estímulo fuera de 200 segundos.

Debido a la evidencia presentada en los estudios anteriores, Smith, Bailey, Chow, Beckmann y Zentall (2016) modificaron el procedimiento clásico de elección subóptima y propusieron un diseño simplificado de este procedimiento que omitió la presentación del estímulo predictor de no reforzamiento. Específicamente, en la alternativa Discriminativa con probabilidad .5 se presentaba una tecla iluminada de rojo que predecía la entrega de un reforzador y con probabilidad .5 se presentaba un Blackout. Por el contrario, en la alternativa No Discriminativa siempre se presentaba una tecla verde que tenía una probabilidad de reforzamiento de .5 (ver figura 5).

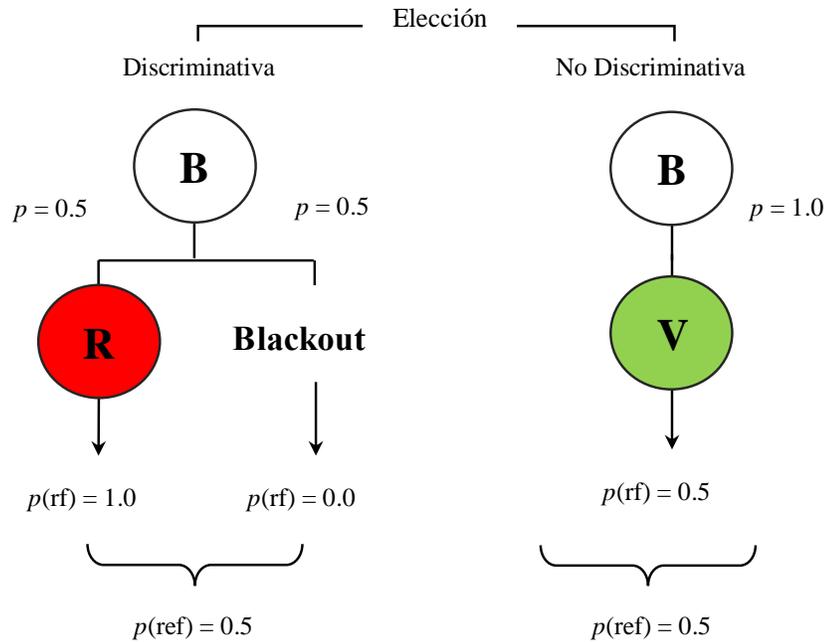


Figura 5 Procedimiento utilizado por Smith et al., (2016).

Los autores reportaron que con esta versión simplificada del procedimiento las palomas mostraron una fuerte preferencia por la alternativa Discriminativa que se mantuvo a pesar de reducir la frecuencia del estímulo predictor de reforzamiento en las fases posteriores al 25% y luego al 12.5%. Este resultado es importante, porque muestra que las palomas se comportan de manera subóptima en las dos versiones del procedimiento ya sea con 2 estímulos o 4 estímulos (Smith et al., 2016; Smith, Beckmann, & Zentall, 2017; T. R. Zentall, Andrews, & Case, 2017).

Las teorías para explicar los resultados encontrados con el procedimiento de “Elección Subóptima” en palomas parecían dar los argumentos suficientes para explicar este comportamiento, sin embargo, la generalidad de dicho procedimiento ha sido cuestionada por diversos trabajos cuyos resultados no encuentran la misma preferencia por la alternativa Discriminativa (subóptima) en otras especies.

CAPÍTULO III

**GENERALIDAD ENTRE ESPECIES EN EL PROCEDIMIENTO DE
"ELECCIÓN SUBÓPTIMA"**

El estudio de la conducta de elección subóptima se ha extendido a humanos, específicamente, Molet et al., (2012) desarrollaron una versión para humanos de la tarea de elección subóptima al incorporar las contingencias del procedimiento en un videojuego. La tarea consistió en un videojuego en el cual los sujetos tenían que elegir entre dos sistemas planetarios, cada uno de estos sistemas con dos planetas que se identificaban por un color diferente. Una vez seleccionado el sistema planetario los participantes debían mover el mouse sobre las naves que aparecían en cada planeta y hacer clic para disparar. Si los participantes elegían la alternativa Discriminativa, el 20% del tiempo fueron enviados a un planeta donde podrían obtener de 9 a 11 puntos, o el 80% del tiempo a un planeta en el cuál no podían obtener ningún punto. Si eligieron la alternativa No Discriminativa, siempre fueron enviados a uno de dos planetas donde podían obtener de 2 a 4 puntos. Por lo tanto, la elección de la alternativa Discriminativa proporcionó un promedio de 2 puntos por ensayo, mientras que la elección de la alternativa No Discriminativa les proporcionó un promedio de 3 puntos por ensayo (ver Figura 6). Los sujetos fueron instruidos para intentar obtener tantos puntos como pudieran.

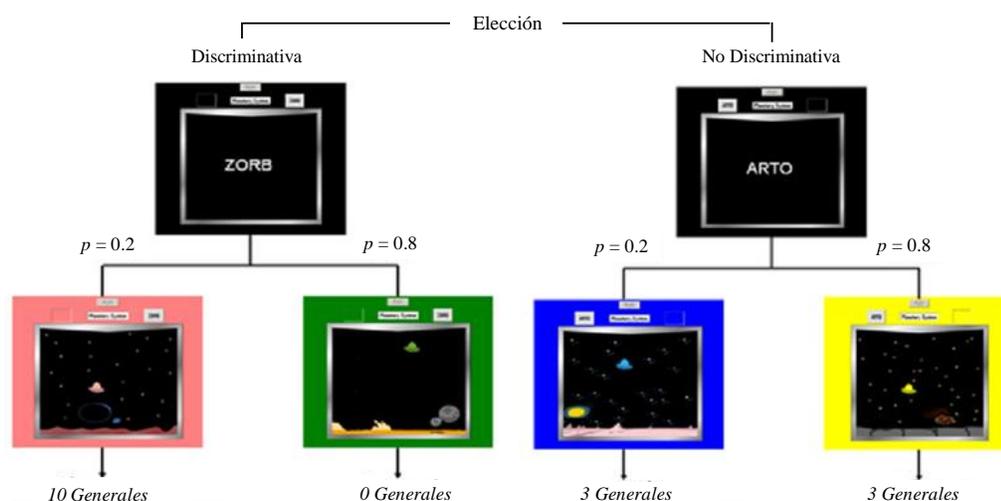


Figura 6 Procedimiento utilizado por Molet et al., (2012).

Además de realizar la tarea por computadora, los participantes debían llenar un autoreporte sobre la frecuencia con la cuál jugaban juegos de azar. Los autores reportaron que los participantes que informaron que participaban regularmente en juegos de azar eligieron la alternativa subóptima significativamente más que los no jugadores. Sin embargo, en comparación con la fuerte preferencia mostrada en palomas (T. R. Zentall & Stagner, 2011), los participantes que reportan que regularmente juegan juegos de azar muestran una preferencia moderada por la alternativa subóptima alrededor del 56.5%. En un estudio reciente, McDevitt, Diller y Pietrzykowski (2019) replicaron los resultados reportados por Molet et al., con humanos no jugadores. En ambos estudios, la proporción de elección por la alternativa Discriminativa fue aproximadamente el 25% de los ensayos de elección.

Dado que bajo condiciones específicas similares las palomas, starlings y los humanos muestran una preferencia por la alternativa subóptima, un paso necesario en la comprensión del fenómeno es evaluar su generalidad entre especies. Por esta razón, Trujano y Orduña (2015) llevaron a cabo con ratas una replicación sistemática de Stagner y Zentall (2010). Específicamente, si los sujetos elegían la alternativa Discriminativa, con .2 de probabilidad se encendía una luz roja después del cual siempre se entregaba un reforzador después de 10 s o con probabilidad .8 se encendía una luz azul, después de la cual nunca se entregaba reforzador. Por el contrario, si los sujetos elegían la alternativa No Discriminativa, se podía encender cualquiera de los dos estímulos (una luz roja o azul) con una probabilidad .5, ambos asociados con una probabilidad de reforzamiento de .5 (ver Figura 7).

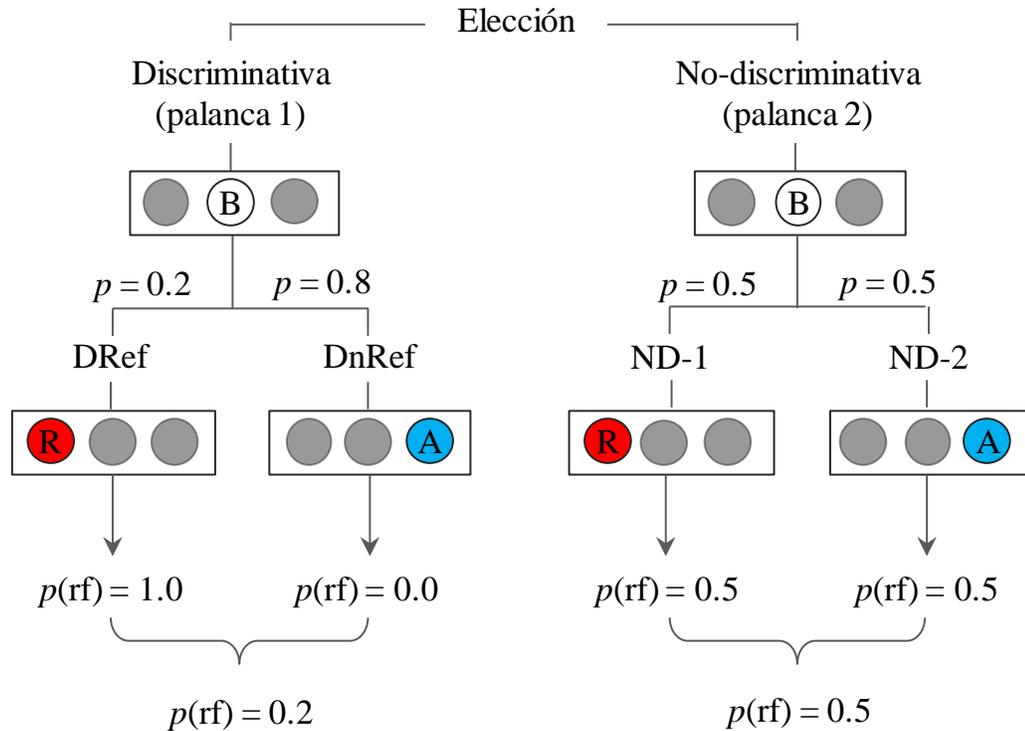


Figura 7 Procedimiento utilizado por Trujano and Orduña (2015).

Los autores reportaron que, contrario a lo reportado con palomas, las ratas tuvieron una preferencia por la alternativa óptima. Además de presentar la proporción de elección de la alternativa Discriminativa, adicionalmente, los autores presentaron el número de respuestas en la alternativa Discriminativa en presencia de los dos estímulos y mostraron que las ratas tuvieron un mayor número de respuestas en presencia del estímulo predictor de reforzamiento que en presencia del estímulo predictor de ausencia de reforzamiento. De esta manera, quedó descartada la posibilidad de que la conducta óptima de las ratas fuese debida a que no fueron sensibles a las consecuencias asociadas a los estímulos discriminativos. En otro trabajo presentado por Ojeda, Murphy y Kacelnik (2018), adaptaron el procedimiento de elección subóptima presentado por Trujano et al., utilizando estímulos auditivos debido a que esta modalidad de estímulos a mostrado ser precisa en la discriminación en ratas

(Heffner & Heffner, 1985). Específicamente, utilizaron diferentes estímulos para señalar las cuatro contingencias posibles: dos tonos para la alternativa Discriminativa (un tono predictor de entrega de un reforzador y un tono diferente predictor de la ausencia de reforzador) y dos tonos para la alternativa No Discriminativa (dos tonos, cada uno de ellos con un 50% de probabilidad de entregar un reforzador). Los autores replicaron el resultado reportado por Trujano et al., es decir, las ratas mostraron una fuerte preferencia por la alternativa No Discriminativa asociada a la tasa global de reforzamiento más alta.

Los resultados del conjunto de estudios descritos anteriormente cuestionan la generalidad entre especies de palomas y ratas en el procedimiento de “elección subóptima”. Específicamente, las palomas muestran una contundente preferencia por la alternativa Discriminativa y por lo tanto eligen la alternativa no óptima, en contraste, las ratas muestran conducta óptima al elegir la alternativa No Discriminativa, pero con una tasa global de reforzamiento más alta. Diversos grupos de investigación han argumentado diversas explicaciones a la conducta de elección subóptima en palomas y óptima en ratas llevado a cabo diversos experimentos para determinar la variable responsable de las diferencias entre especies. A continuación, se describen a detalle estos estudios.

CAPITULO IV

**VARIABLES RESPONSABLES DE LAS DIFERENCIAS ENTRE
ESPECIES EN EL PROCEDIMIENTO DE "ELECCIÓN
SUBÓPTIMA"**

INHIBICIÓN CONDICIONADA Y ELECCIÓN SUBÓPTIMA

Una de las variables que se ha propuesto como relevante en el desarrollo de la elección subóptima es la presencia de un inhibidor condicionado, es decir, el estímulo asociado con la no entrega de un reforzador en la alternativa Discriminativa (Laude, Stagner, et al., 2014). Un estímulo es considerado inhibidor condicionado cuando posee la cualidad de prevenir o disminuir la ocurrencia de una respuesta condicionada (Rescorla, 1969).

Para identificar las variables responsables de las diferencias entre especies en la tarea de elección subóptima, y dado que entre menor impacto tiene el inhibidor condicionado, mayor es la preferencia por la alternativa subóptima en palomas, Trujano, López, Rojas-Leguizamon, y Orduña (2016) evaluaron si la conducta óptima en ratas tenía relación con un fuerte impacto del inhibidor condicionado. Para llevar a cabo esta prueba los autores llevaron a cabo una tarea de sumación es decir presentaron los dos estímulos, uno asociado con la entrega de un reforzador y otro asociado con la ausencia de reforzamiento al inicio del experimento y volvieron a presentar esta prueba en las sesiones finales del experimento.

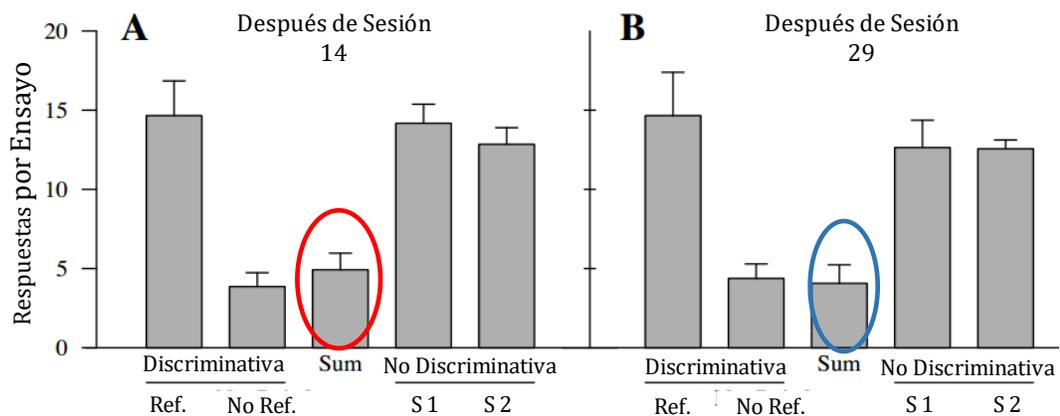


Figura 8 Respuestas al estímulo predictor de reforzamiento en las pruebas de inhibición en el estudio de Trujano et al., (2016).

Los autores reportaron que inhibidor condicionado en la tarea de elección subóptima ejercía mayor efecto en la preferencia de las ratas que el estímulo asociado con reforzamiento y este impacto se presentaba al inicio de la tarea y se mantenían sus efectos al final (ver figura 8).

Este resultado condujo a la hipótesis de que las palomas y las ratas difieren en la sensibilidad al inhibidor condicionado presente en la tarea de elección subóptima. En particular esta explicación sugiere que la diferencia entre especies está mediada por una respuesta diferencial a los estímulos predictores de no reforzamiento (inhibidores condicionados): mientras que las palomas parecen no considerar a este estímulo en su valoración de la alternativa Discriminativa (Laude, Stagner, et al., 2014), las ratas lo consideran con gran fuerza (Trujano et al., 2016), lo cual parece provocar la disminución de valor de la alternativa Discriminativa (subóptima).

No obstante, estas diferencias en cuanto a inhibición condicionada no son las únicas propuestas para explicar las diferencias entre especies en el procedimiento de elección subóptima. Siguiendo con este interés, Chow, Smith, Wilson, Zentall, y Beckmann (2017) sugieren que la conducta óptima reportada por Trujano y Orduña (2015) se debe a la relación que cada especie tiene con algunas propiedades de los estímulos discriminativos utilizados, especialmente su valor incentivo.

VALOR INCENTIVO Y ELECCIÓN SUBÓPTIMA

Cuando un estímulo neutro es contingentemente apareado con un estímulo incondicionado (EI), se convierte en un estímulo condicionado, es decir, adquiere la capacidad de producir una respuesta condicionada que normalmente favorece una reacción adaptativa ante la inminente aparición del EI. Además de esta función predictiva, algunos

estímulos, bajo ciertas circunstancias, activan estados motivacionales y emocionales complejos y se convierten en estímulos incentivos, que se caracterizan por presentar tres propiedades básicas: a) sesgan la atención hacia ellos y elicitan una aproximación hacia ellos, b) son valorados, en el sentido que un organismo trabajará por tener acceso a ellos (se convierten en reforzadores secundarios), y c) activan estados motivacionales relevantes a la obtención del reforzador primario asociado (Anselme, 2016). No todos los estímulos condicionados se convierten en incentivos; por ejemplo, cuando un tono predice la aparición de comida, organismos de diversas especies muestran el aprendizaje de la relación predictiva tono-comida mediante una aproximación hacia el sitio de aparición de la comida. Esta conducta recibe el nombre de “seguimiento de objetivos” y no está asociada al establecimiento de valor incentivo del tono. Sin embargo, cuando otros estímulos, por ejemplo, la aparición de una palanca, se asocian de manera contingente con la presentación de comida, elicitan en algunos organismos una respuesta de aproximación hacia ellos que permiten suponer el establecimiento de valor incentivo. A dicha conducta de aproximación se le conoce como “seguimiento de señales” y es muy común encontrarla tanto en palomas que se aproximan y pican una tecla que predice reforzador, como en ratas que se aproximan, tocan y muerden una palanca que predice reforzador.

Chow et. al., (2017) proponen que las diferencias entre especies encontradas con la tarea de elección subóptima podrían deberse al valor incentivo de los estímulos utilizados en el procedimiento. Tanto en palomas como en ratas, los estímulos discriminativos utilizados en la tarea de elección subóptima son luces, los cuales elicitan conducta de seguimiento de señales en palomas (Brown & Jenkins, 1968) y elicitan conducta de seguimiento de objetivos en ratas (Cleland & Davey, 1983). Por lo tanto, el valor incentivo de los estímulos discriminativos utilizados en la tarea de elección subóptima es mayor para palomas que para

ratas, y posiblemente genera una mayor preferencia por la opción Discriminativa (subóptima) en palomas que en ratas.

En el estudio de Chow et. al., (2017), dos grupos de ratas fueron expuestos a la tarea simplificada del procedimiento clásico de elección subóptima, en la cual se manipuló el valor incentivo de los estímulos presentados. En un grupo se emplearon luces como estímulos discriminativos para elicitación de conducta de seguimiento de objetivos y en otro grupo se utilizaron palancas para elicitación de conducta de seguimiento de señales. Específicamente, para elegir entre la alternativa Discriminativa y la No Discriminativa, los sujetos debían responder en uno de los dos orificios para inserción de nariz (nosepoke) ubicados en el panel trasero. Elegir la alternativa Discriminativa resultaba con .25 de probabilidad la presentación en el panel frontal de un estímulo A asociado con la entrega de un reforzador o con .75 de probabilidad un blackout asociado con la no entrega de reforzador. Por otra parte, elegir la alternativa No Discriminativa resultaba en la presentación de un estímulo B en el panel frontal asociado con una probabilidad de reforzamiento de .5 (ver Figura 9).

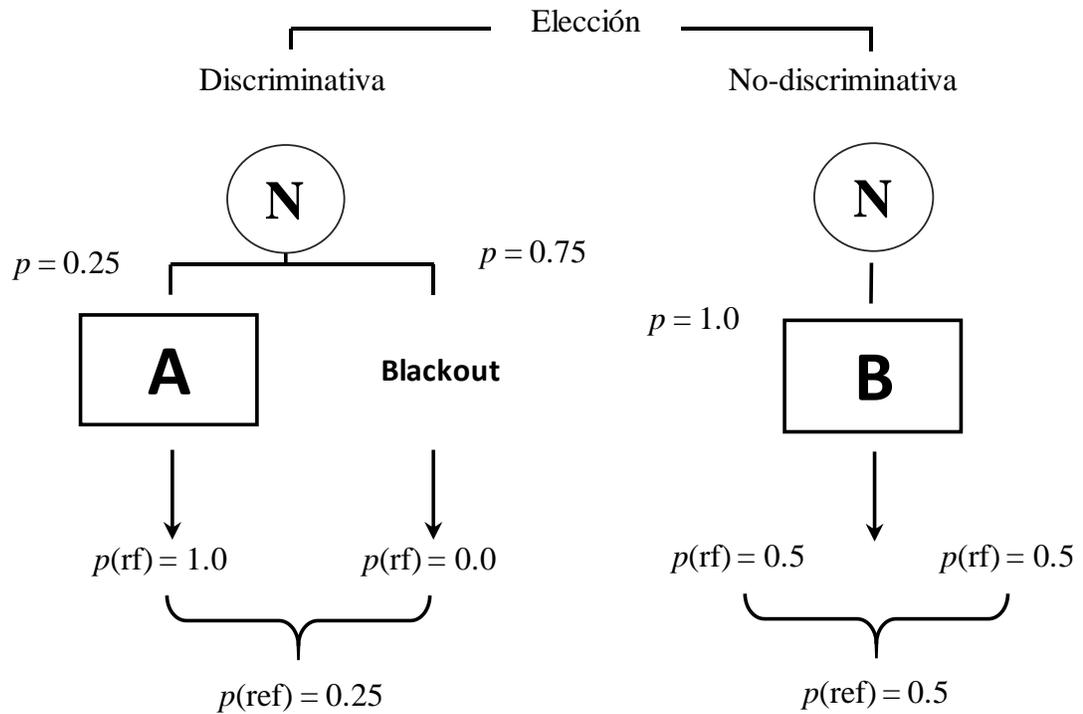


Figura 9 Procedimiento utilizado por Chow et al., (2017).

Los autores reportaron que cuando los estímulos predictores de reforzamiento fueron palancas (es decir, tuvieron valor incentivo), las ratas se comportaron de manera subóptima, mientras que cuando los estímulos discriminativos fueron luces, se replicó la conducta óptima previamente reportada en ratas (Trujano et al., 2016; Trujano & Orduña, 2015). Esta conclusión es de suma relevancia porque parece resolver el problema de las diferencias entre especies en el “procedimiento de elección subóptima”. Sin embargo, a pesar de que aparentemente dicho resultado apoya la hipótesis del valor incentivo, el procedimiento utilizado no empleó estímulos predictores de no reforzamiento, y en su lugar fue presentado un blackout. Dada la relevancia que presumiblemente tiene el inhibidor condicionado, es posible que su ausencia fue la que determinó la suboptimalidad, generando dudas sobre la

conclusión del estudio de Chow et al, es decir, que el valor incentivo de los estímulos discriminativos promovió la elección subóptima.

Posteriormente, Martínez, Alba, Rodríguez y Orduña (2017) utilizaron palancas como estímulos y agregaron al procedimiento una palanca que predecía confiablemente la ausencia del reforzador en la alternativa Discriminativa. Los autores modificaron una cámara operante estándar para tener dos paneles idénticos donde se presentaban los estímulos asociados a las alternativas Discriminativa y No Discriminativa respectivamente. Específicamente, responder en el nosepoke (panel frontal) asociado con la alternativa Discriminativa resultaba en la presentación con .5 de probabilidad de una palanca A asociada con la entrega de un reforzador o con .5 de probabilidad se presentaba una palanca B asociada con ausencia de reforzamiento. Por otro lado, al elegir la alternativa No Discriminativa podía presentarse con una probabilidad de .5 una palanca C y con una probabilidad de .5 la palanca D. Cualquiera que fuese la palanca presente la probabilidad de reforzamiento fue de .75. (ver Figura 10).

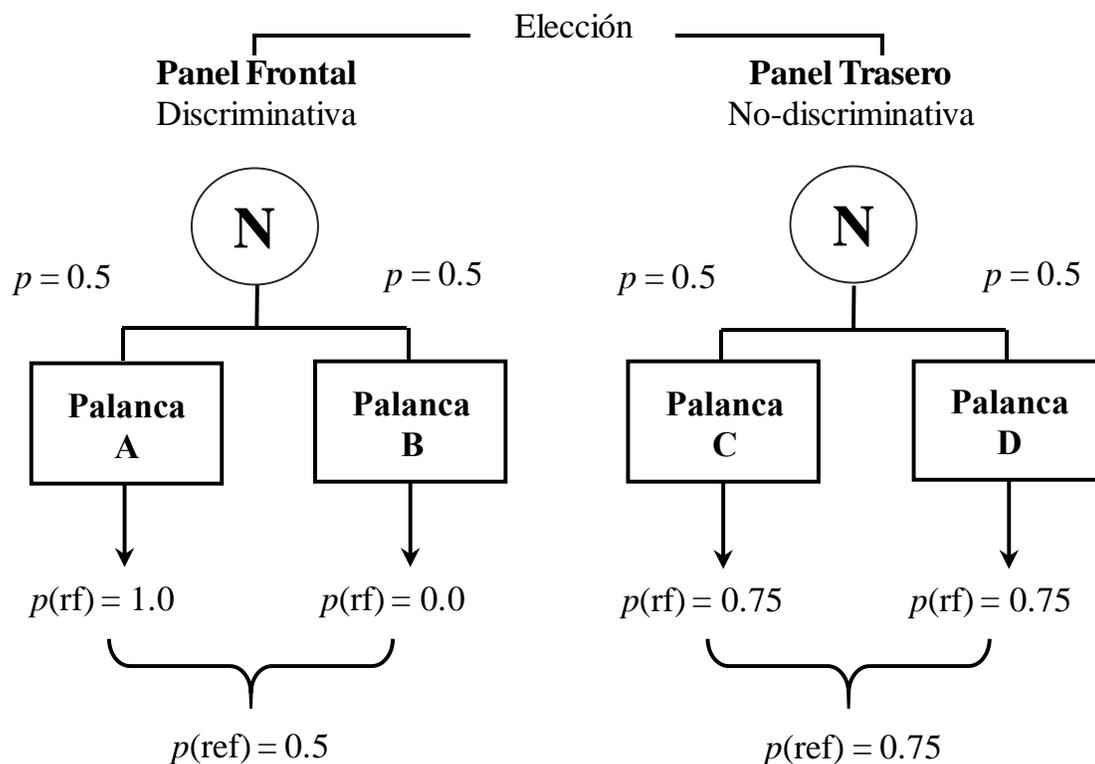


Figura 10 Procedimiento utilizado por Martínez, Alba, Rodríguez y Orduña (2017).

Los autores encontraron que las ratas mostraron una fuerte preferencia por la alternativa No Discriminativa a pesar de que los estímulos utilizados tenían valor incentivo, lo que sugiere que la relación entre el valor incentivo y el comportamiento subóptimo no es tan general como lo implicaron los resultados de Chow et al. Además, los autores reportaron que la palanca que predecía la ausencia de reforzamiento al igual que las luces utilizados en los primeros estudios con ratas funcionó como inhibidor condicionado y dado el tipo de estímulo que utilizaron adquirió un valor de incentivo (negativo). Es importante destacar que el estudio de Martínez et al., es el primero en integrar las dos hipótesis que se han propuesto para explicar las diferencias entre ratas y palomas en el procedimiento de “elección

subóptima”, y el resultado mostrado por estos autores parece descartar que el valor incentivo de los estímulos sea la explicación fundamental de la diferencia entre ratas y palomas.

DIFERENCIAS INDIVIDUALES CON LA CAPACIDAD DE ATRIBUIR VALOR Y ELECCIÓN SUBÓPTIMA

Además de que algunas características del estímulo influyen en la atribución de valor incentivo, existen importantes diferencias individuales que provocan que ante el mismo estímulo (por ejemplo, la aparición de una palanca) algunos animales dentro de la misma especie desarrollen seguimiento de señales y otros, seguimiento de objetivos, a pesar de haber sido entrenados bajo las mismas condiciones (Robinson, Yager, Cogan, & Saunders, 2014).

Específicamente, cuando se presenta una palanca de forma contingente 8 s antes de se entregue un reforzador, algunos sujetos muestran una aproximación hacia el sitio de aparición de la comida. Esta conducta recibe el nombre de “seguimiento de objetivos”. En cambio, en la misma situación, cuando se presenta la palanca algunos organismos muestran una respuesta de aproximación hacia ella que permiten suponer el establecimiento de valor incentivo. A dicha conducta de aproximación se le conoce como “seguimiento de señales” y es muy común encontrarla en ratas que se aproximan, tocan y muerden una palanca que predice reforzador. El seguimiento de señales es tan fuerte, que los animales lo realizan a pesar de que ello implique la pérdida del reforzador primario (Beckmann & Chow, 2015).

Estas diferencias individuales han estimulado mucha investigación que ha demostrado que el fenotipo de seguimiento de señales es más propenso a desarrollar conductas desadaptativas, por ejemplo, son más propensos a desarrollar conductas adictivas

(Robinson & Berridge, 2008; Yager & Robinson, 2015) y muestran más impulsividad (Lovic, Saunders, Yager, & Robinson, 2011).

En un estudio reciente, López, Alba y Orduña (2018), correlacionaron la ejecución de ratas seguidoras de señales y ratas seguidoras de objetivos en el procedimiento de “elección subóptima”, bajo la hipótesis de que los seguidores de señales al ser más propensos a desarrollar conductas desadaptativas mostrarían una preferencia por la alternativa Discriminativa. Específicamente, sometieron a 45 sujetos a un procedimiento de Aproximación Pavloviana Condicionada el cual ha sido ampliamente usado para categorizar a los sujetos en seguidores de señales o seguidores de objetivos (Ahrens, Singer, Fitzpatrick, Morrow, & Robinson, 2016; Pitchers et al., 2015). Este procedimiento consistió en una secuencia de ensayos en que se presentó durante 8 s una palanca de respuesta que fue seguida por la presentación de un reforzador. Esta secuencia se repitió durante 25 ensayos en cada sesión, durante 10 sesiones. Después de esta fase los autores calcularon las siguientes medidas para poder clasificar a los sujetos: a) la diferencia entre la probabilidad de responder a la palanca y la probabilidad de responder al comedero durante la presencia de la palanca, b) la diferencia entre el número de respuestas a la palanca y el número de respuestas a comedero, dividido entre las respuestas totales, y c) la diferencia entre la latencia de respuesta en la palanca y la latencia de entrada a comedero. Teniendo estos valores, los autores promediaron estos valores y los sujetos que obtuvieron valores entre 1.0 y 0.5 fueron clasificados como seguidores de señales, mientras que los que obtuvieron puntajes entre -1.0 y -.5 fueron clasificados como seguidores de objetivos; el resto se consideraron indiferentes (ver Figura 11). Después de entrenar a todos los sujetos en este procedimiento seleccionaron a los 8 sujetos con mayor seguimiento de señales y 8 sujetos con mayor seguimiento de objetivos y posteriormente fueron evaluados en el procedimiento de “elección subóptima”.

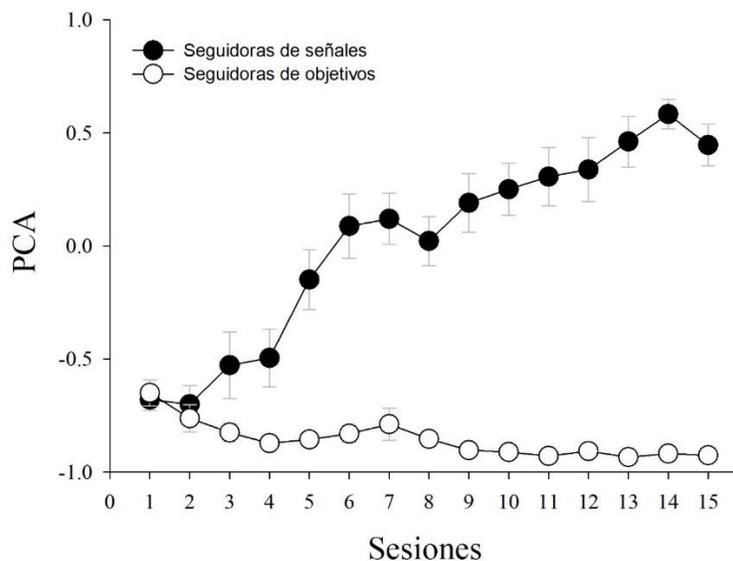


Figura 11 Promedio del índice de PCA del estudio de López, Alba y Orduña (2018).

Los autores reportaron que independientemente de la capacidad de atribuir valor incentivo (seguidor de señales o seguidor de objetivos) las ratas mostraron una fuerte preferencia por la alternativa No Discriminativa (óptima). Por lo tanto, la evidencia actual sugiere que el valor incentivo de los estímulos no es suficiente para desarrollar una preferencia por la alternativa subóptima en las ratas.

La supuesta relación entre la atribución de prominencia de los incentivos y la toma de decisiones desadaptativa es de suma importancia porque parece resolver un debate interesante que ha surgido en los últimos años en torno al "procedimiento de elección subóptima". Sin embargo, los resultados de los experimentos presentados en esta sección muestran que el efecto del valor de incentivo como una explicación de la elección subóptima en ratas no es general y las diferencias entre especies no se deben al valor incentivo de los estímulos.

Los experimentos revisados en este capítulo muestran una clara diferencia entre especies en el procedimiento de elección subóptima. La investigación llevada a cabo hasta el

momento sugiere que la explicación más contundente sobre la razón de estas diferencias parece ser una diferencia en la sensibilidad al estímulo predictor de no reforzamiento en ratas y palomas, particularmente, las ratas muestran una mayor sensibilidad a la inhibición condicionada y en el caso de las palomas esta inhibición se disipa con el transcurso de las sesiones. Por otra parte, la influencia del valor incentivo de los estímulos como una posible explicación de las diferencias entre ratas y palomas en este procedimiento no es del todo concluyente dada la evidencia mostrada hasta el momento.

CONCLUSIONES

Antes de aceptar la idea de que el inhibidor condicionado tiene un impacto mayor sobre la preferencia que el valor incentivo de los estímulos discriminativos, es crucial evaluar algunas diferencias importantes entre los estudios que exploraron de manera directa la hipótesis del valor incentivo. Debido a esto, la primera serie de experimentos de la presente tesis proponen evaluar la generalidad de la conducta óptima presentada por la ratas ante cambios en la probabilidad global de reforzamiento asociadas a cada una de las alternativas; en concreto, dado que el estudio de Chow et al (2017) empleó probabilidades de reforzamiento de .25 vs .5 en las alternativas Discriminativa y No Discriminativa, respectivamente, y los estudios de Martínez et al (2018) y López et al (2018) emplearon probabilidades de reforzamiento de .50 vs .75, es importante evaluar la posibilidad de que esta variable es la relevante para explicar la discrepancia en resultados, en el Experimento 1 exploramos esta posibilidad. Por otra parte, los estudios referidos también difieren en las características del estímulo predictor de no reforzamiento. Mientras que, en el estudio de Martínez et al., y López et al., fue una palanca que presumiblemente adquirió un valor de incentivo (negativo), en el estudio de Chow et al., se presentó un blackout. Por lo tanto, en el Experimento 2 se evaluará la generalidad de la conducta óptima encontrada en el estudio de

Martínez et al., pero en condiciones similares al procedimiento de Chow et al., respecto a las características del estímulo predictor de no reforzamiento. A pesar de que la evidencia presentada con palomas muestra no tener ningún impacto presentar solo un estímulo en la alternativa No Discriminativa, otra diferencia en los procedimientos está relacionada con las características de esta alternativa. Mientras que, en el estudio de Martínez et al., y López et al., se presentaron dos estímulos, en el estudio de Chow et al., solo se presentó un estímulo. En el tercer experimento se evaluará el posible impacto de este aspecto de los procedimientos.

Uno de los temas que más interés ha causado usando este procedimiento es su propuesta como un modelo que simula el juego patológico en humanos. En particular en este procedimiento la diferencia más relevante entre las alternativas no son las probabilidades globales de reforzamiento, si no la diferencia en las magnitudes de reforzamiento. Este procedimiento en particular ha sido empleado principalmente utilizando palomas como sujetos experimentales y no se ha evaluado la generalidad de este procedimiento en otras especies. Debido a esto, el cuarto experimento de esta tesis propone evaluar la generalidad de la conducta óptima de las ratas al utilizar el procedimiento de elección subóptima que ha sido considerado como un modelo animal que permite aproximarse al estudio de la adicción al juego.

La evaluación de la hipótesis del valor incentivo de los estímulos se ha llevado a cabo principalmente usando ratas como sujetos experimentales y a pesar de que la evidencia presentada hasta el momento parece descartar esta hipótesis, la evaluación no ha sido extendida en palomas, en particular, en todos los estudios de elección subóptima con palomas se han utilizado teclas iluminadas (estímulos con alto valor incentivo) y no se han utilizado estímulos con bajo valor incentivo. Por esta razón en el experimento 5 de la presente tesis se

emplearán tonos (estímulos con bajo valor incentivo) en el procedimiento de elección subóptima usando palomas como sujetos experimentales.

Todos los estudios que han manipulado el valor incentivo de los estímulos en el procedimiento de elección subóptima en ratas han mantenido el mismo valor en todos los estímulos y ante esta manipulación las ratas muestran una fuerte preferencia por la alternativa No Discriminativa (óptima). Una explicación de estos resultados es que el valor incentivo de los estímulos fue el mismo en ambas alternativas, por lo que los sujetos le atribuyeron el mismo valor tanto a la alternativa Discriminativa como a la No Discriminativa. En la segunda serie de experimentos de la presente tesis se evaluarán cambios sistemáticos del valor incentivo de los estímulos en el procedimiento de elección subóptima en ratas. En particular, en el experimento 6 se manipulará el valor incentivo de los estímulos de las alternativas Discriminativa y No Discriminativa presentando palancas (estímulos con alto valor de incentivo) en una de ellas y luces (estímulos con bajo valor de incentivo) en la otra, para evaluar su impacto en la elección subóptima. Adicionalmente, en el Experimento 7 se evaluará si las ratas clasificadas como *Seguidoras de señales* muestran una mayor tendencia a la elección subóptima en el "procedimiento de elección subóptima" con esta manipulación.

Los resultados de estos experimentos permitirán obtener evidencia de la generalidad de la optimalidad en ratas y la suboptimalidad en palomas en el procedimiento de elección subóptima. En caso contrario, se aportará evidencia a la hipótesis del valor incentivo de los estímulos como una variable que explica las diferencias entre especies en este procedimiento de elección subóptima. Adicionalmente, los resultados de los experimentos donde se manipulará sistemáticamente el valor incentivo de los estímulos permitirán evaluar de manera global el impacto de esta variable para explicar la conducta óptima en ratas y subóptima en palomas.

CAPITULO V

GENERALIDAD DE LA CONDUCTA ÓPTIMA EN RATAS EN EL PROCEDIMIENTO DE “ELECCIÓN SUBÓPTIMA”

Elección subóptima ocurre cuando los sujetos prefieren una alternativa con estímulos discriminativos que proporciona una menor probabilidad de reforzamiento, que otra alternativa sin esos estímulos, pero con una mayor probabilidad de reforzamiento (McDevitt et al., 2016). Se ha reportado que las palomas muestran consistentemente conducta subóptima (T. R. Zentall, 2016), en contraste, las ratas muestran conducta óptima (López et al., 2018; Martínez et al., 2017; Ojeda et al., 2018; Trujano et al., 2016; Trujano & Orduña, 2015). La única excepción a este resultado fue presentada por Chow, Smith, Wilson, Zentall, y Beckmann (2017) quienes emplearon la versión simplificada del procedimiento de elección subóptima y presentaron estímulos con alto valor incentivo (palancas) como estímulos discriminativos y encontraron que las ratas mostraron un comportamiento subóptimo. Sin embargo, el procedimiento utilizado en este experimento no empleó estímulos predictores de no reforzamiento, y en su lugar fue presentaron un blackout. Dada la relevancia que presumiblemente tiene el inhibidor condicionado en la preferencia en ratas en este procedimiento (Trujano et al., 2016), es posible que su ausencia fue la que determinó la suboptimalidad reportada en ratas en este estudio. Martínez, Alba, Rodríguez y Orduña (2017) utilizaron la versión clásica del procedimiento de elección subóptima y añadieron un inhibidor condicionado en la opción Discriminativa; emplearon los mismos estímulos con alto valor incentivo en el procedimiento de “elección subóptima”. Los autores encontraron que las ratas se comportaron óptimamente, resultado contrario a lo reportado por Chow et al.

Sin embargo, aunque el estudio de Martínez et al., se diseñó para evaluar la relevancia de la presencia del inhibidor condicionado, su procedimiento presentó otras características que podrían haber explicado parcialmente la diferencia en los resultados con el estudio de Chow et al. En particular, las probabilidades globales de reforzamiento

empleadas en Chow et al., fueron .25 en la alternativa Discriminativa y .5 en la alternativa No Discriminativa; en contraste, Martínez et al. (2017) emplearon probabilidades de .5 vs .75. Por lo tanto, existe la posibilidad de que las ratas no sean óptimas cuando se evalúan con la primera combinación de probabilidades. El apoyo a esta hipótesis proviene de estudios con palomas, en el que la preferencia por la alternativa subóptima ha sido mayor cuando las probabilidades de reforzamiento en las alternativas Discriminativas vs No Discriminativas fueron .2 vs .5 (Stagner & Zentall, 2010) que cuando fueron .5 vs .75. (Gipson et al., 2009). El primer objetivo del presente estudio es evaluar esta posibilidad.

Los estudios referidos también difieren el procedimiento utilizado de elección subóptima, Chow et al., utilizaron la versión simplificada del procedimiento (donde solo se presentan dos estímulos) y Martínez et al., utilizaron el procedimiento clásico (4 estímulos). Una variable importante en la cual difieren ambos estudios es en las características del estímulo predictor de no reforzamiento. Mientras que, en el estudio de Martínez et al., fue una palanca que presumiblemente adquirió un valor de incentivo (negativo), en el estudio de Chow et al., se presentó un blackout. Por lo tanto, en el segundo experimento de esta tesis se evaluó la generalidad de la conducta óptima encontrada en el estudio de Martínez et al., pero en condiciones similares al procedimiento de Chow et al., respecto a las características del estímulo predictor de no reforzamiento, pero utilizando la versión clásica del procedimiento de elección subóptima.

Aunque menos relevante teóricamente, otra diferencia en los procedimientos está relacionada con las características de la alternativa No Discriminativa. Mientras que, en el estudio de Martínez et al., se presentaron dos estímulos, en el estudio de Chow et al., solo se presentó un estímulo. El tercer experimento tiene como objetivo evaluar el posible impacto de este aspecto de los procedimientos.

Dentro de este contexto, se realizarán tres experimentos que tienen el objetivo principal de evaluar la generalidad del comportamiento óptimo en ratas en el "procedimiento de elección subóptima". En los tres experimentos se utilizarán palancas como estímulos, con el propósito de evaluar más a fondo la relación entre el valor incentivo y la elección subóptima.

EXPERIMENTO 1: ANÁLISIS DE LA RELACIÓN ENTRE LAS PROBABILIDADES .2 VS .5 Y EL VALOR INCENTIVO DE LOS ESTÍMULOS

Método

Sujetos

En el presente experimento se utilizaron ocho ratas macho de la cepa Wistar obtenidas del vivero del Instituto de Fisiología Celular de la UNAM, de aproximadamente 180 días de edad. Los sujetos tenían experiencia en este procedimiento de elección subóptima con el mismo tipo de estímulos discriminativos, pero no con las probabilidades de reforzamiento empleadas en el presente estudio (Martínez et al., 2017). Todas las ratas fueron alojadas de forma grupal en cajas de cuatro animales en la que tuvieron al agua ad libitum, y fueron mantenidas a una temperatura constante de $21^{\circ}\text{C} \pm 1^{\circ}\text{C}$, y en un ciclo de luz-oscuridad de 12:12 horas. Todas las ratas tuvieron un tiempo de acceso a la comida de 50 minutos diarios después de las sesiones experimentales. El cuidado de los sujetos se realizó en conformidad con la norma NOM-062-ZOO-1999 “Especificaciones técnicas para la producción, cuidado y uso de los animales de laboratorio”.

Aparatos

Se ocuparon cuatro cajas de condicionamiento operante modificadas de 32cm (largo) x 25.5cm (ancho) x 33.5cm (alto), situadas dentro de cajas sonoamortiguadas. Las paredes laterales fueron de plástico transparentes y el suelo fue una rejilla formada por 19 barras de acero inoxidable de .48cm de diámetro. Como pared trasera y pared frontal, se colocaron dos paneles operativos idénticos (ver Figura 12).

Cada panel contó con un receptáculo de pellets (MED Associates, Inc, Model ENV-200R2M) ubicado a 1.5cm del suelo y a 9.3cm de las paredes laterales, en el cual se

entregaron pellets de 45mg (BioServ, Product F0165) desde un módulo dispensador circular (MED Associates, Inc., Model ENV-203). A 3cm sobre el receptáculo y a 11cm de las paredes laterales, se colocó un orificio para una respuesta de inserción de nariz (nosepoke) de 2.5cm de diámetro. Para medir las entradas al receptáculo de pellets y las respuestas nosepoke, se contó con fotorreceptores tanto en el receptáculo de pellets como en el orificio para nosepoke. Finalmente, a 3.3cm del receptáculo de pellets y a 5.7cm del suelo, se colocaron dos palancas retráctiles de 5cm de largo (MED Associates, Inc., Model ENV-112CM).

En los orificios para nosepoke, se colocaron un par de focos color amarillo que funcionaron luz en el orificio para nosepoke, respectivamente. La caja contó con un extractor de aire que proporcionaba ventilación y servía como ruido blanco para enmascarar los ruidos del exterior.



Figura 12 Fotografía de caja operante modificada.

Para operar el equipo se utilizó una computadora personal conectada a una interfase MED-PC (Med Associates, St. Albans, VT.).

Procedimiento

En la tarea de elección subóptima, los sujetos fueron expuestos a un programa concurrente encadenado en el que tuvieron que elegir entre la opción Discriminativa, y la No Discriminativa, cuyas consecuencias fueron programadas en los diferentes paneles. Al grupo A se le presentó la alternativa Discriminativa en el panel frontal, y la alternativa No Discriminativa en el panel trasero, mientras que para el grupo B estas posiciones fueron invertidas.

Todos los sujetos fueron expuestos a 40 sesiones experimentales en esta condición. Cada sesión consistió en 60 ensayos divididos de la siguiente forma: 20 ensayos forzados para la alternativa Discriminativa, 20 ensayos forzados para la alternativa No Discriminativa y 20 ensayos de elección, todos presentados de manera aleatoria.

En cada ensayo forzado se iluminó únicamente un orificio de nosepoke, ya sea en el panel frontal o en el panel trasero, de tal manera que los sujetos solamente tenían la opción de responder en esa alternativa y, por lo tanto, servían para familiarizarse con las contingencias de reforzamiento en ambas alternativas. Las respuestas en el orificio de nosepoke contrario fueron inefectivas.

En los ensayos de elección, ambas luces de los orificios de nosepoke estaban encendidas hasta que el sujeto respondía en uno de ellos. El elegir la alternativa Discriminativa resultaba en un programa de reforzamiento TF10 s en el panel correspondiente, durante el cual podía presentarse con una probabilidad de .2 la palanca izquierda (Estímulo A) y con con una probabilidad de .8 la palanca derecha (Estímulo B).

Al concluir el periodo de 10 s en presencia de la palanca izquierda, siempre se entregaba un pellet como reforzador, mientras que al concluirlo en presencia de la palanca derecha nunca se entregaba reforzador. Así, la probabilidad de reforzamiento en esta alternativa fue de .2.

El elegir la alternativa No Discriminativa, por otra parte, resultaba en un TF10 s en el panel contrario, durante el cual podía presentarse con una probabilidad de .2 la palanca izquierda (Estímulo C) y con una probabilidad de .8 la palanca derecha (Estímulo D). Cualquiera que fuese la palanca presente durante el TF10 s, la probabilidad de reforzamiento fue de .5. Entre cada ensayo hubo un intervalo de 10 s en el que ningún estímulo fue presentado (ver Figura 13).

Elección Subóptima: Reversión.

Para descartar que la preferencia de los sujetos estuviese dada por sesgo hacia una posición y no por las consecuencias asociadas a cada alternativa, se llevó a cabo una condición de reversión, en la cual se cambió la posición de los paneles correspondientes a la alternativa Discriminativa y a la No Discriminativa de la condición anterior. Por lo tanto, la alternativa Discriminativa se ubicó en el panel trasero para el primer grupo, y en el panel frontal para el segundo. Un cambio adicional fue cambiar la posición del estímulo predictor de reforzamiento de la alternativa Discriminativa, de manera que el reforzador sólo fue entregado una vez concluido el TF10 s en presencia de la palanca derecha del panel correspondiente. Por lo demás, la probabilidad de presentación de los estímulos y la probabilidad global de reforzamiento en ambas alternativas se mantuvo igual a la primera condición.

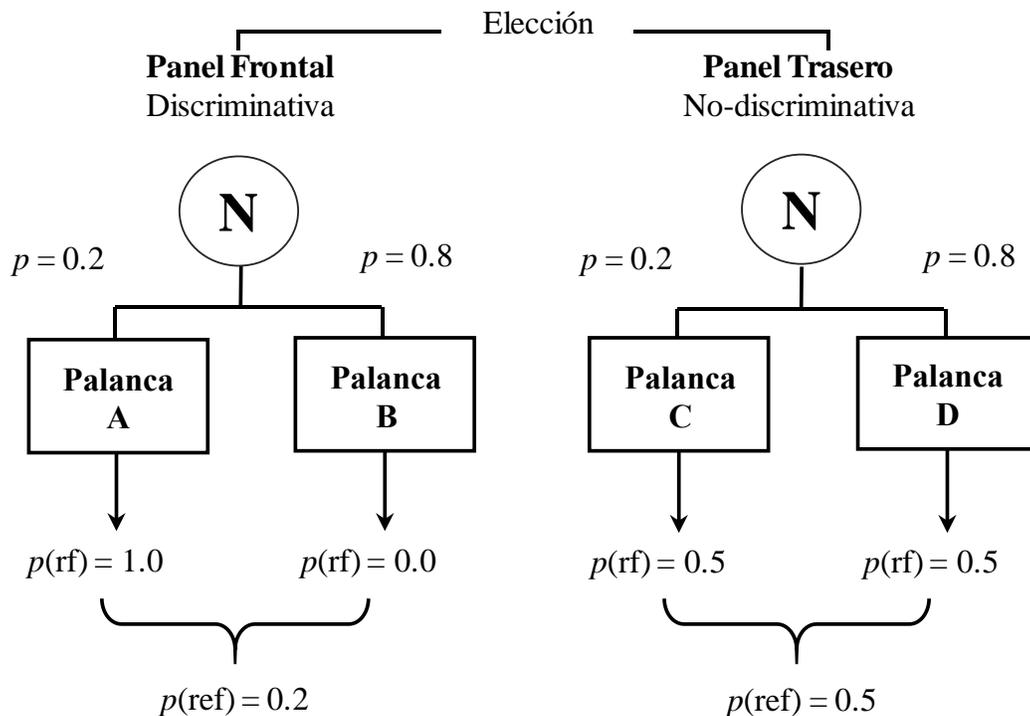


Figura 13 Diagrama del procedimiento usado en el presente experimento. Las dos alternativas Discriminativa y No Discriminativa se presentaron en paneles opuestos. Elegir la alternativa Discriminativa resultaba en la presentación de la palanca A con $p=.2$, durante 10 s después de los cuales se entregó un reforzador o se presentó la palanca B con $p=.8$, que predecía la ausencia de reforzamiento. Elegir la alternativa No Discriminativa resultó en la presentación de la palanca C con $p=.2$, o la palanca D con $p = 0.8$, ambas asociadas con una probabilidad de refuerzo de $.5$.

Resultados

Todos los análisis estuvieron basados en las últimas 5 sesiones de cada condición. Para cada sujeto fue calculada la proporción de elección de la alternativa Discriminativa en ambas condiciones, lo cual se tomó como índice de elección subóptima. Éste se obtuvo dividiendo el número de ensayos en los que el sujeto eligió la opción Discriminativa sobre el número total de elecciones durante la sesión.

La Fig. 14 muestra el promedio de la proporción de elección de la alternativa Discriminativa en los ensayos de elección. Al inicio de la condición el valor de dicha proporción estuvo alrededor de $.5$, es decir, los sujetos no mostraron preferencia por una

alternativa particular. Sin embargo, con el paso de las sesiones tal valor disminuyó de manera que, en las últimas cinco sesiones de la primera condición, los sujetos mostraron una clara preferencia por la alternativa óptima.

Durante las últimas cinco sesiones de la condición de entrenamiento, la proporción de elección para la alternativa Discriminativa fue de $.22 \pm .02$. Esta proporción fue significativamente diferente de indiferencia según una prueba t de dos colas ($t(7) = 14.16$, $p < .0001$).

En la primera sesión de la condición de reversión (ver Fig. 14), el promedio de la proporción de elección de la alternativa Discriminativa tuvo un valor alrededor del .7. No obstante, en las sesiones posteriores, se encontró que el valor de la proporción de elección recuperó los valores de la condición anterior, y durante las últimas cinco sesiones, la proporción de elección para la alternativa Discriminativa fue de $.23 \pm .03$, la cual en promedio también fue significativamente diferente de indiferencia ($t(7) = 10.52$, $p < .0001$). Con este resultado se descartó que la preferencia observada en la primera condición fuese resultado de un sesgo de posición. Adicionalmente, se llevaron a cabo pruebas Binomiales comparando la proporción de elección de las últimas cinco sesiones contra una probabilidad de elección de .5. Los resultados indicaron que tanto en la primera condición como en la segunda condición las ocho ratas mostraron una preferencia significativa para la alternativa óptima ($p's < .0001$ y $p's < .01$ entrenamiento y reversión, respectivamente).

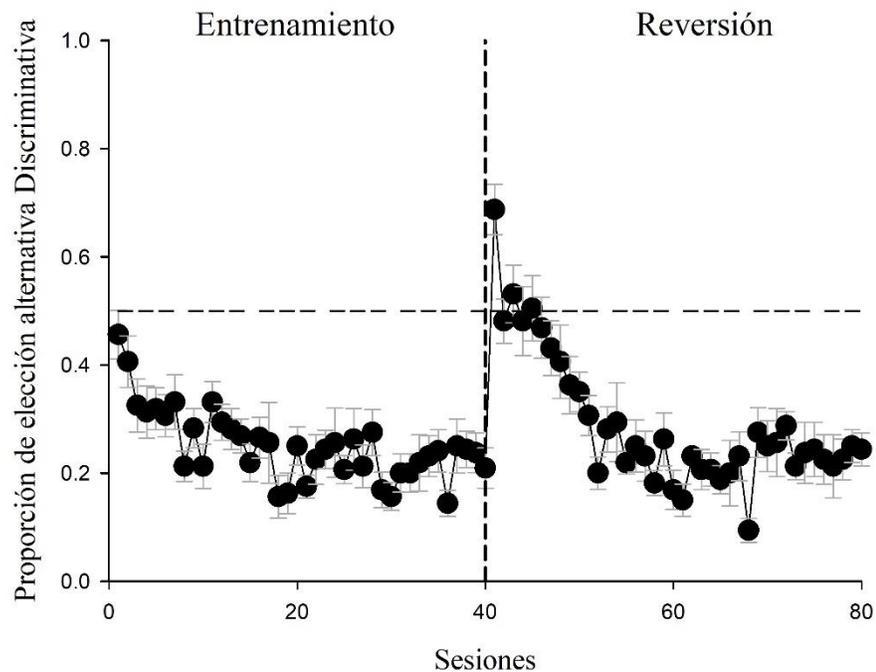


Figura 14. Experimento 1: Promedio de la proporción de elección de la alternativa Discriminativa a lo largo de las 40 sesiones de entrenamiento en el procedimiento de elección subóptima, y las 40 sesiones de la fase de reversión. La línea punteada horizontal señala la indiferencia entre ambas alternativas y la línea punteada vertical indica el inicio de la condición de reversión.

Con el propósito de evaluar si los sujetos fueron sensibles a las consecuencias programadas para los estímulos de la alternativa Discriminativa y No Discriminativa, se registró en ambas condiciones el número de respuestas en palanca y el número de entradas a comedero emitidas durante el TF10s de los ensayos forzados.

En la fig. 15 se puede observar que al inicio de la primera condición el promedio de respuestas fue similar ante los dos estímulos asociados a la alternativa Discriminativa. No obstante, con el paso de las sesiones, en promedio los sujetos respondieron más en presencia del estímulo predictor de reforzamiento que ante el estímulo predictor de la ausencia de reforzamiento. En las primeras sesiones de la condición de reversión, la tendencia general mostrada por los sujetos en cuanto a las respuestas en palanca de la alternativa discriminativa se invirtió. Sin embargo, con el paso de las sesiones nuevamente

mostraron un mayor número de respuestas ante el estímulo predictor de reforzamiento que ante el estímulo predictor de no reforzamiento.

Específicamente, para la alternativa Discriminativa, las respuestas en presencia del estímulo predictor de reforzamiento durante las últimas cinco sesiones en ambas condiciones (5.06 ± 1.45) fueron más altas que las respuestas a la palanca asociada con la ausencia de reforzamiento (0.27 ± 0.09). Con un ANOVA 2 (condición) x 2 (estímulos) de medidas repetidas utilizando el promedio de las últimas cinco sesiones, se confirmó que hubo diferencias estadísticamente significativas entre el número de respuestas en palanca ante el estímulo predictor de reforzamiento y ante el estímulo predictor de ausencia de reforzador ($F(1, 7) = 11.03, p < 0.05$). El efecto condición no fue significativo ($F(1, 7) = .004, p = 0.94$) y su interacción con los estímulos no fue significativo ($F(1, 7) = .16, p = 0.69$).

El mismo análisis se llevó a cabo para contrastar el número de respuestas en palanca ante los estímulos presentados en la alternativa No Discriminativa y, como se puede apreciar en la figura 15, se encontró que en promedio los sujetos no respondieron diferencialmente ante los estímulos asociados con dicha alternativa ($F(1, 7) = 2.40, p = 0.165$). El efecto por condición ($F(1, 7) = 3.10, p = 0.12$) y su interacción con los estímulos ($F(1, 7) = 2.41, p = 0.164$) tampoco fueron significativos.

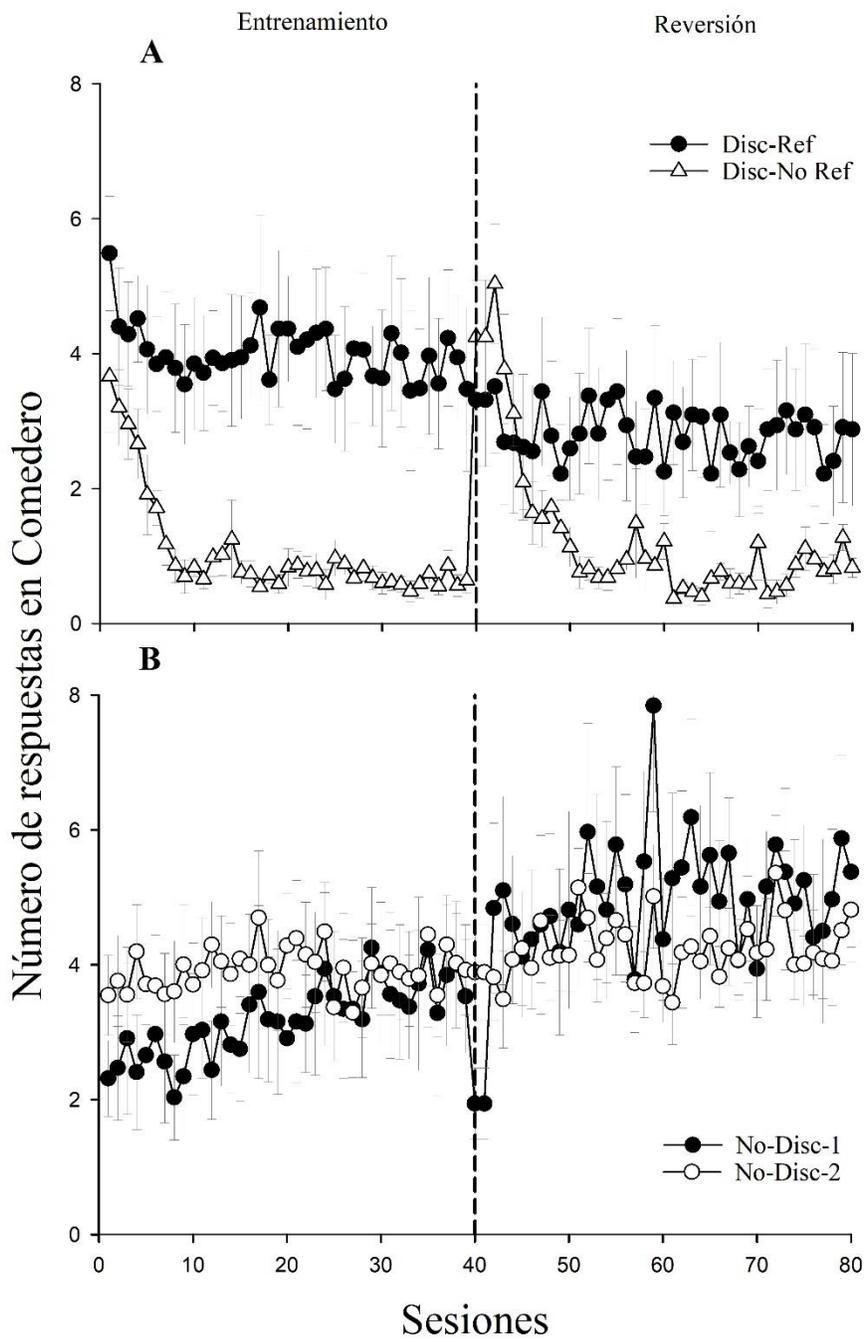


Figura 16. Experimento 1: Número de entradas al comedero en los ensayos forzados en todas las sesiones durante la fase de entrenamiento y reversión. El panel A muestra el promedio de entradas al comedero en la alternativa Discriminativa. Los símbolos cerrados representan el número de entradas al comedero en presencia del estímulo asociado con el reforzador (Disc-Ref) y los símbolos abiertos representan las entradas al comedero en presencia del estímulo asociado con la ausencia de reforzamiento (Disc-No Ref). En el panel B se muestra el promedio de entradas al comedero en la alternativa No Discriminativa. Los símbolos cerrados representan las entradas al comedero en presencia de la palanca C (No-Disc-1) y los símbolos abiertos representan las entradas al comedero en presencia de la palanca D (No-Disc-2).

En la figura 16 se presenta el número de entradas a comedero en presencia de los estímulos presentados en la alternativa Discriminativa (panel superior) y la alternativa No Discriminativa (panel inferior), durante todas las sesiones de entrenamiento y de la fase de reversión. Aunque se observó un mayor número de entradas al comedero en presencia de la palanca asociada con el reforzador (3.18 ± 0.89) que en presencia de la palanca asociada con la ausencia del reforzador (1.14 ± 0.10), especialmente durante la primera condición, esta diferencia no fue estadísticamente significativa, según un ANOVA realizado con datos de las últimas cinco sesiones de ambas condiciones (estímulos ($F(1, 7) = 5.34, p = .05$), condición ($F(1, 7) = 4.44, p = .07$) e interacción ($F(1, 7) = .95, p = .36$).

Con respecto a las entradas al comedero en presencia de los estímulos de la alternativa No Discriminativa, un ANOVA indicó que no había un efecto principal estadísticamente significativo del factor estímulos ($F(1, 7) = .003, p = .95$), pero los efectos de la condición ($F(1, 7) = 11.55, p = .01$), y de la interacción entre la condición y los estímulos ($F(1, 7) = 6.86, p = .03$) fueron significativos.

Discusión

El objetivo principal del presente experimento fue evaluar la posibilidad de que la conducta óptima mostrada por las ratas en el procedimiento de elección subóptima reportado en investigaciones previas (Martínez et al., 2017) sea el resultado de la relación particular entre las probabilidades globales de reforzamiento en las alternativas Discriminativa versus la alternativa No Discriminativa. Para lograr esto, las ratas se evaluaron en el procedimiento de “elección subóptima” usando las probabilidades globales de reforzamiento de .2 vs .5 en la alternativa Discriminativa frente a la No Discriminativa respectivamente, porque estas probabilidades en particular generaron el mayor nivel de suboptimidad en palomas (Stagner y Zentall 2010; Gipson et al., 2009). Los resultados

indicaron que las ratas se comportaron de manera óptima, eligiendo la alternativa asociada con la mayor probabilidad global de reforzamiento, tal vez incluso más que con la proporción de .5 vs .75 que se empleó en investigaciones previas (López, Alba y Orduña 2018; Martínez et al., 2017).

Dada la fuerte influencia que juega el estímulo predictor de no reforzamiento en el comportamiento de las ratas (Trujano et al., 2016), no es sorprendente que cuando la probabilidad global de reforzamiento en la alternativa Discriminativa es de .20 (y, por lo tanto, la probabilidad de presentar el estímulo predictor de no reforzamiento fue .8) se encontró una alta preferencia por la alternativa No discriminativa.

Las Figuras 14 y 15 (paneles superiores) muestran que los sujetos discriminaron la función de los estímulos asociados a la alternativa Discriminativa, realizando un mayor número de respuestas en presencia del estímulo que predijo el reforzador que en presencia del estímulo que predijo su ausencia. En contraste, el número de respuestas realizadas en presencia de los dos estímulos asociados con la alternativa No Discriminativa (figuras 14 y 15, paneles inferiores) fueron similares. Se observó un patrón inusual en el número de respuestas a las palancas durante la fase de reversión (figura 14, panel inferior). Aunque las respuestas a las palancas en la alternativa No Discriminativa no fueron estadísticamente significativas, se observó una diferencia en el número de respuestas en presencia de las palancas asociadas con esta alternativa. Es posible que las múltiples reversiones a las que fueron sometidas las ratas (como se señaló en la sección Método, los sujetos habían participado en otro experimento (Martínez et al., 2017)) favorecieron esta insensibilidad a las nuevas contingencias de reforzamiento, y facilitaron que los animales generalizaran la relación entre la ubicación espacial y la función predictiva que prevaleció en la alternativa Discriminativa, ubicada en el panel opuesto. Es importante señalar que esta transferencia

fue particular para el presente experimento, ya que no se ha observado en otros estudios con el mismo procedimiento (López, Alba y Orduña 2018; Martínez et al., 2017).

EXPERIMENTO 2: EVALUACIÓN DEL IMPACTO DE LA INHIBICIÓN CONDICIONADA Y EL VALOR INCENTIVO DE LOS ESTÍMULOS

Método

Sujetos

Se utilizaron ocho ratas macho Wistar como sujetos experimentales obtenidas del vivero del Instituto de Fisiología Celular de la UNAM. Todos los sujetos eran experimentalmente ingenuos, y de aproximadamente 60 días de edad. Los sujetos fueron alojados y mantenidos en las mismas condiciones que en el Experimento 1.

Aparatos

Los aparatos utilizados en este experimento fueron los mismos que en el Experimento 1.

Procedimiento

Habitación:

En esta primera fase, todos los sujetos fueron colocados dentro de la caja operante por 20 min. Durante ese periodo, tuvieron acceso a 30 pellets distribuidos equitativamente en el receptáculo trasero y en el receptáculo frontal. Una vez que los animales consumieron todos los pellets en ambos receptáculos, pasaron a la siguiente fase.

Entrenamiento a comedero:

En esta sesión los sujetos fueron expuestos a un programa de reforzamiento tiempo fijo (TF) de 60 s disponible aleatoriamente en el panel frontal o en el panel trasero, de manera que pudiesen asociar el sonido de la caída de pellet en cada receptáculo con la entrega de un reforzador. Cuando los animales consumieron 30 pellets en total (15 obtenidos en el panel frontal y 15 obtenidos en el panel trasero), se dio por concluido el entrenamiento a comedero. No hubo ninguna luz presente.

Entrenamiento a Nosepoke:

Las ratas fueron entrenadas a responder en el orificio de nosepoke utilizando un programa de razón fija (RF) 1 en el panel frontal o en el panel trasero de manera aleatoria. Para señalar en qué panel se encontraba disponible el programa de reforzamiento, al inicio del ensayo se encendía la luz del orificio del nosepoke correspondiente. Con la primera respuesta en el nosepoke, es decir, con la interrupción del haz infrarrojo de los fotorreceptores, la luz del orificio de nosepoke se apagaba y se entregaba un pellet como reforzador.

Cada sesión consistió en 50 ensayos en total, 25 ensayos para el nosepoke del panel trasero y 25 ensayos para el nosepoke del panel frontal. Entre cada ensayo hubo un intervalo entre ensayos (IEE) de 10 s, en los cuales la luz del nosepoke se mantuvo apagada. Esta fase fue concluida después de cinco sesiones, una vez que la latencia de los sujetos fue menor a 3 s y la diferencia de su latencia en el nosepoke frontal y el nosepoke trasero fue menor a 1.5 s.

Procedimiento

El procedimiento fue similar al empleado en el experimento anterior. La presentación de los ensayos forzados y de elección durante cada sesión experimental fueron presentados con las mismas contingencias descritas para el experimento 1. La principal diferencia fue que la palanca asociada con la ausencia del reforzador se sustituyó por un blackout. Además, las probabilidades globales de reforzamiento para las alternativas Discriminativas y No Discriminativas fueron .50 y .75, respectivamente. La secuencia de eventos asociados con cada alternativa fue la siguiente: Al emitir una respuesta al nosepoke en la alternativa Discriminativa con $p = .5$ se presentó la palanca izquierda A durante 10 segundos, después de la cual el reforzador siempre fue entregado; con $p = .5$ se presentó un blackout de 10 segundos y no se entregó ningún reforzador. Una sola respuesta al nosepoke en la alternativa No Discriminativa dio lugar a los siguientes eventos: la palanca izquierda C o la palanca derecha D se presentaron con $p = .5$ durante 10 s y cada palanca estuvo asociada con una probabilidad de reforzamiento de .75. (ver figura 17).

Después de 40 sesiones de entrenamiento bajo estas condiciones, durante la siguiente fase, se cambió la ubicación de las alternativas, con el fin de detectar cualquier posible sesgo de posición durante la fase previa.

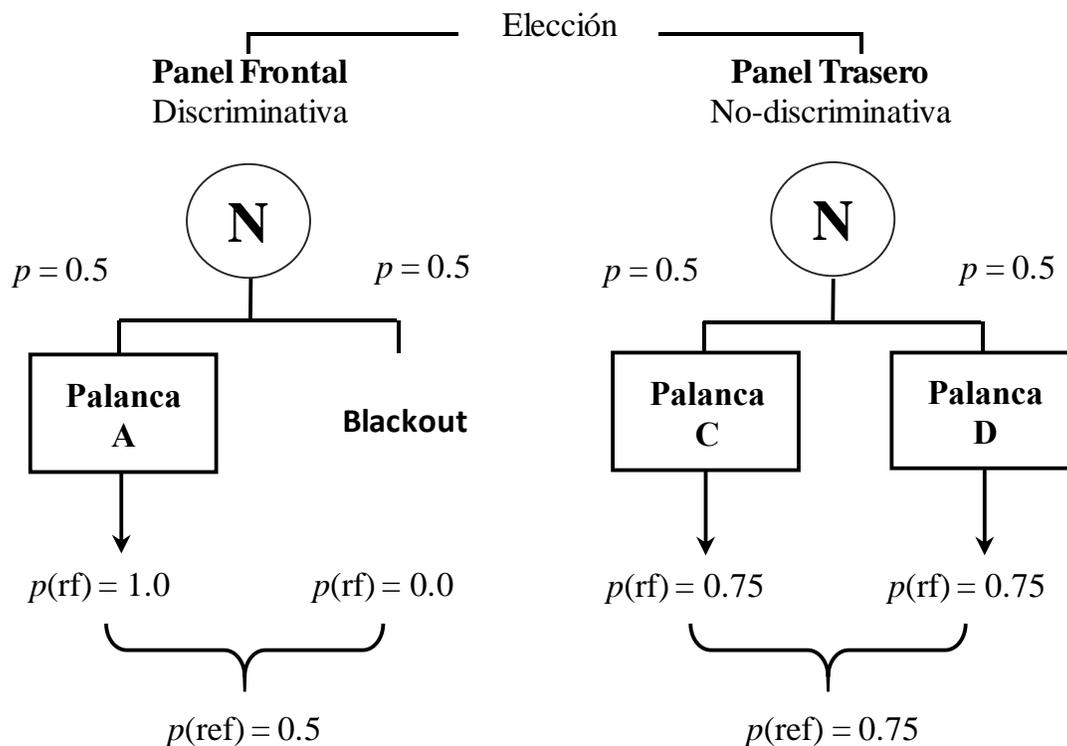


Figura 17. Diagrama del procedimiento usado en el presente experimento. Las dos alternativas Discriminativa y No Discriminativa se presentaron en paneles opuestos. Elegir la alternativa Discriminativa resultaba en la presentación de la palanca A con $p=0.5$, durante 10 s después de los cuales se entregó un reforzador o se presentó un Blackout con $p=0.8$, que predecía la ausencia de reforzamiento. Elegir la alternativa No Discriminativa resultó en la presentación de la palanca C o la palanca D con $p = 0.5$, ambas asociadas con una probabilidad de reforzo de $.75$.

Resultados

La Figura 18 muestra el promedio de la proporción de elección para la alternativa subóptima en las dos fases del experimento. Durante las últimas cinco sesiones de entrenamiento, las ratas mostraron preferencia por la alternativa No discriminativa (óptima); específicamente, el promedio de la proporción de elección para la alternativa Discriminativa fue $.26 \pm .04$, esta proporción fue estadísticamente diferente del azar según una prueba t de dos colas ($t(7) = 6.55, p < .0001$). Cuando las contingencias de reforzamiento se revirtieron, la preferencia por la alternativa No discriminativa (óptima) se recuperó rápidamente y el promedio de la proporción de elección durante las últimas cinco

sesiones para la alternativa Discriminativa fue de $.33 \pm .03$, que nuevamente fue significativamente inferior al azar ($t(7) = 5.85, p < .001$). Las pruebas Binomiales indicaron que la preferencia por la alternativa optima fue significativa para 7 de las 8 ratas en ambas condiciones del presente experimento ($p's < .0001$ y durante la reversión $p's \leq .01$). Un sujeto fue indiferente en cada condición ($p = .38$ y $p = .53$, respectivamente).

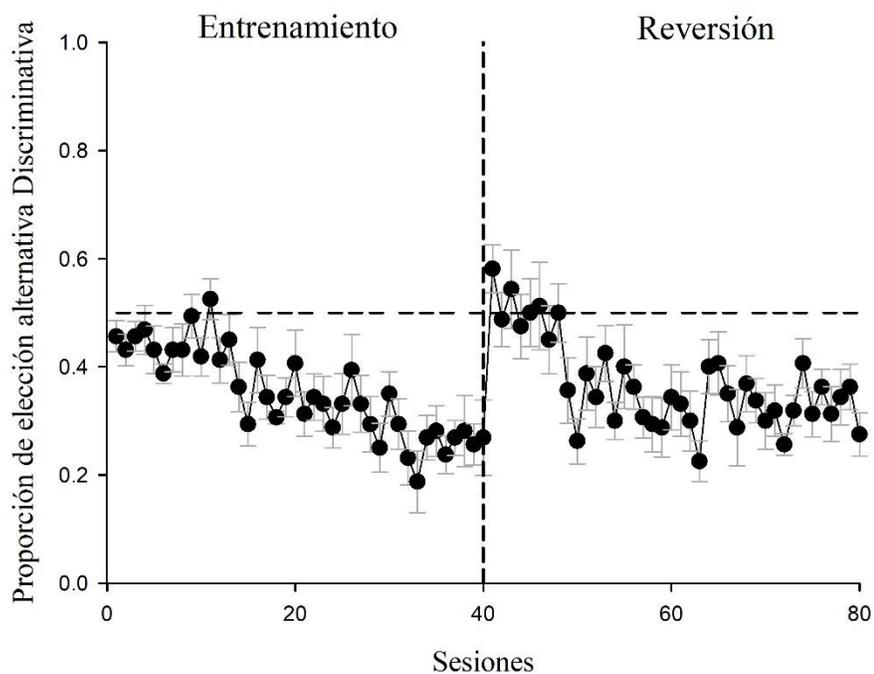


Figura 18. Experimento 2: Promedio de la proporción de elección de la alternativa Discriminativa a lo largo de las 40 sesiones de entrenamiento en el procedimiento de elección subóptima, y las 40 sesiones de la fase de reversión. La línea punteada horizontal señala la indiferencia entre ambas alternativas y la línea punteada vertical indica el inicio de la condición de reversión.

Discusión

En el presente experimento, el estímulo asociado con ausencia de reforzamiento fue sustituido por un blackout. Estas condiciones proporcionaron una aproximación más cercana al procedimiento empleado por Chow et al. (2017), el único estudio que reportó un comportamiento subóptimo en ratas. A pesar de esta manipulación, las ratas mostraron una preferencia por la alternativa óptima al mismo nivel que en el experimento 1.

Este resultado sugiere que la ausencia de un estímulo asociado explícitamente con la ausencia de reforzamiento no fue la variable relevante para generar un comportamiento subóptimo en el estudio de Chow et al. (2017). En el presente experimento, los resultados son compatibles con la idea de que el blackout adquirió propiedades aversivas, lo que disminuyó el valor de la alternativa Discriminativa. Sin embargo, es importante tener en cuenta que la relación entre las probabilidades globales de reforzamiento fueron diferentes a las utilizadas por Chow et al. (2017) por lo que existe la posibilidad que estas variables interactúen para promover la elección subóptima. Esta posibilidad se evaluará en el siguiente experimento, que es una aproximación aún más cercana al procedimiento empleado por Chow et al. (2017).

Un aspecto importante en la diferencia entre los procedimientos empleados en los dos estudios son las características de la alternativa No Discriminativa. En particular en el estudio de Chow et al., la alternativa No Discriminativa se asociaba con un solo estímulo en lugar de los dos presentados en los experimentos anteriores. Un arreglo similar se empleó en la primera demostración de elección subóptima en palomas (Kendall 1974), y se ha empleado recientemente en varios estudios con palomas que utilizan la versión simplificada del procedimiento de elección subóptima, y en todos estos estudios no se han reportado diferencias en la proporción de elección en comparación con el procedimiento tradicional

de 4 estímulos (Smith et al., 2016; Smith y Zentall 2016; Gipson et al., 2009). En el experimento 3, los sujetos serán expuestos a esta configuración del procedimiento y se evaluarán las probabilidades globales de reforzamiento utilizadas en el estudio de Chow et al., y Martínez et al., en la primera condición ($p = .50$ vs $p = .75$), y ($p = .20$ vs $p = .50$) en la segunda condición.

EXPERIMENTO 3: ANÁLISIS DEL IMPACTO DE LAS CARACTERÍSTICAS DE LA ALTERNATIVA NO DISCRIMINATIVA Y SU RELACIÓN CON LA PROBABILIDAD DE REFORZAMIENTO

Método

Sujetos

Se usaron cuatro ratas macho Wistar obtenidas del vivero del Instituto de Fisiología Celular de la UNAM. Todos los sujetos eran experimentalmente ingenuos, de aproximadamente 60 días de edad. Los sujetos fueron alojados y mantenidos en las mismas condiciones que en los experimentos previos.

Aparatos

Los aparatos utilizados en este experimento fueron los mismos que en los Experimentos anteriores.

Procedimiento

Antes del entrenamiento de elección subóptima, los sujetos fueron expuestos a las mismas condiciones de habituación, entrenamiento a comedero y entrenamiento al nosepoke con las mismas características que en los experimentos anteriores.

El procedimiento fue similar al empleado en los experimentos anteriores, a excepción de las características de la alternativa No Discriminativa (ver Figura 19). En la

condición 1 del presente experimento en los ensayos forzados en la alternativa Discriminativa, solo se iluminó un solo nosepoke, y al emitir una respuesta con $p = .5$ la palanca izquierda A se presentó durante 10 s, después de la cual se entregó un reforzador; con $p = .5$, se presentó un blackout durante 10 s seguido de ningún reforzador. Durante los ensayos forzados asignados a la alternativa No Discriminativa, una respuesta al nosepoke resultó en la presentación con $p = 1$ de la palanca B durante 10 s, después de la cual el reforzador fue entregado con $p = .75$. Bajo estas condiciones, las probabilidades generales de reforzamiento fueron .5 y .75 para las alternativa Discriminativa y No Discriminativa, respectivamente. Luego de 40 sesiones, se inició la condición 2 del presente experimento, que consistió en la modificación de las probabilidades globales de reforzamiento: en la alternativa Discriminativa se disminuyó de .5 a .2, mientras que para la alternativa No Discriminativa se disminuyó de .75 a .5. Adicionalmente, la posición de ambas alternativas se cambió al panel contrario de donde fueron presentadas. Finalmente, en la tercera condición, la posición de ambas alternativas nuevamente se cambió para evaluar un posible sesgo de posición.

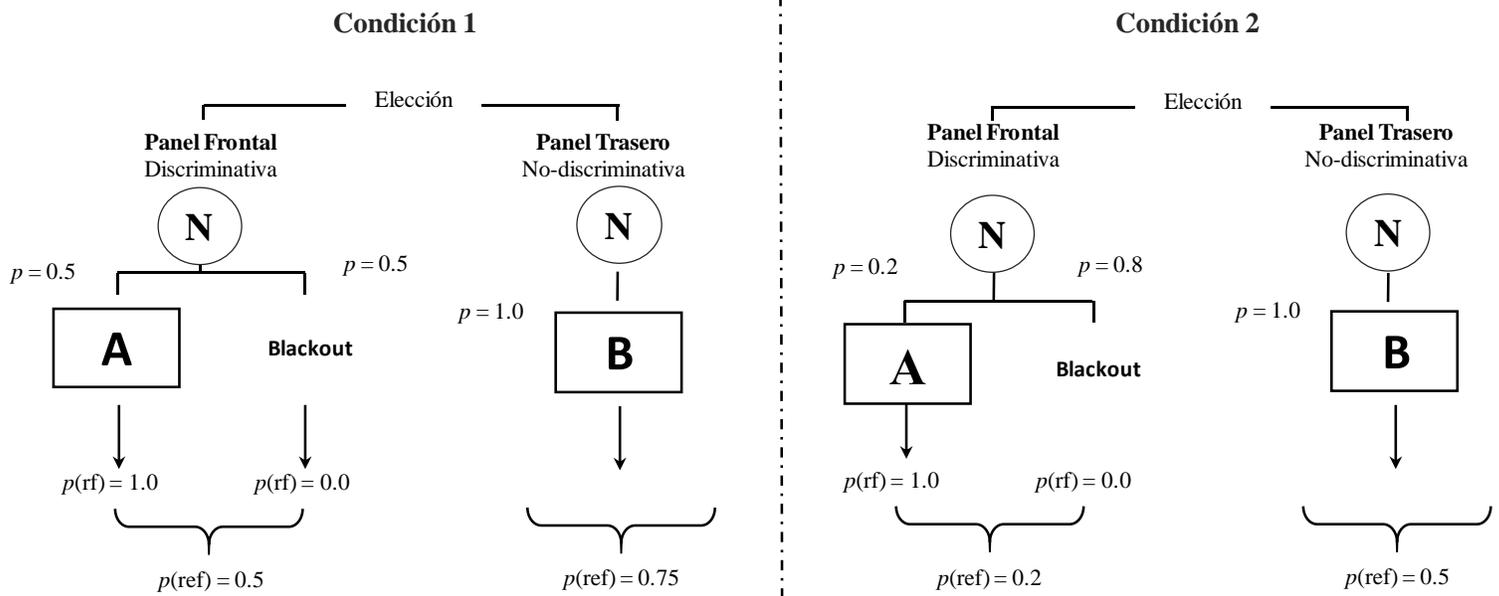


Figura 19. Diagrama del procedimiento usado durante las diferentes condiciones del presente experimento. Las dos alternativas Discriminativa y No Discriminativa se presentaron en paneles opuestos. En la condición 1 elegir la alternativa Discriminativa resultaba en la presentación de la palanca A con $p=.5$, durante 10 s después de los cuales se entregó un reforzador o se presentó un Blackout con $p=.5$, que predecía la ausencia de reforzamiento. Elegir la alternativa No Discriminativa resultó en la presentación de la palanca B en todos los ensayos, asociada con una probabilidad de reforzamiento de .75. En la condición 2 y 3 se manipuló la probabilidad global de reforzamiento, de tal forma que elegir alternativa Discriminativa resultaba en la presentación de la palanca A con $p=.2$, durante 10 s después de los cuales se entregó un reforzador o se presentó un Blackout con $p=.8$, que predecía la ausencia de reforzamiento. Elegir la alternativa No Discriminativa resultó en la presentación de la palanca B en todos los ensayos, asociada con una probabilidad de reforzamiento de .5.

Resultados

La figura 20 muestra el promedio de la proporción de elección de la alternativa Discriminativa durante las tres condiciones del presente experimento. Cuando las probabilidades de reforzamiento en las alternativas Discriminativa y No Discriminativa fueron .50 y .75, respectivamente, los sujetos mostraron una tendencia a preferir la alternativa óptima. Durante las últimas cinco sesiones de esta condición, el promedio de la proporción de elección para la alternativa Discriminativa fue de $.37 \pm .04$; sin embargo, una prueba t del promedio de la proporción de elección durante los últimos cinco días contra la

indiferencia no fue estadísticamente significativo ($t(3) = -3.12, p = 0.052$). Las pruebas Binomiales indicaron que las preferencias de dos sujetos fueron significativas hacia la optimalidad ambas con $p's < .001$, y dos sujetos fueron indiferentes ($p = .18$ y $.10$, respectivamente). Durante la segunda condición, en la que la combinación de probabilidades en las alternativas Discriminativa y No Discriminativa fueron $.20$ vs $.50$, el promedio de la proporción de elección para la alternativa Discriminativa disminuyó rápidamente y durante los últimos cinco días de la condición la proporción fue $.26 \pm .07$, que fue estadísticamente significativo por debajo de la indiferencia ($t(3) = 3.55, p < .05$), todos los sujetos mostraron una preferencia significativa por la alternativa optima de acuerdo a las pruebas Binomiales ($p's < .05$). Finalmente, durante las últimas cinco sesiones de la fase de reversión, la proporción de elección para la alternativa Discriminativa fue $.12 \pm .01$, y nuevamente fue estadísticamente significativa por debajo de la indiferencia ($t(3) = 25.28, p < .01$), las pruebas Binomiales indicaron que todos los sujetos tuvieron una preferencia significativa por la alternativa optima ($p's < .0001$).

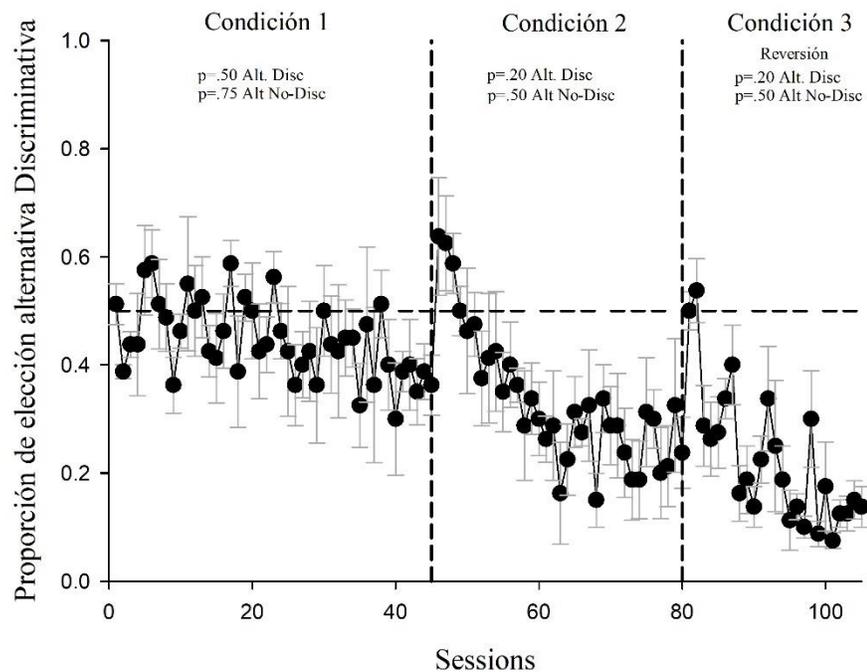


Figura 20. Experimento 3: Promedio de la proporción de elección de la alternativa Discriminativa durante las tres condiciones del presente experimento. La línea punteada vertical indica el inicio de una nueva condición; en la primera condición la probabilidad global de reforzamiento fue .50 para la alternativa Discriminativa y .75 para la alternativa No Discriminativa. En la condición 2 y 3 las probabilidades globales de reforzamiento fueron .20 y .50, respectivamente. La línea punteada horizontal señala la indiferencia entre ambas alternativas.

Discusión

El objetivo del presente experimento fue proporcionar una aproximación más cercana al procedimiento empleado por Chow et al. (2017). La fuerte preferencia por la alternativa No Discriminativa es consistente con la encontrada en los experimentos 1 y 2 de la presente tesis, y con estudios previos con ratas (López, Alba y Orduña 2018; Martínez et al., 2017; Trujano et al., 2016; Trujano y Orduña 2015), y sugiere que tener uno o dos estímulos asociados con la alternativa No Discriminativa no influye en el grado de elección óptima en ratas. Además, los resultados actuales respaldan las conclusiones del experimento 1, que sugiere que las probabilidades globales de reforzamiento de .25 vs .50

empleadas por Chow et al. (2017) no fue la variable responsable del comportamiento subóptimo que reportaron.

Discusión General

Una propuesta reciente para explicar la diferencia entre especies en el "procedimiento de elección subóptima" se basó en la idea de que los estímulos discriminativos empleados con palomas y ratas diferían en su valor de incentivo. Según esta hipótesis, las ratas mostrarían un comportamiento subóptimo cuando el procedimiento empleara estímulos discriminativos con alto valor incentivo. Aunque la primera evaluación de esta hipótesis la apoyó (Chow et al., 2017), dos experimentos recientes encontraron el resultado opuesto a pesar de que emplearon los mismos estímulos discriminativos con alto valor incentivo (López, Alba y Orduña 2018; Martínez et al., 2017). El objetivo principal del presente conjunto de experimentos fue proporcionar una evaluación adicional de esta hipótesis mediante el análisis de la generalidad del comportamiento óptimo de las ratas en el procedimiento de "elección subóptima". En concreto, se evaluó el impacto de tres variables que podrían ser relevantes para explicar la diferencia entre los resultados con palomas y ratas, y también con los resultados discrepantes de los estudios realizados con ratas.

Con respecto al primer punto, un análisis de la literatura con palomas (Gipson et al., 2009; Stagner y Zentall 2010) indica que la relación entre las probabilidades globales de reforzamiento en las alternativas Discriminativas frente a las No Discriminativas es un fuerte determinante de elección. Se ha reportado que las palomas muestran un mayor nivel de comportamiento subóptimo cuando esta relación es más extrema (es decir, cuando la pérdida en la tasa global de reforzamiento por preferir la alternativa Discriminativa es mayor), por ejemplo .20 vs .50 en la alternativa Discriminativa frente a las No

Discriminativa (Stagner y Zentall 2010), que cuando las tasas globales de reforzamiento son .50 vs .75 (Gipson et al., 2009). Dado que el único estudio que informó un comportamiento subóptimo en ratas empleó una tasa global de reforzamiento de .20 vs .50 (Chow et al., 2017), y que la mayoría de los estudios que informaron un comportamiento óptimo emplearon una tasa global menor (.50 vs .75; López, Alba y Orduña 2018; Martínez et al., 2017; Trujano et al., 2016), parecía razonable evaluar directamente el efecto de esta variable en condiciones donde los estímulos empleados en el procedimiento tuvieran alto valor de incentivo, similar al estudio que reportó la suboptimidad en ratas. Los resultados de los experimentos 1 y 3 coinciden en sugerir que el nivel del comportamiento óptimo no disminuyó y probablemente aumentó cuando se emplearon las probabilidades globales de reforzamiento de .2 vs .5. Aunque estas probabilidades ya se habían evaluado empleando ratas en el procedimiento clásico con luces como estímulos discriminativos (Trujano y Orduña 2015) y tonos (Ojeda et al., 2018), los resultados actuales son relevantes porque era posible que las probabilidades empleadas y el valor incentivo de los estímulos discriminativos interactuaran y promovieran la elección subóptima. Esta hipótesis no está respaldada por los resultados actuales, que también sugieren que la conducta óptima de las ratas es un fenómeno robusto.

A diferencia del experimento 1 en el que se demostró que los sujetos discriminaban la función de los dos estímulos de la alternativa Discriminativa, ya sea por un mayor número de respuestas a la palanca asociada con el reforzador o por un mayor número de entradas al comedero en presencia de esta palanca, en los experimentos 2 y 3 no fue posible realizar ninguno de estos análisis debido a que no había una palanca asociada con la ausencia del reforzador. Aunque el número de entradas al comedero podría en principio usarse para analizar dicha discriminación, la presencia de una palanca cuando se entregó un

reforzador, y su ausencia cuando no se entregó, generó una competencia entre las entradas al comedero y las respuestas a la palanca en el primer caso, pero no en el último, lo que hace que la diferencia en el número de entradas al comedero sea una medida no válida de discriminación. Sin embargo, los datos de estudios previos sugieren que es poco probable que los sujetos no sean sensibles a la función diferencial de estos estímulos, dada la discriminación alcanzada en experimentos anteriores (López, Alba y Orduña 2018; Martínez et al., 2017; Trujano et al., 2016; Trujano y Orduña 2015).

Martínez y cols. (2017) propusieron recientemente que la conducta óptima mostradas por las ratas en este procedimiento se ve favorecida por la presencia de inhibidores condicionados y propusieron que una posible explicación de la suboptimalidad encontrada por Chow et al. (2017) es que su procedimiento no empleo un inhibidor condicionado. Sin embargo, los resultados de los presentes experimentos (experimentos 2 y 3) argumentan en contra de tal explicación y sugieren que la presencia de un predictor de no reforzamiento, ya sean palancas con alto valor incentivo o un simple apagón, es suficiente para que las ratas disminuyan su preferencia de elegir la alternativa Discriminativa. Esto contrasta enormemente con los estudios con palomas y otras aves, en los que se ha demostrado que el predictor de no reforzamiento se ignora en el cálculo del valor de la alternativa Discriminativa (Smith et al., 2016; Vasconcelos et al., 2015). Es posible que la diferencia entre ratas y palomas en el procedimiento de "elección subóptima" esté mediada por esta sensibilidad diferencial al predictor de no reforzamiento.

Otra diferencia entre el estudio de Chow et al. (2017) y el conjunto actual de experimentos está relacionada con el número de sesiones realizadas: mientras que Chow et al. (2017) informaron un comportamiento subóptimo después de 14 sesiones (después de la exposición a aproximadamente 336 ensayos forzados para cada alternativa), todos los

experimentos actuales emplearon al menos 40 sesiones (800 ensayos forzados para cada alternativa). Como se puede observar en las figuras 14 y 18, el nivel de comportamiento óptimo después de 17 sesiones (cuando se había experimentado el mismo número de ensayos que en el estudio de Chow et al.) fue considerablemente menor que después de 40 sesiones. Es posible que las complejas contingencias de reforzamiento y el hecho de que los sujetos deben atender dos dimensiones de cada alternativa (discriminación y probabilidad de reforzamiento) contribuyan a retrasar la adquisición de un comportamiento óptimo. Además, se ha demostrado que la ausencia de un estímulo que indique un resultado retrasa la adquisición de un comportamiento estable; Vasconcelos et al. (2015) mostraron que cuando se omitió la señal de reforzamiento, la adquisición de un comportamiento subóptimo se retrasó considerablemente.

En resumen, todos los resultados obtenidos en el presente conjunto de experimentos son consistentes con la idea de que la diferencia fundamental entre ratas y palomas radica en su sensibilidad diferencial al estímulo predictor no reforzamiento, y que el valor incentivo de los estímulos discriminativos tiene un papel menos influyente.

CAPITULO VI

**GENERALIDAD DE LA CONDUCTA
OPTIMA EN RATAS EN EL PROCEDIMIENTO
DE ELECCIÓN SUBÓPTIMA PROPUESTO
COMO UN MODELO ANIMAL DE JUEGO
PATOLOGICO**

EXPERIMENTO 4: EVALUACIÓN DEL COMPORTAMIENTO EN RATAS EN EL PROCEDIMIENTO DE ELECCIÓN SUBÓPTIMA PROPUESTO COMO UN MODELO ANIMAL DE JUEGO PATOLOGICO

Además de la relevancia de este fenómeno para la comprensión de los mecanismos básicos de la toma de decisiones, el tema ha adquirido gran relevancia porque ha sido considerado como un modelo animal que permite aproximarse al estudio de la adicción al juego que es padecida por una proporción significativa de la población humana (Molet et al., 2012). La adicción al juego es un fenómeno complejo que puede conducir a graves consecuencias perjudiciales para un individuo en las esferas familiar, social y financiera (para una revisión, ver Raylu & Oei, 2002). Una gran cantidad de investigaciones ha identificado diversos factores relacionados con el juego, incluidos los factores sociales (Gupta & Derevensky, 1997) y neurobiológicos (Goudriaan, Oosterlaan, de Beurs, & Van den Brink, 2004; Potenza, 2008).

Una característica consistente en la adicción al juego es la tendencia a sobrevalorar la ocurrencia de una victoria y a subvalorar la ocurrencia (más frecuente) de una pérdida. Esta tendencia puede tener sus raíces en la disponibilidad heurística (Tversky & Kahneman, 1974), según la cual las victorias influyen más en el comportamiento que las pérdidas porque las primeras son más fáciles de recordar que las segundas. Esta asimetría entre la influencia de las ganancias y las pérdidas, a su vez, puede derivarse de un proceso pavloviano que asigna un valor diferencial a los estímulos de acuerdo con su relevancia y los resultados con los que están asociados; mientras que un valor fuerte y positivo (reforzador condicionado) se atribuye a los estímulos ambientales altamente salientes que acompañan a las ganancias, el valor negativo (inhibición condicionada) asignado a los

estímulos que están presentes cuando se experimentan pérdidas o se omiten las ganancias, probablemente porque esos los estímulos son menos sobresalientes (Wagenaar, 2016).

Como se describió en la introducción de la presente tesis, las palomas como los jugadores humanos, muestran una preferencia por la alternativa Discriminativa (subóptima). Además, como los jugadores humanos, las palomas parecen ignorar el estímulo asociado con ningún reforzador o pérdida (Fortes, Vasconcelos y Machado 2016; Laude, Stagner y Zentall 2014; McDevitt et al., 2016), por lo que el valor de la alternativa Discriminativa está determinado exclusivamente por el estímulo asociado con el reforzador (victorias). Debido a estas similitudes se ha sugerido que los mismos procesos relacionados con la función de los estímulos influyen en el comportamiento desadaptativo de las palomas y los humanos (Zentall 2011; Zentall y Stagner 2011), una explicación basada en la evolución que permite la posibilidad de que este patrón de comportamiento esté presente en otras especies. Sin embargo, la suboptimalidad mostrada por las palomas y otras aves no se ha replicado en ratas (Martínez et al., 2017; López, Alba y Orduna 2018; Trujano, López, et al., 2016; Ojeda, Murphy y Kacelnik 2018; Trujano y Orduña 2015; cf, Chow et al., 2017), cuestionando la amplia generalidad implicada por la suposición de un proceso común como causa de la similitud entre el comportamiento de las palomas y los humanos.

A pesar de esta ausencia de generalidad, las diferencias entre el comportamiento de las palomas y las ratas podrían proporcionar algunas ideas en la búsqueda de las variables que determinan la adicción al juego en los humanos. Por ejemplo, la diferencia informada en el comportamiento entre palomas y ratas en el "procedimiento de elección subóptima" es consistente con la idea de que estas especies difieren en su sensibilidad a los estímulos que predicen la ausencia de un reforzador: mientras que las palomas parecen ignorarlos (Laude, Stagner y Zentall 2014; McDevitt, Spetch y Dunn 1997), las ratas muestran una alta

sensibilidad a estos estímulos (Trujano, Lopez, et al., 2016). Es posible, entonces, que la sensibilidad a los inhibidores condicionados sea un predictor de la tendencia a jugar en humanos.

Recientemente, se ha propuesto (Zentall 2011) que el procedimiento clásico de elección subóptima no es el mejor análogo del juego patológico y que un mejor procedimiento debería considerar que en el juego patológico, la diferencia más relevante entre las alternativas (jugar versus no jugar) no son las probabilidades globales de reforzamiento, si no la diferencia en las magnitudes de reforzamiento. Zentall y Stagner (2011) desarrollaron un nuevo procedimiento que mantiene las características principales del procedimiento clásico, pero presenta dos modificaciones importantes: a) los diferentes estímulos asociados con la alternativa Discriminativa, predicen distintas magnitudes de reforzamiento, además de diferentes probabilidades, b) en alternativa No Discriminativa ambos estímulos predicen la misma magnitud, entregada con la misma probabilidad (1.0), por lo tanto, se elimina la incertidumbre asociada con esta alternativa y se puede descartar su influencia en la preferencia.

En este procedimiento, la elección de una alternativa (Valor condicionado diferencial) da como resultado, en el 20% de los ensayos, un estímulo asociado con la entrega de 10 pellets, y en el otro 80%, en un estímulo asociado con cero pellets (ganancias esperadas: dos pellets por ensayo). Elegir la otra alternativa (Sin valor condicionado diferencial) da como resultado la aparición de cualquiera de dos estímulos asociados con la entrega de tres pellets (ganancias esperadas: tres pellets por ensayo). Como se describió en la introducción, cuando se utilizó este procedimiento la mayoría de las palomas rápidamente desarrollaron una fuerte preferencia por la alternativa subóptima (Zentall y Stagner 2011; Laude et al., 2014; Laude, Stagner y Zentall 2014). Dado que la alternativa

preferida, además de tener valor condicionado diferencial, también presenta resultados variables Zentall y Stagner (2011) describieron que es necesario analizar si la variabilidad del resultado o el valor condicionado diferencial era la variable que determinaba la preferencia de las palomas. Por esta razón, en un segundo experimento, el valor condicionado diferencial se eliminó de la alternativa subóptima proporcionando 10 pellets con probabilidad .2 (y 0 pellets con $p = .2$) 10 s después de la presentación de cualquiera de los dos estímulos. Bajo estas condiciones, las palomas cambiaron su preferencia a la alternativa óptima, demostrando que era el valor condicionado diferencial es lo que determina su preferencia (Zentall y Stagner 2011, exp. 2).

Siguiendo nuestra idea de que la comparación entre especies podría proporcionar información útil sobre las variables que controlan el comportamiento de juego patológico en humanos, en el presente experimento evaluamos la generalidad de la conducta optima presentada por las ratas en esta versión del "procedimiento de elección subóptima".

Método

Sujetos

Se usaron ocho ratas macho Wistar obtenidas del vivero del Instituto de Fisiología Celular de la UNAM. Todos los sujetos eran experimentalmente ingenuos, de aproximadamente 60 días de edad. Los sujetos fueron alojados y mantenidos en las mismas condiciones que en los experimentos previos.

Aparatos

Los aparatos utilizados en este experimento fueron los mismos que en los Experimentos anteriores.

Procedimiento

Los sujetos fueron expuestos a las mismas condiciones de habituación, entrenamiento a comedero y entrenamiento al nosepoke con las mismas características que en los experimentos anteriores.

Posteriormente, durante 40 sesiones, la alternativa con valor condicionado diferencial (VCD) estuvo vigente en el panel frontal y la alternativa sin valor condicionado diferencial (SVCD) estaba vigente en el panel trasero (ver figura 21, panel A). Una respuesta en el nosepoke frontal resultó en la presentación de la palanca izquierda A con probabilidad .2, durante 10 segundos y se entregaron 10 pellets; o se presentó la palanca derecha B con probabilidad .8, seguida de ningún reforzador. Una respuesta en el nosepoke trasero resultaba en la presentación de la palanca izquierda C con probabilidad .2, o la palanca derecha D con probabilidad .8, durante 10 segundos seguidas por la entrega de tres pellets. El intervalo entre ensayos fue de 10 segundos, y hubo un total de 60 ensayos por sesión, de los cuales 20 fueron ensayos forzados a la alternativa con valor condicionado diferencial, 20 forzados a la alternativa sin valor condicionado diferencial y 20 de elección. Las sesiones continuaron hasta alcanzar relativa estabilidad en la proporción de elección, seguido de lo cual se realizó una fase de reversión en la cual el panel trasero pasó a ser la alternativa con valor condicionado diferencial; y el frontal, la alternativa sin valor condicionado diferencial, un cambio adicional fue que se revirtió la posición de la palanca que predecía la entrega de 10 pellets en la alternativa con valor condicionado diferencial (ver figura 21).

Condición control

Al finalizar el procedimiento de elección subóptima comenzó una nueva condición en el experimento en la cual se incrementó la probabilidad de aparición de los estímulos tanto en la alternativa con valor condicionado diferencial como en la alternativa sin valor condicionado diferencial, en particular la probabilidad de aparición de los estímulos fue de .5. De esta manera la palanca que anuncia la entrega de 10 pellets en la alternativa con valor condicionado diferencial tenía una probabilidad de aparición de .5, de manera que la tasa global de reforzamiento en esta alternativa fue mayor que en la alternativa sin valor condicionado diferencial y, por lo tanto, se convirtió en la opción “óptima” (ver figura 21, panel B). De igual forma, se presentaron 20 ensayos de cada tipo de manera aleatoria. El objetivo de esta nueva fase fue definir si la elección de los sujetos estaba determinada por evitar los resultados variables más que por la tasa global de reforzamiento de cada alternativa.

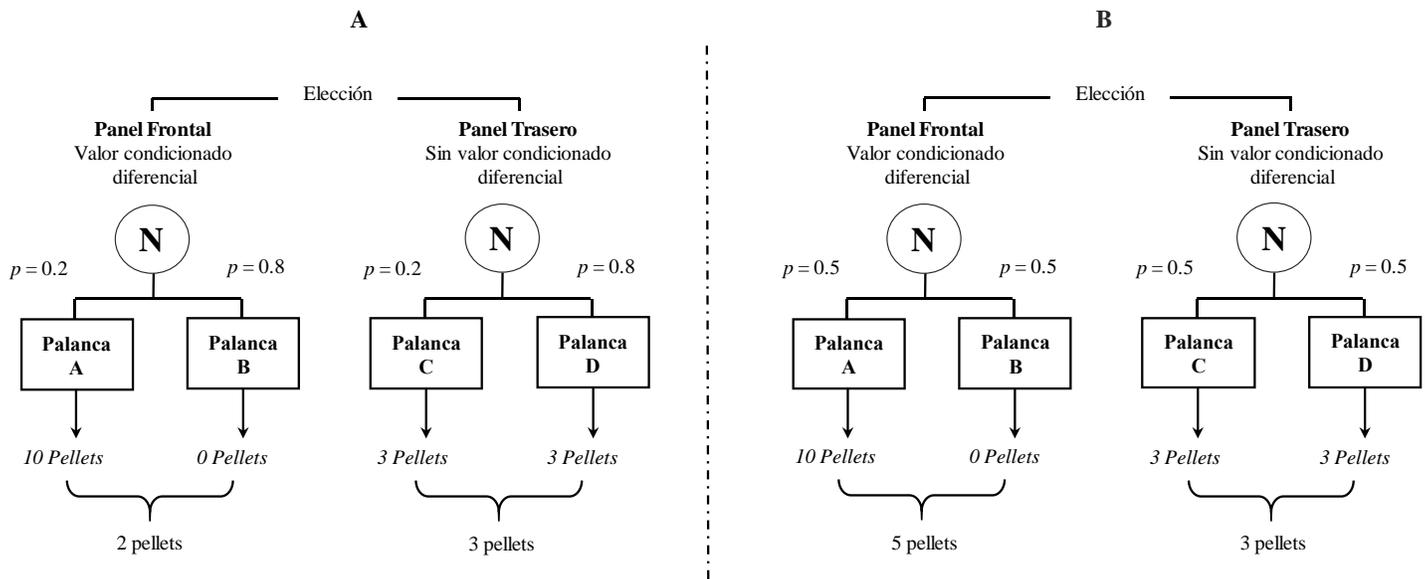


Figura 21. Diagramas de los procedimientos usados en los dos experimentos del presente estudio. Dos alternativas, Valor Condicionado Diferencial (VCD) y Sin Valor Condicionado Diferencial (SVCD) se presentaron en paneles opuestos. En condición principal (panel A) elegir la alternativa VCD resultaba en la presentación de la palanca A con $p=.2$, durante 10 s después de los cuales se entregaron 10 pellets o se presentó la palanca B con $p=.8$, que predecía la ausencia de reforzamiento. Elegir la alternativa SVCD resultó en la presentación de la palanca C con $p=.2$ y la palanca D $p=.8$ ambas asociadas con la entrega de 3 pellets. La alternativa SVCD era la alternativa óptima. En la Condición Control la alternativa VCD la probabilidad de presentar la palanca A se incrementó a $.5$ y la probabilidad de presentar la palanca B se disminuyó a $.5$. En la alternativa SVCD la probabilidad de presentar la palanca C y la palanca D fue de $.5$. Bajo estas condiciones la alterativa VCD era la alternativa óptima.

Resultados

La Figura 22 se presenta el promedio de la proporción de elección de la alternativa con Valor Condicionado Diferencial durante cada una de las 40 sesiones de entrenamiento y las 40 sesiones de la fase reversión. Durante las últimas cinco sesiones de la condición de entrenamiento, cuando se logró la estabilidad, el promedio de la proporción elección para la alternativa con Valor Condicionado Diferencial (subóptimo) fue de 0.35 ± 0.07 . Esta proporción no fue significativamente diferente de indiferencia según una prueba t de dos colas ($t(7) = 2.24, p = 0.059$). Las pruebas Binomiales indicaron que 6 de los 8 sujetos

mostraron una preferencia significativa por la alternativa óptima (p 's < 0.003); y los dos sujetos restantes mostraron una preferencia significativa por la alternativa subóptima (ambas p 's < 0.04).

Durante la reversión, el promedio de la proporción de elección de la alternativa con Valor Condicionado Diferencial alcanzó aproximadamente los mismos niveles que en la condición anterior, y durante las últimas cinco sesiones la proporción fue de 0.34 ± 0.06 , que fue significativamente inferior a la indiferencia ($t(7) = 2.82$, $p < 0.05$), y al igual que en la condición de entrenamiento las pruebas Binomiales mostraron que 6 de los 8 sujetos mostraron una preferencia significativa por la alternativa óptima (p 's < 0.003); una fue indiferente ($p=0.31$), y el otro mostro una preferencia significativa por la alternativa subóptima ($p=0.01$).

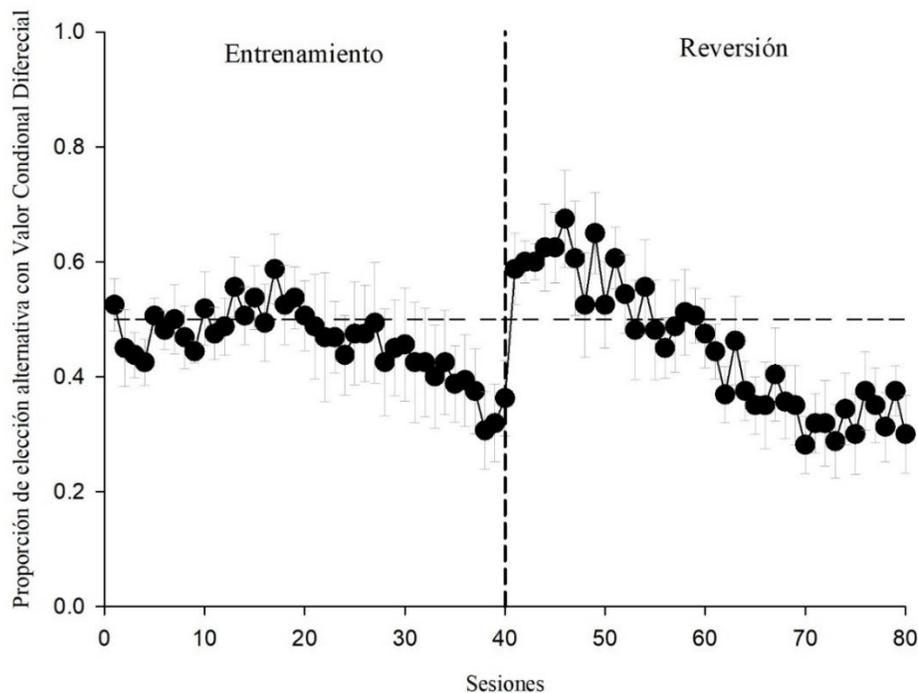


Figura 22. Experimento 4: Promedio de la proporción de elección de la alternativa Con Valor Condicionado Diferencial a lo largo de las 40 sesiones de entrenamiento en el procedimiento de elección subóptima, y las 40 sesiones de la fase de reversión. La línea punteada horizontal señala la indiferencia entre ambas alternativas y la línea punteada vertical indica el inicio de la condición de reversión.

La Figura 23 muestra el número de respuestas en presencia de las palancas en ensayos forzados para la alternativa con Valor Condicionado Diferencial VCD (panel superior) y la alternativa Sin Valor Condicionado Diferencial SVCD (panel inferior). Para la alternativa VCD, el promedio de respuestas durante las últimas cinco sesiones en ambas condiciones (10.36 ± 1.09) a la palanca asociadas con la entrega de 10 pellets fue mayor que el número de respuestas a la palanca asociada con la ausencia del reforzador (0.43 ± 0.16). Un ANOVA realizado para evaluar el efecto de la condición (entrenamiento versus reversión) y estímulos (predictor de 10 pellets versus predictor de cero pellets) indicó que el factor estímulo fue significativo ($F(1, 7) = 70.20, p < 0.0001$), mientras que el factor condición no fue estadísticamente significativo ($F(1, 7) = 0.05, p = 0.83$). La interacción entre el factor condición y el factor estímulo tampoco fue significativa ($F(1, 7) = 0.10, p = 0.76$). Para la alternativa SVCD (figura 22, panel inferior), durante las últimas cinco sesiones de cada condición, el promedio de respuestas fue similar en presencia de los dos estímulos que predijeron tres pellets. Un ANOVA con los mismos factores reveló que no hubo efectos estadísticamente significativos del factor estímulos ($F(1, 7) = 0.004, p = 0.95$) o condición ($F(1, 7) = 1.93, p = 0.21$), ni la interacción fue significativa ($F(1, 7) = 0.01, p = 0.92$).

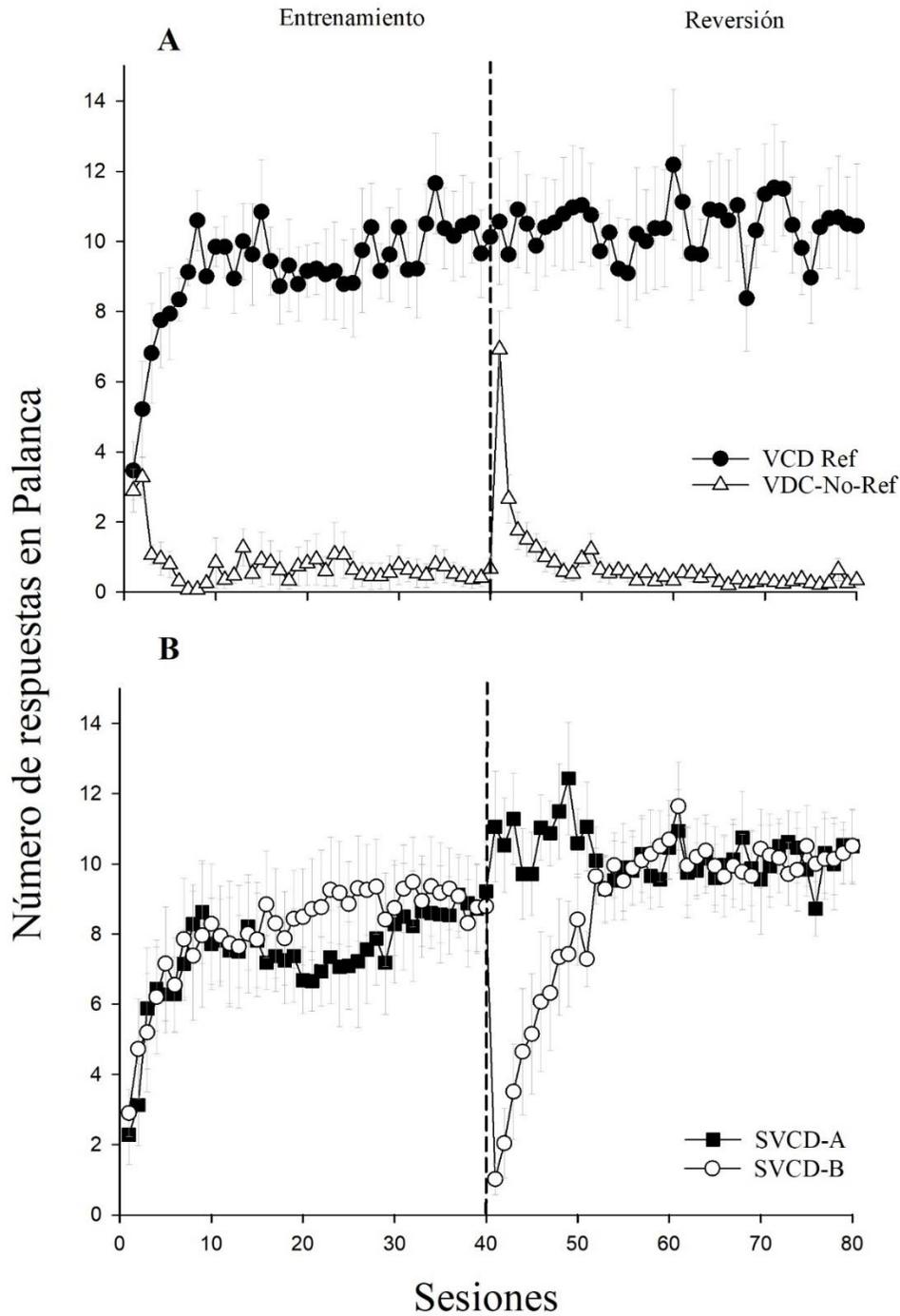


Figura 23. Experimento 4: Número de respuestas a las palancas en los ensayos forzados en todas las sesiones durante la fase de entrenamiento y reversión. El panel A muestra el promedio de respuestas en la alternativa Con Valor Condicional Diferenciado. Los círculos representan el número de respuestas en presencia del estímulo asociado con la entrega de 10 pellets (VCD Ref) y los triángulos representan las respuestas en presencia del estímulo asociado con la ausencia de reforzamiento (VCD-No-Ref). En el panel B se muestra el promedio de respuestas en la alternativa Sin Valor Condicional Diferenciado. Los cuadrados representan las respuestas en la palanca C (SVCD A) y los círculos representan las respuestas en la palanca D (SVCD B).

Condición Control

La Figura 24 muestra el promedio de la proporción de elección para la alternativa VCD en las 40 sesiones que duro esta condición. Desde las primeras sesiones, los valores de esta proporción se alejaron rápidamente de la indiferencia y mostraron una preferencia confiable por la alternativa VCD. Durante las últimas cinco sesiones de esta condición, el promedio de la preferencia por esta alternativa fue de $0,74 \pm 0,03$, que fue estadísticamente significativo superior a la indiferencia ($t(7) = 7,55, p < 0,001$). Las pruebas Binomiales indicaron que los 8 sujetos mostraron una preferencia significativa por la alternativa VCD (p 's $< 0,03$).

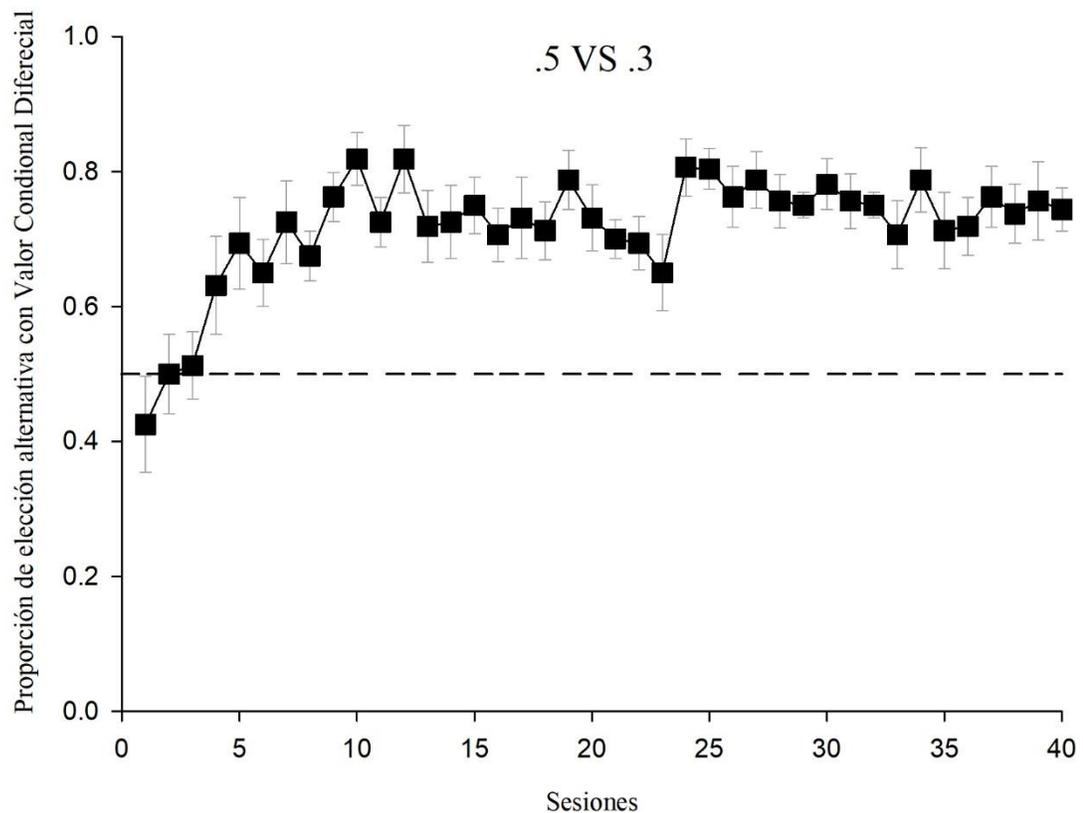


Figura 24. Experimento 4: Promedio de la proporción de elección de la alternativa Con Valor Condicionado Diferencial a lo largo de las 40 sesiones. La línea punteada horizontal señala la indiferencia entre ambas alternativas.

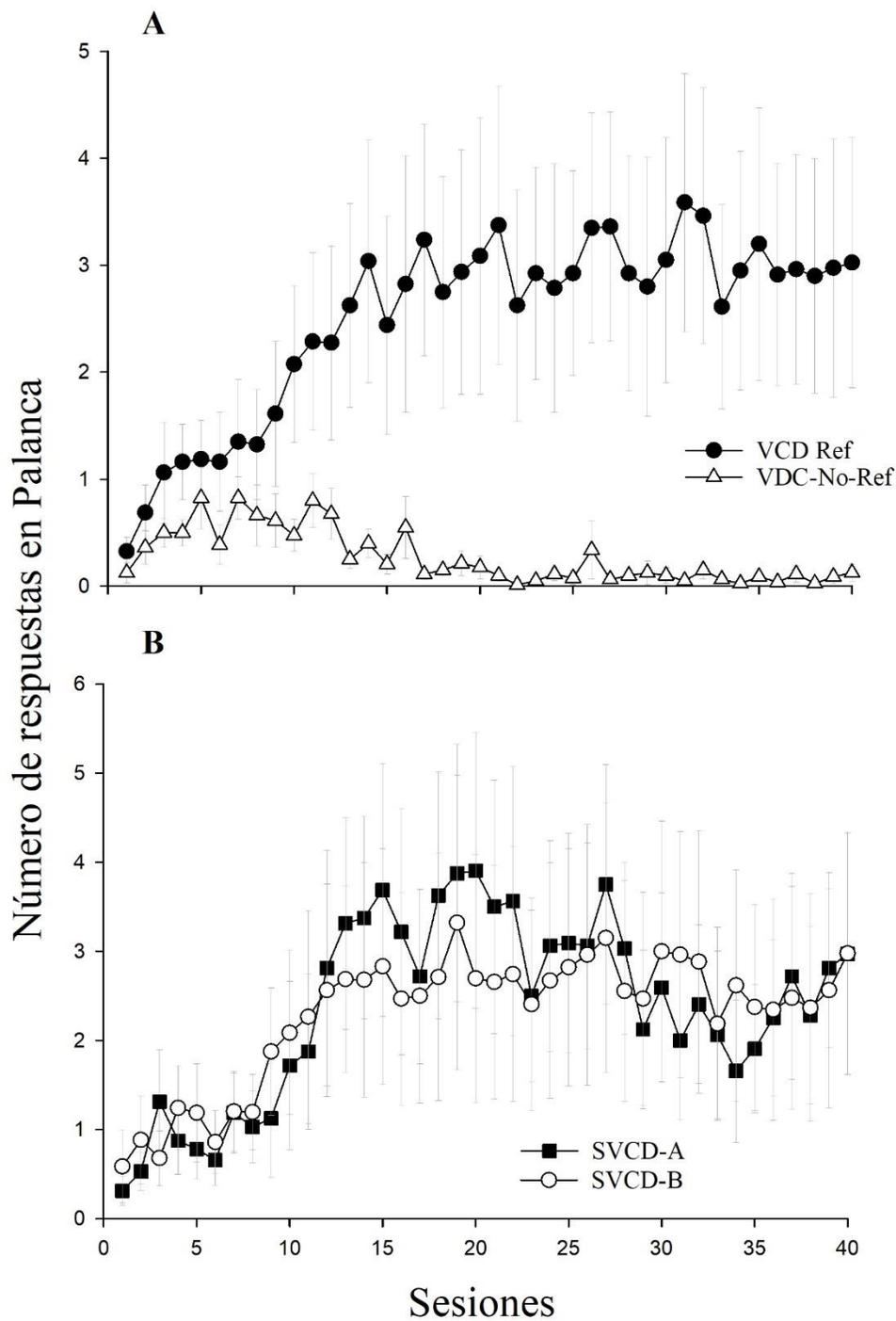


Figura 25. Experimento 4: Número de respuestas a las palancas en los ensayos forzados durante las 40 sesiones de la condición control. El panel A muestra el promedio de respuestas en la alternativa Con Valor Condicionado Diferenciado. Los círculos representan el número de respuestas en presencia del estímulo asociado con la entrega de 10 pellets (VCD Ref) y los triángulos representan las respuestas en presencia del estímulo asociado con la ausencia de reforzamiento (VCD-No-Ref). En el panel B se muestra el promedio de respuestas en la alternativa Sin Valor Condicionado Diferenciado. Los cuadrados representan las respuestas en la palanca C (SVCD A) y los círculos representan las respuestas en la palanca D (SVCD B).

La Figura 25 muestra el número de respuestas en presencia de las palancas en ensayos forzados para la alternativa con Valor Condicionado Diferencial VCD (panel superior A) y la alternativa Sin Valor Condicionado Diferencial SVCD (panel inferior B). Como se puede observar al inicio el promedio de respuestas fue similar ante los dos estímulos asociados a la con Valor Condicionado Diferencial. No obstante, con el paso de las sesiones, en promedio los sujetos respondieron más en presencia del estímulo predictor de 10 pellets que ante el estímulo predictor de la ausencia de reforzamiento. Un ANOVA realizado para evaluar el efecto estímulos (predictor de 10 pellets versus predictor de cero pellets) con el promedio de respuestas de las últimas 5 sesiones indicó que se encontraron diferencias estadísticamente significativas ($F(1, 7) = 6.75, p = .03$).

El mismo análisis se llevó a cabo para contrastar el número de respuestas en palanca ante los estímulos presentados en la alternativa Sin Valor Condicionado Diferencial SVCD y, como se puede apreciar en la figura 25, se encontró que en promedio los sujetos no respondieron diferencialmente ante los estímulos asociados con dicha alternativa ($F(1, 7) = .01, p = 0.89$).

Discusión General

En el presente experimento, evaluamos el comportamiento de elección de las ratas en un procedimiento que ha sido considerado como un análogo animal del comportamiento de juego humano. Los resultados obtenidos aquí son relevantes para la gran cantidad de datos obtenidos con palomas y otras aves evaluadas con procedimientos similares, y también para estudios realizados con ratas en otros procedimientos similares a los juegos de azar.

La comparación más directa de los datos actuales es con los resultados obtenidos con palomas usando el mismo procedimiento (Laude et al., 2014; Laude, Stagner y Zentall 2014; Zentall y Stagner 2011). A diferencia de lo reportado con palomas, encontramos que la mayoría de las ratas (75%) mostraron una clara preferencia por la alternativa sin Valor Condicionado Diferencial (óptima). Un aspecto crucial de estos resultados es el hallazgo de que los sujetos discriminaron la función diferencial de los estímulos (uno asociado con la entrega de diez pellets y el otro asociado con cero pellets), porque se ha demostrado que las palomas hacen elecciones óptimas cuando la alternativa subóptima está compuesta por dos estímulos no discriminativos (Zentall y Stagner 2011, exp.2) y es importante porque excluye la posibilidad de que el comportamiento óptimo presentado por las ratas en este experimento se derivó de la incapacidad de discriminar los estímulos de la alternativa subóptima. Además, los datos de la condición de control sugieren que la aversión potencial a los resultados variables asociados con la alternativa subóptima no jugó un papel determinante en la preferencia por la alternativa óptima.

Es importante destacar que la preferencia óptima observada en el presente experimento se produjo a pesar de que los estímulos utilizados tenían alto valor incentivo (palancas); como se describió en la introducción Chow et al. (2017) plantearon la hipótesis

de que la razón de las diferencias entre las palomas y las ratas en el procedimiento de elección subóptima era que las teclas iluminadas empleadas en los estudios con palomas tenían alto valor incentivo, mientras que las luces y los tonos empleados en los primeros estudios con ratas (Trujano, López, et al., 2016; Trujano y Orduña 2015; Ojeda et al., 2018) no tenían valor incentivo (para un argumento relacionado basado en la teoría de los sistemas de comportamiento, ver Zentall, Smith y Beckmann 2019). Sin embargo, los resultados actuales, junto con otros (López, Alba y Orduña 2018; Martínez et al., 2017; Exp. 1,2 y 3 de la presente tesis), sugieren que el valor incentivo de los estímulos discriminativos no juega un papel importante en el desarrollo de la elección subóptima.

Indudablemente, los modelos animales han contribuido a comprender los roles que desempeñan los estímulos que indican victorias y pérdidas. En particular, el análisis de las diferencias entre ratas y palomas en el procedimiento de elección subóptima ha resaltado algunas variables que tienen un gran potencial para ser incorporadas en el estudio del comportamiento de juego en humanos, por ejemplo, el valor incentivo de los estímulos (Chow et al., 2017), el impacto de la sensibilidad a los estímulos que señalan pérdidas (Laude, Stagner y Zentall 2014; Trujano, López, et al., 2016) y, más estrechamente relacionado con los resultados actuales, el impacto de la información de los estímulos discriminativos.

Es importante mencionar, que en cada una de las condiciones del presente experimento 2 ratas mostraron un comportamiento de elección subóptima. Este resultado es congruente con lo encontrado en un procedimiento distinto que igualmente se ha propuesto como un modelo animal de conducta de juego patológico: una adaptación para roedores de la Iowa Gambling Task (IGT). En este procedimiento, al evaluar a ratas Wistar se ha encontrado una proporción similar de sujetos (25%) que escoge confiablemente la

alternativa más desfavorable (Rivalan, Ahmed, & Dells-Hagedorn, 2009). Estas similitudes entre los dos procedimientos podría representar que estos sujetos que no se comportan de manera óptima capturan un aspecto del modelo animal del comportamiento de juego en humanos, lo que respalda la idea de que el estudio de las diferencias individuales tiene un gran potencial para avanzar en el estudio del comportamiento subóptimo (Rivalan et al., 2009).

Los resultados del presente experimento junto con los datos reportados en los experimentos anteriores de la presente tesis muestran que la conducta óptima presentada por las ratas en este procedimiento es tan fuerte y confiable, como la subóptimalidad en palomas. Además, el hecho de que las ratas en este experimento mostraran una preferencia por la alternativa óptima a pesar de emplear estímulos con alto valor incentivo muestra que la hipótesis del valor incentivo como una variable que modula la conducta de elección en este procedimiento en ratas no es general como fue en principio propuesto.

CAPITULO VII

IMPACTO DEL VALOR INCENTIVO DE LOS ESTÍMULOS EN LA ALTERNATIVAS DEL PROCEDIMIENTO DE “ELECCIÓN SUBÓPTIMA EN RATAS Y PALOMAS”

EXPERIMENTO 5: EL PAPEL DEL VALOR INCENTIVO DE LOS ESTÍMULOS EN LA CONDUCTA SUBÓPTIMA EN PALOMAS.

En una situación de elección que involucra dos alternativas que difieren en la tasa global de reforzamiento y la información que predicen los estímulos asociados con cada alternativa, las palomas muestran de manera contundente una fuerte preferencia por las alternativas que proporcionan información sobre cuándo se va a entregar un reforzador y cuando no se recibirá un reforzador, en lugar de elegir las alternativas asociadas con mayores tasas de reforzamiento (McDevitt et al., 2016; T. R. Zentall, 2016; T. R. Zentall & Stagner, 2011). Debido a que este resultado se ha reportado de manera contundente en más de una docena de estudios, se le ha nombrado a este procedimiento como “procedimiento de elección subóptima”.

Cuando se presenta una situación similar de elección, pero las alternativas solo difieren en la tasa global de reforzamiento, debido a que los estímulos presentados en cada alternativa no proporcionan ninguna información respecto a la presencia o ausencia de reforzamiento, las palomas eligen consistentemente las alternativas asociadas con la mayor tasa global de reforzamiento (ver Fig. 2; Stagner and Zentall, 2010). Esto muestra que las palomas son sensibles a este parámetro y a pesar de ello, prefieren una menor tasa de reforzamiento a cambio de estímulos que les permitan discriminar la entrega o no entrega de un reforzador.

Se han propuesto varias teorías para explicar por qué las palomas eligen estas alternativas y no las alternativas que las llevarían a comportarse de manera óptima. En particular estas teorías han propuesto que la fuerte preferencia de las palomas por la alternativa subóptima se debe a la importancia que le dan las palomas al estímulo que

predice la presentación de un reforzador y el poco efecto que tienen los estímulos que predicen la ausencia de reforzamiento (Fortes et al., 2016; Laude, Stagner, et al., 2014; Smith & Zentall, 2016; Vasconcelos et al., 2015).

A pesar de encontrar generalidad de este resultado en diversos estudios, este hallazgo se ha reportado en su mayoría usando palomas como sujetos experimentales, y cuando se ha buscado la generalidad de este hallazgo en otras especies, se ha reportado que las ratas la especie más utilizada en el análisis experimental de la conducta muestran el comportamiento opuesto presentado por las palomas, es decir, las ratas eligen las alternativas que las llevan a obtener la mayor tasa global de reforzamiento (López et al., 2018; Martínez et al., 2017; Ojeda et al., 2018; Trujano & Orduña, 2015).

En este contexto se han propuesto algunas hipótesis para explicar estas diferencias entre especies siendo la hipótesis del valor incentivo de los estímulos la propuesta que más ha llamado la atención en los últimos años (Chow et al., 2017). Esta hipótesis básicamente propone que la razón de las diferencias de comportamiento entre las palomas y las ratas en el procedimiento de elección subóptima radica en el valor incentivo de los estímulos discriminativos empleados con cada especie: mientras que para las palomas los estímulos son, sin excepciones, teclas iluminadas que tienen un alto valor incentivo, los estímulos empleados en los primeros estudios con ratas fueron luces (Trujano et al., 2016; Trujano y Orduña 2015) o tonos (Ojeda, Murphy y Kacelnik 2018), que no tienen alto valor incentivo para esta especie.

Recientemente se ha probado esta hipótesis utilizando estímulos con alto valor incentivo en ratas, y los resultados reportados muestran que esta hipótesis no tiene resultados concluyentes debido a que las ratas se comportan de manera óptima a pesar de usar estímulos con alto valor de incentivo en el procedimiento de elección subóptima

(López, Alba y Orduña 2018; Martínez et al., 2017; Exp. 1,2, 3 y 4 de la presente tesis). Por lo tanto, la hipótesis del valor incentivo de los estímulos como la variable responsable de promover comportamiento subóptimo en ratas no es la variable que explica esta diferencia entre especies.

A pesar de que esta evidencia parece descartar la importancia del valor incentivo de los estímulos como la variable que promueve el comportamiento subóptimo, es importante destacar que esta hipótesis ha sido ampliamente estudiada en ratas, sin embargo, en todos los estudios del procedimiento de elección subóptima usando palomas, se han empleado teclas iluminadas (estímulos con alto valor incentivo) lo cual no permite descartar que esta variable si module el comportamiento de las palomas en este procedimiento.

Siguiendo la lógica de la hipótesis del valor incentivo de los estímulos planteamos que las palomas elegirán las alternativas asociadas con la mayor tasa global de reforzamiento (óptimas) cuando los estímulos discriminativos tengan bajo valor incentivo. En este contexto, en el presente experimento usaremos tonos como estímulos discriminativos, en lugar de las teclas iluminadas.

Método

Sujetos

Los sujetos fueron 6 palomas experimentalmente ingenuas de aproximadamente 2 años de edad. Se mantuvo a los sujetos al 80% de su peso en alimentación libre y se alojaron en jaulas individuales con acceso libre a agua. en una habitación con temperatura controlada, bajo un ciclo de luz oscuridad de 12-12 hr.

Aparatos

Se utilizó una caja experimental para palomas (Med Assoc. Mod. ENV-007) modificada, situada dentro de una caja sonoamortiguada. Las paredes laterales fueron de plástico transparentes y el suelo fue una rejilla formada por 19 barras de acero inoxidable de .48 cm de diámetro. Como pared trasera y pared frontal, se colocaron dos paneles operativos idénticos (ver Figura 26).

Cada panel contó con un receptáculo de pellets MED Associates, Inc, Model ENV-200R2M) ubicado a 1.5 cm del suelo y a 9.3 cm de las paredes laterales, en el cual se entregaron pellets de 14 mg (BioServ, Product F05684) desde un módulo dispensador circular (MED Associates, Inc., Model ENV-203-14P). Las entregas de los pellets fueron acompañadas por una iluminación de un LED ultra-brillante blanco durante 3 s. Para medir las entradas al receptáculo de pellets se contó con fotorreceptores en el receptáculo de pellets. En la pared de cada panel se colocó un foco de 28 V que proporcionó la iluminación general de la caja (Med Assoc. Mod. ENV-215 M) y 3 cm debajo de cada foco se instaló una bocina conectada a un generador de tonos.

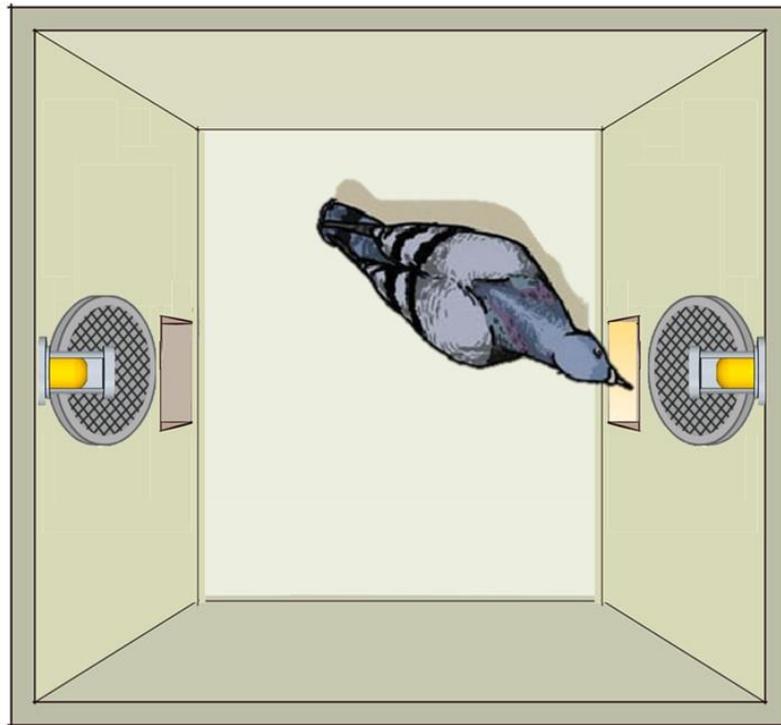


Figura 26 Diseño de la caja operante modificada.

Para operar el equipo se utilizó una computadora personal conectada a una interfase MED-PC (Med Associates, St. Albans, VT.).

Entrenamiento a comedero

En dos sesiones de 25 minutos cada una, se entrenó a las palomas a comer del receptáculo de pellets con el siguiente procedimiento: al principio de cada sesión se colocaron 15 pellets en cada receptáculo de pellets y se mantuvo encendido el led de cada receptáculo. Cuando los animales consumieron 30 pellets en total (15 obtenidos en el panel frontal y 15 obtenidos en el panel trasero) se dio por concluido el entrenamiento a comedero.

Entrenamiento para responder al receptáculo de pellets.

Una vez finalizado el entrenamiento a comedero, se expuso a las palomas a un procedimiento de reforzamiento continuo (RfC), hasta obtener la respuesta de meter la

cabeza al receptáculo de pellets por comida. Al inicio de cada sesión experimental se encendió la luz del receptáculo de pellets y una entrada al receptáculo resultó en la entrega de tres pellets y la luz del LED se mantuvo encendida durante 3 s. Después de la entrega de la comida se inició un nuevo ciclo. En esta fase cada sesión estuvo vigente durante 30 minutos o 40 ensayos lo que ocurriera primero. Se expuso a las palomas a este procedimiento durante 1 sesión y en sesiones posteriores se incrementó el requisito de respuesta a RF2, RF4, RF6 y RF10. Se dio por concluido el entrenamiento a responder al receptáculo de pellets cuando las palomas cumplieron con el último valor del RF.

Pre-entrenamiento

Una vez que las palomas respondieron consistentemente a ambos receptáculos de pellets, se llevaron a cabo sesiones de pre-entrenamiento hasta que se observaron latencias similares en las respuestas a los receptáculos, a fin de reducir la posibilidad de sesgo de posición. Específicamente, un programa de reforzamiento RF1 estaba presente en el panel frontal o en el panel trasero de manera aleatoria. Para señalar en qué panel se encontraba disponible el programa de reforzamiento, al inicio del ensayo se encendía el LED del receptáculo de pellets correspondiente. Con la primera respuesta en el receptáculo, es decir, con la interrupción del haz infrarrojo de los fotorreceptores, se entregaban tres pellets como reforzador.

Cada sesión consistió en 50 ensayos en total, 25 ensayos para el receptáculo del panel trasero y 25 ensayos en el receptáculo del panel frontal. Entre cada ensayo hubo un intervalo entre ensayos (IEE) de 10s, en los cuales no se presentó ningún estímulo. Esta fase fue concluida después de cinco sesiones, una vez que la latencia de los sujetos fue menor a 3 s y la diferencia de su latencia en el receptáculo frontal y el receptáculo trasero fue menor a 1.5 s.

Procedimiento

Cada una de las sesiones consistió en 60 ensayos divididos de la siguiente forma: 20 ensayos forzados para la alternativa Discriminativa, 20 ensayos forzados para la alternativa No Discriminativa y 20 ensayos de elección, todos presentados de manera aleatoria. El inicio de un ensayo forzado inició con la iluminación del LED del receptáculo de pellets. En el caso de la alternativa Discriminativa, se encendió el LED del receptáculo de pellets frontal y una respuesta a este receptáculo en el 20% de los ensayos, se presentó un tono de 1.5 kHz y se encendió la luz general del panel frontal durante 10 s, después de este tiempo la primera respuesta al receptáculo fue seguida por la entrega de tres pellets (IF 10 s). En el 80% restante de los ensayos, se presentó un Clicker que se encendía y apagaba cada .25 s acompañado con la iluminación de la luz general del panel frontal, y después de 10 s, el ensayo termino sin entregar ningún reforzador. En esta alternativa, la probabilidad global de reforzamiento fue de .2. En los ensayos forzados de la alternativa No Discriminativa, se encendió el LED del receptáculo de pellets del panel trasero y una sola respuesta al receptáculo presentó el 100% de las veces un tono de 2.5 kHz y después de que se cumplió el tiempo de un programa IF 10 s la primera respuesta en el receptáculo después de este tiempo resultó en la entrega de tres pellets con probabilidad .50. En esta alternativa, la probabilidad global de reforzamiento fue de .5 (ver Figura 27).

En los ensayos de elección, se encendieron ambos LEDs de los receptáculos de pellets y la primera respuesta en cualquiera de los dos presentó las mismas contingencias que se describieron para los ensayos forzados. Cuando finalizó un ensayo, se presentó un ITI de 10 s durante el cual no se presentó ningún estímulo, y se inició un nuevo ensayo después. Los estímulos presentados y la posición de la alternativa Discriminativa (panel frontal o panel trasero) se contrabalancearon entre los sujetos.

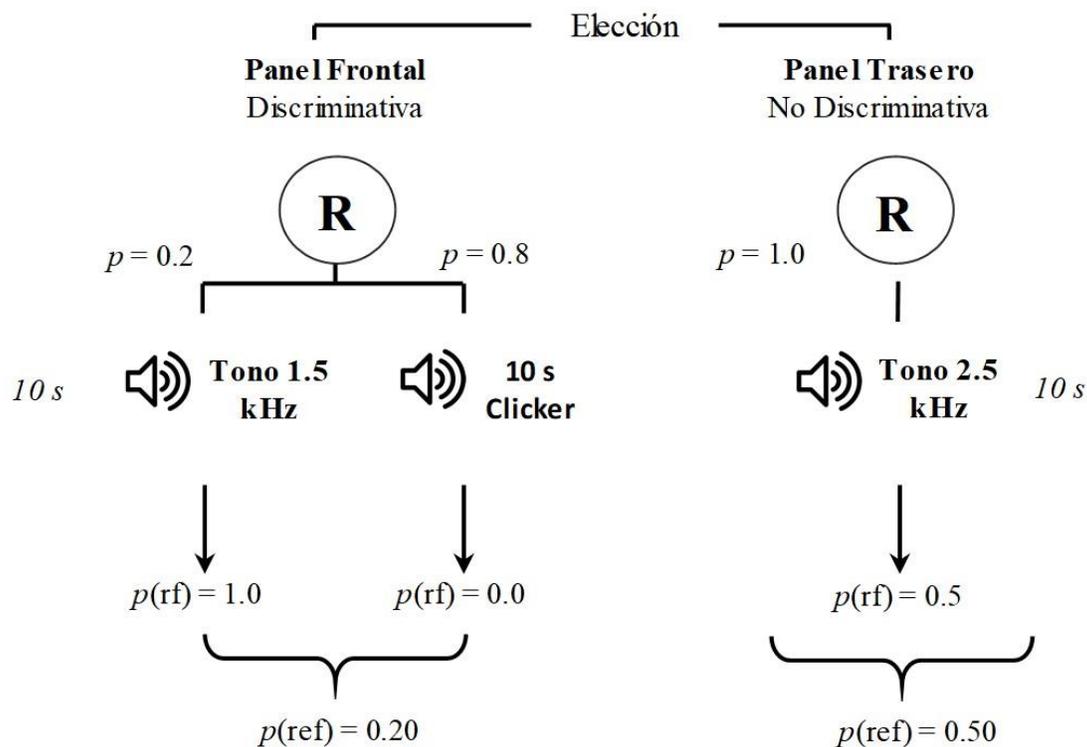


Figura 27 Diagrama del procedimiento usado en el presente experimento. Las dos alternativas Discriminativa y No Discriminativa se presentaron en paneles opuestos. Elegir la alternativa Discriminativa resultaba en la presentación de Tono de 1.5 kHz con $p=.2$, durante 10 s después de los cuales se entregaron 3 pellets o se presentó un clicker (encendía y apagaba en intervalos de .5 s) con $p=.8$, que predecía la ausencia de reforzamiento. Elegir la alternativa No Discriminativa resultó en la presentación un Tono de 2.5 kHz en todos los ensayos asociado con una probabilidad de reforzamiento de .5. Todos los tonos presentados estuvieron controladas a 65 decibeles.

Resultados

La figura 28 muestra el promedio de la proporción de elección de la alternativa Discriminativa durante todo el experimento. Al inicio del experimento el valor de dicha proporción estuvo alrededor de .58, es decir, los sujetos no mostraron preferencia por una alternativa particular. Sin embargo, con el paso de las sesiones tal valor disminuyó de manera que, en las últimas cinco sesiones, los sujetos mostraron una clara preferencia por la alternativa óptima. Para descartar que la preferencia mostrada por los sujetos fuese

producto del azar, se empleó una prueba T para una muestra en la que se comparó el promedio de la proporción de elección de las últimas cinco sesiones contra la indiferencia (.5). Se encontró que el promedio de la proporción de elección fue estadísticamente diferente de la indiferencia ($t(5) = -2.878$, $p < .03$). Las pruebas Binomiales indicaron que 4 de los 6 sujetos mostraron una preferencia significativa por la alternativa óptima (p 's < 0.001); y los dos sujetos restantes mostraron indiferencia ($p=0.89$), y ($p=0.133$).

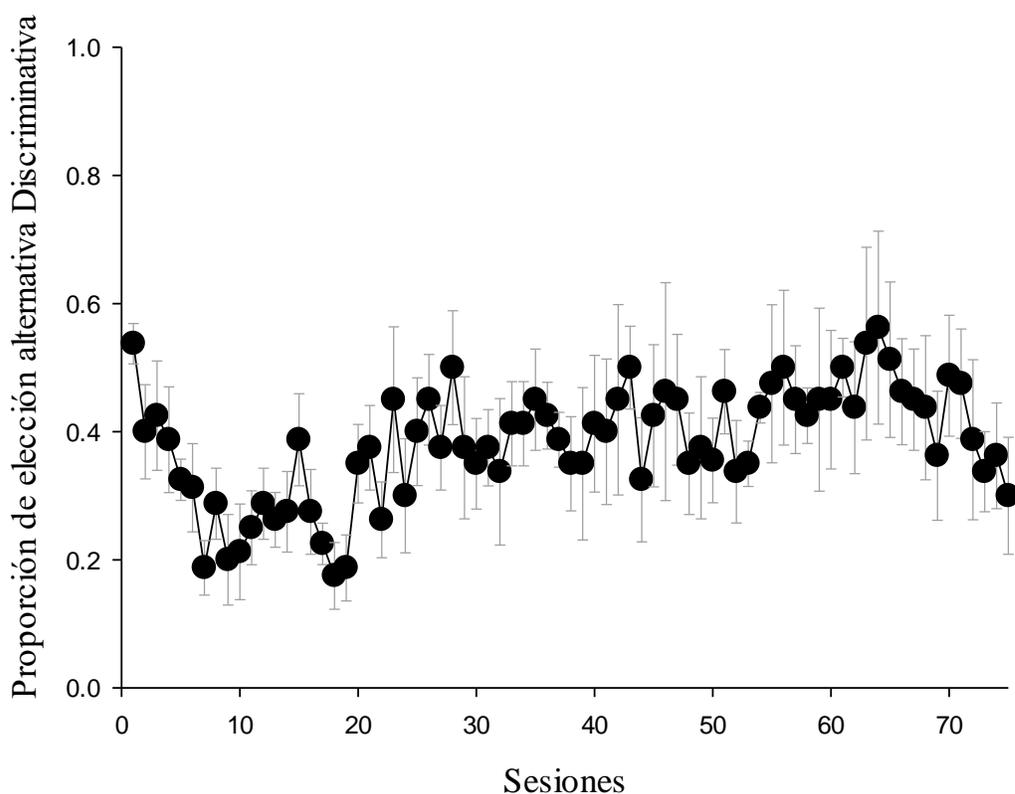


Figura 28. Experimento 5: Promedio de la proporción de elección de la alternativa Discriminativa a lo largo de las 75 sesiones. La línea punteada horizontal señala la indiferencia entre ambas alternativas.

Con el propósito de evaluar si los sujetos fueron sensibles a las consecuencias programadas en los estímulos de la alternativa Discriminativa y No Discriminativa, se registró el número de entradas al receptáculo de pellets emitidas durante el los IF10s y el TF10 de los ensayos forzados. En la figura 29 se puede observar que al inicio el promedio de respuestas de todos los sujetos fue similar ante los dos estímulos asociados a la alternativa Discriminativa. No obstante, con el paso de las sesiones, en promedio los sujetos respondieron más en presencia del estímulo predictor de reforzamiento que ante el estímulo predictor de la ausencia de reforzamiento.

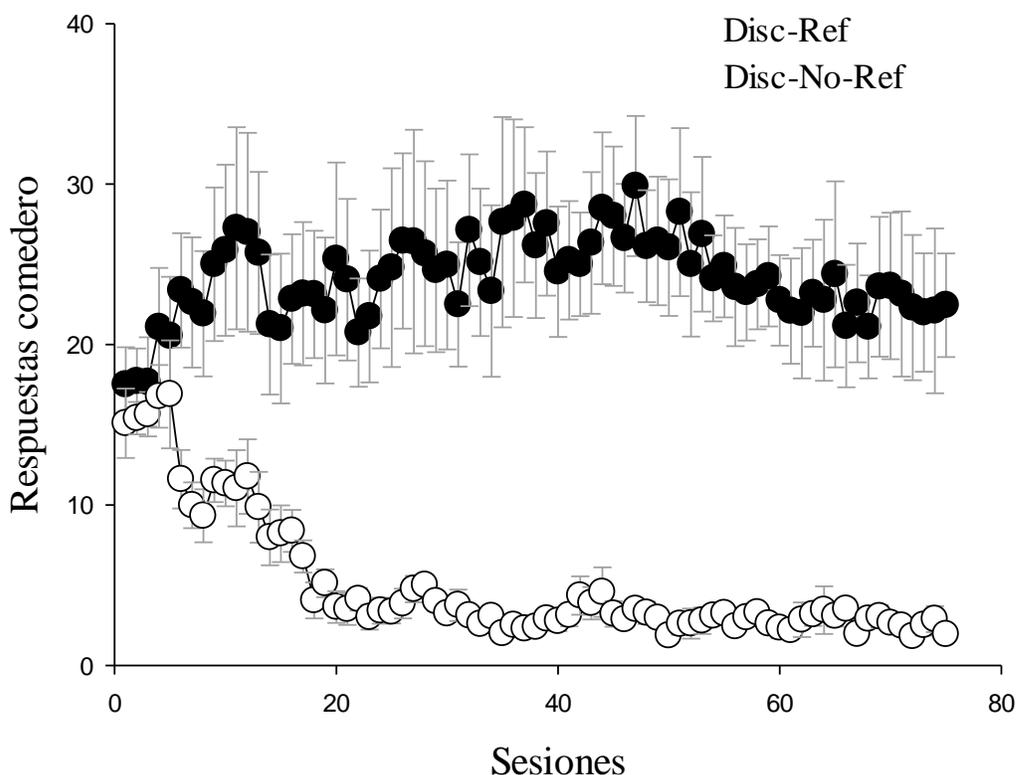


Figura 29. Experimento 5: Número de entradas al comedero en los ensayos forzados de la alternativa Discriminativa en todas las sesiones del presente experimento. Los símbolos cerrados representan el número de entradas al comedero en presencia del estímulo asociado con el reforzador (Disc-Ref) y los símbolos abiertos representan las entradas al comedero en presencia del estímulo asociado con la ausencia de reforzamiento (Disc-No Ref).

Un Anova del número de respuestas a al receptáculo de pellets en presencia de los estímulos de la alternativa Discriminativa informó diferencias estadísticamente significativas ($F(1, 5) = 32.819$ $p < 0.002$). Con el objetivo de garantizar que los resultados observados fuesen atribuibles a la discriminación de las contingencias asociadas a cada estímulo, se calculó el índice de discriminación de las respuestas al receptáculo de pellets que fue definido como el cociente de la división del número de respuestas emitidas en presencia del estímulo predictor de reforzamiento, dividido entre el número total de respuestas.

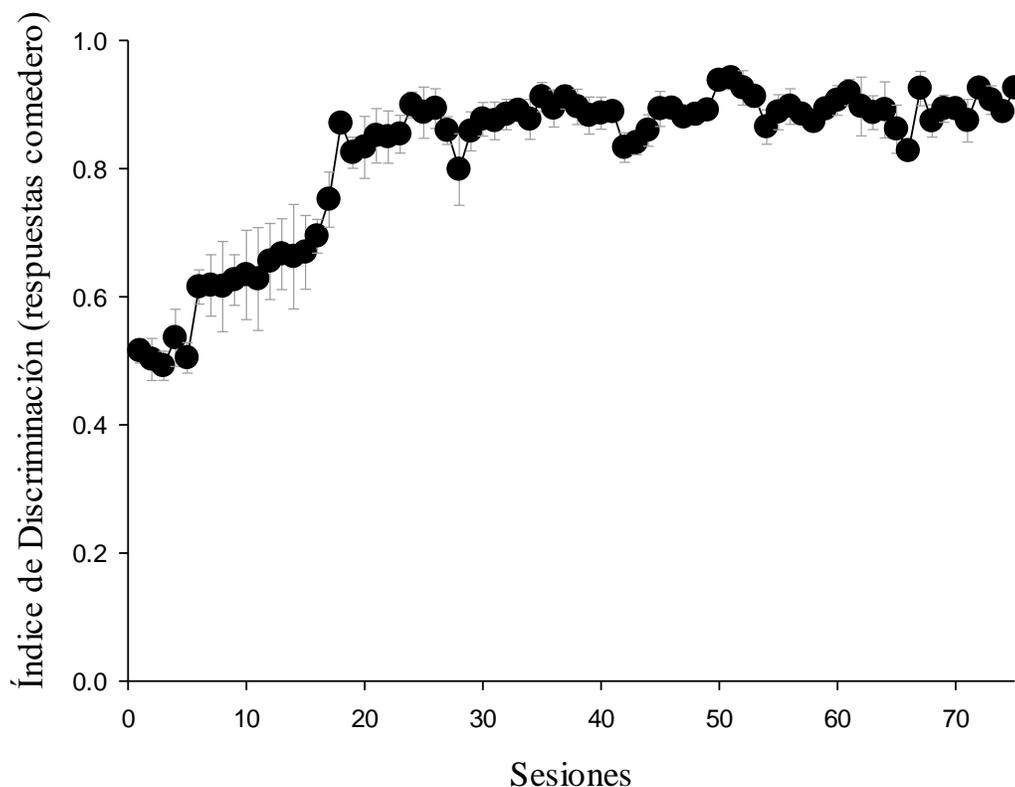


Figura 30. Experimento 5: Promedio del índice de discriminación a lo largo de las 75 sesiones del presente experimento.

El índice podía tomar valores de 0 a 1, donde 1 significaría discriminación perfecta (emisión exclusiva de respuestas ante el estímulo predictor de reforzamiento), y 0.5

significaría ausencia de discriminación. Como se observa en la Figura 30 el promedio del valor del índice de discriminación estuvo por arriba del .5 desde las primeras sesiones, este valor fue aumentando conforme el experimento transcurrió hasta alcanzar valores notablemente altos (0.90 ± 01) y consistente en todos los sujetos (rango .88-.93).

Discusión

En el presente experimento se buscó aportar evidencia que contribuya al debate acerca de posibles diferencias entre palomas y ratas en el procedimiento de elección subóptima. Mientras que algunos antecedentes indican que existen diferencias entre ratas y palomas y que éstas pueden deberse a un impacto diferencial de los inhibidores condicionados (Laude et. al., 2014; Trujano et. al., 2016), investigaciones recientes sugieren que cuando ambas especies son evaluadas con estímulos discriminativos con alto valor incentivo, sus diferencias son eliminadas (Chow et. al., 2017). Sin embargo, la evidencia presentada en contra de esta hipótesis es contundente en ratas, debido que muestran un comportamiento óptimo sin importar el valor incentivo de los estímulos que se empleen en la tarea de elección subóptima (López, Alba y Orduña 2018; Martínez et al., 2017; Exp. 1,2, 3 y 4 de la presente tesis). Sin embargo, antes de descartar la hipótesis del valor incentivo de los estímulos como una explicación a estas diferencias entre especies, para una adecuada evaluación de esta hipótesis, resulta necesario incluir a ambas especies en esta evaluación. En todos los estudios que utilizan palomas como sujetos experimentales en este procedimiento utilizan estímulos con alto valor incentivo. Lo cual no permite descartar que el valor incentivo de los estímulos module el comportamiento subóptimo en palomas, debido a que solo se ha evaluado usando estímulos con alto valor de incentivo.

El presente experimento pretendió complementar la evaluación de esta hipótesis en palomas, utilizando estímulos con bajo valor incentivo en el procedimiento de elección

subóptima. Por esa razón, en el procedimiento empleado en el presente experimento se utilizaron tonos en lugar de teclas iluminadas como estímulos discriminativos. Los resultados indicaron que las palomas se comportaron de manera óptima, eligiendo la alternativa asociada con la mayor probabilidad global de reforzamiento. Este resultado es contradictorio a todos los resultados reportados con palomas utilizando este procedimiento (Smith et al., 2016; Stagner & Zentall, 2010; Zentall & Stagner, 2011).

La principal diferencia procedimental del presente experimento respecto a los estudios previos es el valor incentivo de los estímulos y que las consecuencias programadas de cada alternativa fueron presentadas en distintos paneles operativos. No obstante, el número de respuestas en presencia de los estímulos asociados a la opción Discriminativa, dan cuenta de que los sujetos lograron distinguir la función particular de cada estímulo. La Figura 30 muestra que el índice de discriminación para la alternativa Discriminativa durante todo el experimento (respuestas al receptáculo de pellets en presencia del estímulo que predecía la entrega de reforzamiento / (respuestas al receptáculo en presencia del estímulo predictor de reforzamiento + respuestas en el receptáculo en presencia del estímulo que predecía la ausencia de un reforzador)) fue notablemente alto (0.90 ± 0.03) y consistente entre todos los sujetos (rango .84 - .90). En concreto, los sujetos emitieron un mayor número de respuestas en presencia del estímulo predictor de reforzamiento que ante el estímulo predictor de la ausencia de reforzamiento, lo cual demuestra que los sujetos pudieron discriminar las contingencias asociadas a los estímulos y la optimalidad encontrada fue producto de su sensibilidad a las probabilidades globales de reforzamiento.

La conclusión a la que llevan los presentes resultados es que la hipótesis del valor incentivo que postula que las palomas van a mostrar conducta subóptima al utilizar estímulos con alto valor incentivo y conducta óptima al utilizar estímulos con bajo valor

incentivo si explica el comportamiento de las palomas en este procedimiento, contrario a los datos encontrados con ratas donde se ha reportado que el valor incentivo no promueve el desarrollo de la conducta subóptima.

Los resultados del presente experimento aportan evidencia a favor de la hipótesis del valor incentivo de los estímulos como una explicación de la conducta subóptima en palomas. Sin embargo, van en contra de las diferentes hipótesis que se han planteado en la literatura para explicar este comportamiento subóptimo [SiGN (Margaret A McDevitt et al., 2016), Contraste (Zentall, 2016), el Modelo Ecológico (Fortes, Machado, & Vasconcelos, 2017; Fortes et al., 2016; Vasconcelos et al., 2015)].

Específicamente, McDevitt et al., (2016) propuso que el desarrollo de la preferencia de la alternativa subóptima se debía a la discrepancia entre reforzamiento primario y condicionado: el reforzamiento primario favorece siempre a la alternativa No discriminativa (óptima); sin embargo, en esa alternativa el reforzamiento condicionado no tiene ningún efecto debido a que los estímulos no señalan confiablemente la entrega de ningún reforzador. Basándose en la hipótesis de la reducción de la demora de Fantino, propuso que en la alternativa No Discriminativa (óptima) los estímulos no señalan ninguna reducción del tiempo previo a la entrega de reforzamiento. En cambio, en la alternativa Discriminativa el estímulo que señala la entrega inminente de un reforzador “señal de buenas noticias” sí funciona como reforzador condicionado y por lo tanto señala una disminución en la demora de reforzamiento. De este modo, el tiempo en la reducción de la demora de entrega de las recompensas favorece a la alternativa Discriminativa. Los resultados del presente experimento son contradictorios con esta propuesta, debido a que a pesar de que la alternativa No discriminativa no señala ninguna reducción del tiempo previo a la entrega de una recompensa, las palomas prefirieron consistentemente esta

alternativa a pesar de que en la alternativa Discriminativa (subóptima) el estímulo predictor de reforzamiento señalo mayor reducción de la demora de los reforzadores.

La hipótesis de contraste describe que las palomas prefieren la alternativa Discriminativa (subóptima) debido al contraste entre el valor esperado de reforzador (20% esperado) y el valor obtenido al recibir un reforzador (100% obtenido) dada la elección de la alternativa subóptima, mientras que habría poco contraste con la alternativa óptima (100% de refuerzo esperado y 100% de refuerzo obtenido). Sin embargo, esta hipótesis no explica los datos del presente experimento en el cual las palomas prefirieron la alternativa óptima a pesar del poco contraste de esta alternativa. Una posible explicación de este resultado podría ser el número de sesiones utilizadas en la tarea (Case & Zentall, 2018), sin embargo, las palomas desarrollaron una preferencia por la alternativa óptima en el mismo número de sesiones que los estudios que postularon esta hipótesis, lo cual deja descartada que esta sea la razón de estos resultados contradictorios.

Por último, Vasconcelos et al., (2015) propusieron que en el ambiente natural los animales no toman en cuenta los estímulos que señalan la ausencia de una recompensa, en su lugar utilizan esta información para buscar otros estímulos que si les indiquen la entrega de alimento. Bajo esta propuesta el estímulo que señala la entrega de una recompensa en la alternativa Discriminativa favorece este mecanismo y, por lo tanto, el desarrollo de una preferencia por las alternativas que brindan información. Contrario a esta hipótesis los datos del presente experimento no son explicados por este mecanismo, debido a que las palomas desarrollaron una preferencia por la alternativa No Discriminativa que no cuenta con información confiable sobre cuando se va a recibir un reforzador.

Las tres hipótesis antes mencionadas, tienen en común el poco valor que le atribuyen al estímulo que predice la ausencia de una recompensa en la preferencia de las

palomas en este procedimiento y esto está fundamentado por la evidencia que señala que las palomas al inicio del entrenamiento, durante las sesiones en las que los sujetos muestran una preferencia por la alternativa óptima, la inhibición condicionada generada por el estímulo asociado a la no entrega de reforzador tiene una fuerte influencia en la elección del sujeto. Sin embargo, el efecto de la inhibición condicionada disminuye con el paso de las sesiones, correlacionando con un incremento en la preferencia por la alternativa subóptima (Laude et.al., 2014). Una posibilidad del poco efecto que tiene este estímulo como inhibidor condicionado es que una vez que se identifica el estímulo negativo, las palomas se alejan de él, lo que limita su efectividad.

Se han reportado en la literatura manipulaciones en las características del estímulo predictor de la ausencia de una recompensa en un intento por aumentar su valor. Por ejemplo, Stagner, Laude, & Zentall (2011) diseñaron un experimento en el que fuera imposible evitar el estímulo predictor de la ausencia del reforzador, reemplazando la luz en la tecla por una luz ambiental que iluminaba toda la caja operante, obligando así la exposición del animal al estímulo. Los resultados de este experimento fueron que los sujetos expuestos a la luz general como estímulo inhibitorio adquirieron de igual manera la preferencia por la opción subóptima. Los datos del presente experimento tienen en común con este estudio que el estímulo que se utilizó como predictor de la ausencia de una recompensa se presenta de manera contextual, por lo cual el organismo no puede evitar ser expuesto a este estímulo, sin embargo, los datos del presente estudio sugieren que contrario a los resultados de Stagner et al., las palomas si desarrollaron una fuerte sensibilidad al inhibidor condicionado y por lo tanto mostraron una fuerte preferencia por la alternativa óptima que es congruente a lo reportado con ratas que muestran una fuerte preferencia por

la alternativa óptima al tomar en cuenta en su elección al inhibidor condicionado presente en la alternativa Discriminativa.

Esta diferencia de resultados puede deberse a las características de los estímulos empleado en cada estudio, en particular, Stagner et al., utilizó teclas (con alto valor incentivo) en todos los estímulos del procedimiento con excepción del estímulo predictor de no reforzamiento, por lo tanto, el contraste entre el valor incentivo de los estímulos (alto para el predictor de reforzamiento y bajo para el predictor de la ausencia de reforzamiento) pudo originar que las palomas siguieran prefiriendo la alternativa Discriminativa (subóptima) debido al bajo valor incentivo que tiene la luz general comparado con el alto valor que tenía el predictor de reforzamiento. En cambio, en el presente experimento todos los estímulos utilizados en el procedimiento tuvieron bajo valor incentivo que no originó este contraste entre los estímulos de la alternativa Discriminativa y los sujetos fueron sensibles al inhibidor condicionado y eligieron la alternativa óptima. Esta evidencia sugiere que el valor incentivo de los estímulos del procedimiento de elección subóptima tiene un efecto en el impacto que tiene cada alternativa en la elección de los sujetos.

EXPERIMENTO 6 (A y B): EL VALOR INCENTIVO DE LOS ESTÍMULOS SESGA LAS PREFERENCIAS DE LAS RATAS EN EL PROCEDIMIENTO DE "ELECCIÓN SUBÓPTIMA".

En los últimos años, el procedimiento de elección subóptima se ha empleado ampliamente para estudiar los mecanismos de elección en humanos (Molet et al., 2012), palomas (Stagner y Zentall 2010), ratas (Trujano y Orduna 2015) y otras especies (Vasconcelos, Monteiro y Kacelnik 2015), probablemente debido a su relación con algunos casos de comportamiento desadaptativo en humanos, como la adicción al juego (Molet et al., 2012).

Sin embargo, una de las cuestiones no resueltas con respecto a este procedimiento es la gran diferencia que se ha encontrado al comparar el rendimiento de las palomas y las ratas: como se indicó anteriormente, en docenas de experimentos las palomas han mostrado una fuerte preferencia por la alternativa Discriminativa-subóptima (Kendall 1974; Smith y Zentall 2016; Stagner y Zentall 2010). En contraste, la evidencia acumulada muestra que las ratas prefieren la alternativa óptima, a pesar de que discriminan fácilmente los diferentes estímulos asociados con ella (López, Alba y Orduña 2018; Martínez et al., 2017; Ojeda, Murphy y Kacelnik 2018; Trujano y Orduña 2015; Trujano et al., 2016). Aunque se han sugerido diferentes hipótesis para explicar esta diferencia entre especies (Chow et al., 2017; Trujano et al., 2016; Laude et al., 2014; Zentall, Smith y Beckmann 2019), hoy en día no hay un acuerdo sobre su causa.

En marcado contraste con estos resultados, en un área de investigación relacionada se ha demostrado que el valor incentivo de los estímulos tiene un fuerte impacto sobre la preferencia, por ejemplo, Beckman and Chow (2015), emplearon un procedimiento de Aproximación Pavloviana Condicionada para evaluar simultáneamente en el mismo sujeto

el valor incentivo de dos estímulos con diferente modalidad, una palanca y un tono.

Después de mostrar que la palanca generaba un comportamiento de seguimiento de señales, mientras que el tono generaba un comportamiento de seguimiento de objetivos, y que la palanca adquiría las propiedades de un reforzador condicionado, mientras que el tono no, los sujetos podían elegir entre dos alternativas: En una de estas alternativas se presentaba una palanca durante 8 s que era seguida de la entrega de un reforzador, caso contrario, la otra alternativa estaba asociada con la aparición de un tono durante 8 s después del cual se entregaba un reforzador. En esta condición, los sujetos tenían una fuerte preferencia (95%) por la alternativa asociada con la palanca. Posteriormente, disminuyeron la probabilidad de reforzamiento después de la presentación de palanca a .50, y reportaron que los sujetos prefirieron esta alternativa (alrededor del 75% de los ensayos), mostrando una elección subóptima.

Es posible que la diferencia entre el estudio de Beckman y Chow (2015) y aquellos que no han encontrado una relación entre la elección subóptima y el valor incentivo de los estímulos (Martínez et al., 2017; López , Alba y Orduña 2018; Exp 1, 2, 3 y 4 de la presente tesis) se debe al hecho de que en el primer estudio las alternativas diferían en el valor incentivo de los estímulos asociados, mientras que en el segundo grupo de estudios el valor incentivo era el mismo en ambas alternativas, por lo que es posible que el efecto que tiene el valor incentivo de los estímulos se cancele.

En el presente experimento, manipulamos el valor incentivo de las alternativas Discriminativas y No Discriminativas empleando palancas en una de ellas y luces en la otra, para evaluar su impacto en la elección subóptima (Experimento 6A). Además, esta manipulación se repitió empleando dos grupos de sujetos, que diferían en su capacidad de

atribución de valor incentivo (experimento 6B), bajo la hipótesis de que la influencia del valor incentivo sería mayor en el grupo de ratas clasificadas como seguidoras de señales.

Método

Sujetos

Se usaron ocho ratas macho Wistar obtenidas del vivero del Instituto de Fisiología Celular de la UNAM. Todos los sujetos eran experimentalmente ingenuos, de aproximadamente 60 días de edad. Los sujetos fueron alojados y mantenidos en las mismas condiciones que en los experimentos previos.

Aparatos

Los aparatos utilizados en este experimento fueron los mismos que en los Experimentos anteriores.

Procedimiento

Los sujetos fueron expuestos a las mismas condiciones de habituación, entrenamiento a comedero y entrenamiento al nosepoke con las mismas características que en los experimentos anteriores.

Procedimiento de elección subóptima.

Las ratas se dividieron en dos grupos, que diferían en el orden de exposición a las dos condiciones del experimento. Cada grupo fue entrenado en una versión diferente del procedimiento de elección subóptima (Fig. 31): en una versión (Panel A), la alternativa Discriminativa se asoció con un estímulo con alto valor incentivo (una palanca), mientras que en la alternativa No Discriminativa se asoció con un estímulo con bajo valor incentivo (una luz). En la otra versión (Panel B), estas relaciones se invirtieron.

En ambas versiones, se presentaron 40 ensayos forzados (20 para cada alternativa) y 20 ensayos de elección. Para los ensayos forzados de la alternativa Discriminativa, una respuesta en el nosepoke; resultaba en el 50% de los ensayos en la presentación del estímulo asociado con el reforzador (una palanca para el grupo 1, una luz para el grupo 2) y el reforzador se entregó 10 s más tarde. En el otro 50% de los ensayos se presentó un blackout y 10 s más tarde el ensayo terminó sin entregar ningún reforzador para ambos grupos.

En los ensayos forzados de la alternativa No Discriminativa, después de una respuesta al nosepoke, se presentó un estímulo con probabilidad 1.0 (luz para el Grupo 1, una palanca para el grupo 2) y 10 segundos más tarde se entregó un reforzador con probabilidad .75.

En los ensayos de elección, se iluminaron ambos nosepokes y la primera respuesta en cualquiera de los dos presentó las mismas contingencias que se describieron para los ensayos forzados. Cuando finalizó un ensayo, se presentó una ITI de 10 s durante el cual no se presentó ningún estímulo, y se inició un nuevo ensayo después. El tipo de ensayo se asignó de forma cuasialeatoria con la única restricción de que no se presentaran más de 4 ensayos del mismo tipo (ensayos forzados Discriminativos, forzados No Discriminativos o de elección) secuencialmente. Para ambos grupos, la probabilidad global de reforzamiento para la alternativa Discriminativa fue del 50% y del 75% para la alternativa No Discriminativa.

Después de 40 sesiones, la relación entre el tipo de alternativa (Discriminativa y No Discriminativa) y el valor incentivo de los estímulos (alto, bajo) se revirtió durante 40 sesiones adicionales. Específicamente, el grupo que en la condición anterior tenía el estímulo con alto valor de incentivo en la alternativa Discriminativa ahora experimentó el

estímulo con alto valor de incentivo en la alternativa No discriminativa (ver figura #, panel B) y viceversa. Un cambio adicional fue la ubicación de las alternativas Discriminativa y No Discriminativa (panel frontal o panel trasero). En resumen, las condiciones A y B se realizaron en el orden A-B para el Grupo 1 y B-A para el grupo 2.

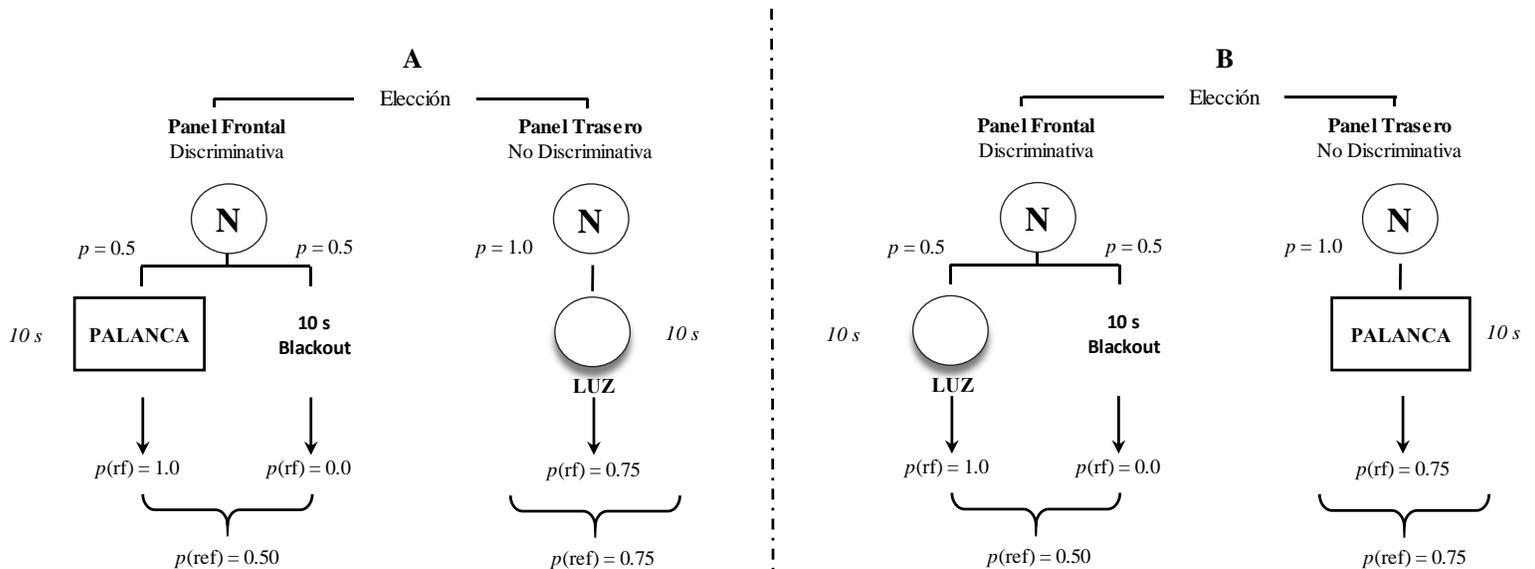


Figura 31. Diagramas de las dos versiones del procedimiento de "elección subóptima" empleado en los presentes experimentos. Dos alternativas, Discriminativa y No Discriminativa se presentaron en paneles opuestos. En la primera versión (Panel A) elegir la alternativa Discriminativa resultaba en la presentación de una palanca con $p=.5$, durante 10 s después de los cuales se entregó un reforzador o se presentó un Blackout con $p=.5$, que predecía la ausencia de reforzamiento. Elegir la alternativa No Discriminativa resultó en la presentación de una luz en todos los ensayos, asociada con una probabilidad de reforzamiento de .75. En la segunda versión (Panel B) elegir la alternativa Discriminativa resultó en la presentación de una luz con $p = .5$, que era un predictor confiable de reforzamiento, o un Blackout con $p = 0.5$, que predecía la ausencia de reforzamiento. La elección de la alternativa No Discriminatoria resultó en la presentación de una palanca asociada con una probabilidad de reforzamiento de .75.

Resultados

La Figura 32 muestra el promedio de la proporción de elección para la alternativa Discriminativa para ambos grupos durante las 40 sesiones de cada condición. Durante las últimas cinco sesiones de cada condición, la proporción de elección para la alternativa Discriminativa fue mayor cuando el estímulo con alto valor incentivo estaba en la alternativa Discriminativa que cuando este estímulo se asoció con la alternativa No Discriminativa ($.49 \pm .05$ y $.41 \pm .07$, para los grupos 1 y 2, respectivamente) que cuando este estímulo se asoció con la alternativa No Discriminativa ($.15 \pm .01$ y $.10 \pm .01$).

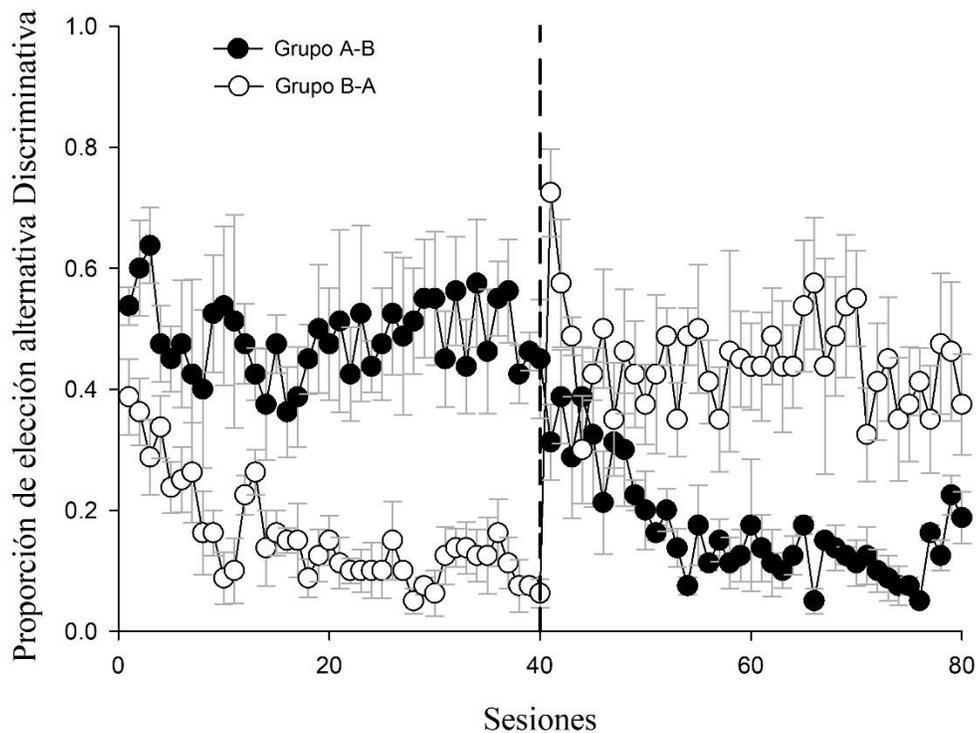


Figura 32. Experimento 6: Promedio de la proporción de elección de la alternativa Discriminativa a lo largo de las 40 sesiones de la condición A y las 40 sesiones de la condición B para cada uno de los grupos. Las condiciones A y B se llevaron a cabo en el orden A-B para el Grupo 1 indicado por círculos negros y B-A para el Grupo 2 indicado por círculos blancos. La línea vertical señala la transición de las condiciones.

Un ANOVA de medidas repetidas exploró el efecto Condición (A, B) y el orden en que se llevaron a cabo las condiciones para los diferentes grupos (A-B, B-A). Los resultados indicaron que el factor Condición fue significativo ($F(1, 6) = 47.56$, $p < 0.001$), mientras que el factor Orden y su interacción con el factor Condición no lo fueron ($F(1, 6) = 2.27$, $p = 0.18$ y $F(1, 6) = 0.06$, $p = 0.82$, respectivamente).

Discusión

El propósito del Experimento 6 era probar la hipótesis de que, en el procedimiento de elección subóptima, las ratas serían sensibles al valor incentivo de los estímulos asociados a las diferentes alternativas. Los resultados demostraron que cualquier alternativa que tuviera los estímulos con alto valor incentivo, tenía un fuerte impacto en su preferencia. Este efecto se mostró intra y entre grupos.

Sin embargo, el valor incentivo atribuido a un estímulo no depende exclusivamente de las propiedades físicas del estímulo, sino de su interacción con diferentes factores, por ejemplo, la especie a la que pertenece el organismo y el fenotipo particular de un organismo dentro de una especie determinada (Meyer et al., 2012). En relación con el último factor, se ha demostrado repetidamente que las ratas difieren en su capacidad de atribuir valor incentivo a las palancas empleadas en el procedimiento de Aproximación Pavloviana Condicionada. En este procedimiento, una palanca se presenta de forma contingente 8 s antes de que se entregue un reforzador, y como consecuencia, la palanca evoca respuestas condicionadas. Sin embargo, las respuestas condicionadas difieren entre los sujetos y aproximadamente el 35% de ellos se aproximan y entran en contacto con la palanca (seguidores de señales), el 30% se acerca al dispensador de comida (seguidores de objetivos) y el 35% restante son intermedios (Flagel et al., 2011; Meyer et al., 2012). Dentro de este contexto, es razonable plantear la hipótesis de que el sesgo encontrado en el

presente experimento debería ser más fuerte para las ratas que tienen una mayor capacidad de atribución de incentivos (seguidoras de señales). La evaluación de esta hipótesis es el objetivo del siguiente experimento.

Método

Sujetos

Se usaron trece ratas macho Wistar de aproximadamente 180 días de edad obtenidas del vivero del Instituto de Fisiología Celular de la UNAM. Los sujetos fueron alojados y mantenidos en las mismas condiciones que en los experimentos previos. Los sujetos habían sido clasificados previamente de acuerdo con su capacidad de atribución de incentivo (López et al., (2018) y tenían experiencia en el procedimiento de elección subóptima (para una descripción detallada de la categorización de acuerdo con su índice de PCA, y de la experiencia en el procedimiento, ver López et al. (2018)). Seis de estos sujetos cumplieron los requisitos para ser clasificados como seguidores de señales, mientras que los otros 7 como seguidores de objetivos (Meyer et al. 2012). Todos los sujetos se mantuvieron en el 85% de su peso de alimentación libre.

Procedimiento de elección subóptima

Fases 1 y 2: Probabilidad global de reforzamiento .20 (Discriminativa) vs .50 (No Discriminativa)

Durante las fases 1 y 2, las condiciones A y B descritas previamente para el experimento 6 (ver figura 31) se presentaron sucesivamente a todos los sujetos en el mismo orden (A-B). La única diferencia con el experimento 6A fue que se modificaron las probabilidades globales de reforzamiento para las alternativas; específicamente, la probabilidad global de reforzamiento para la alternativa Discriminativa se redujo de .50 a

.20, y para la alternativa No Discriminativa de .75 a .50. Estas condiciones se realizaron durante 45 y 20 sesiones, respectivamente. Cuando los sujetos cambiaron de la fase 1 a la fase 2, hubo una reversión de las posiciones de las alternativas Discriminativas y No Discriminativas.

Fases 3 y 4: Probabilidad global de reforzamiento .50 (Discriminativa) vs .75 (No Discriminativa)

Para evaluar el efecto de la frecuencia de presentación de los estímulos con alto valor incentivo, se presentaron nuevamente las condiciones A y B, pero con un aumento en la probabilidad de aparición del estímulo con alto valor incentivo; La probabilidad global de reforzamiento para la alternativa Discriminativa se incrementó de .20 a .50, mientras que para la alternativa No Discriminativa se incrementó de .50 a .75. Al comienzo de cada fase, las posiciones de las alternativas se revirtieron.

Resultados

La Figura 33 muestra el promedio de la proporción de elección para la alternativa Discriminativa para ambos grupos durante todas las fases del experimento. Durante las últimas cinco sesiones de la fase 1, el promedio de la proporción de elección para la alternativa Discriminativa fue de $.47 \pm .09$ para los seguidores de señales y de $.25 \pm .04$ para los seguidores de objetivos. Esta diferencia fue estadísticamente significativa, de acuerdo con un ANOVA que evaluó el efecto de grupo (seguidores de señales versus seguidores de objetivos) y las sesiones sobre la proporción de elección para la alternativa Discriminativa; el factor grupo fue estadísticamente significativo ($F(1, 11) = 5.43, p = 0.03$), mientras que el factor sesiones ($F(4,44) = 0.70, p = 0.60$) y la interacción Grupo-

Sesiones no fueron significativos ($F(4, 44) = 0.75, p = 0.57$). Al final de la fase 2, la preferencia cambió en la dirección de la alternativa No Discriminativa, que ahora presentaba el estímulo con alto valor incentivo. Durante las últimas cinco sesiones, la proporción de elección para la alternativa Discriminativa fue de $.06 \pm .02$ para los seguidores de señales y de $.13 \pm .02$ para los seguidores de objetivos. Un Anova reportó que el factor Grupo fue estadísticamente significativo ($F(1, 11) = 7.99, p = 0.02$), mientras que el factor Sesiones y su interacción con el Grupo no lo fueron ($F(4, 44) = 0.16, p = .96$ y $F(4, 44) = 0.68, p = .60$, respectivamente). Al final de la Fase 3, la proporción de elección para la alternativa Discriminativa fue mayor para los seguidores de señales ($.81 \pm .06$) que para los seguidores de objetivos ($.48 \pm .07$). Esta diferencia fue estadísticamente significativa ($F(1, 11) = 12.42, p = 0.004$) mientras que el factor sesiones y la interacción con el factor grupo no tuvieron efectos significativos ($F(4,44) = 0.17, p = 0.95$ y $F(4, 44) = 0.83, p = 0.510$, respectivamente). Finalmente, durante los últimos cinco sesiones de la Fase 4, ambos grupos mostraron una baja preferencia por la alternativa Discriminativa que ahora estaba asociada con un estímulo con bajo valor incentivo ($.15 \pm .04$ y $.17 \pm .05$ para los seguidores de señales y seguidores de objetivos, respectivamente). Un ANOVA reveló que no hubo efectos significativos del factor Grupo ($F(1, 11) = 0.09, p = 0.77$), Sesiones ($F(4, 44) = 0.19, p = 0.94$, ni su interacción ($F(4, 44) = 2.00, p = 0.11$).

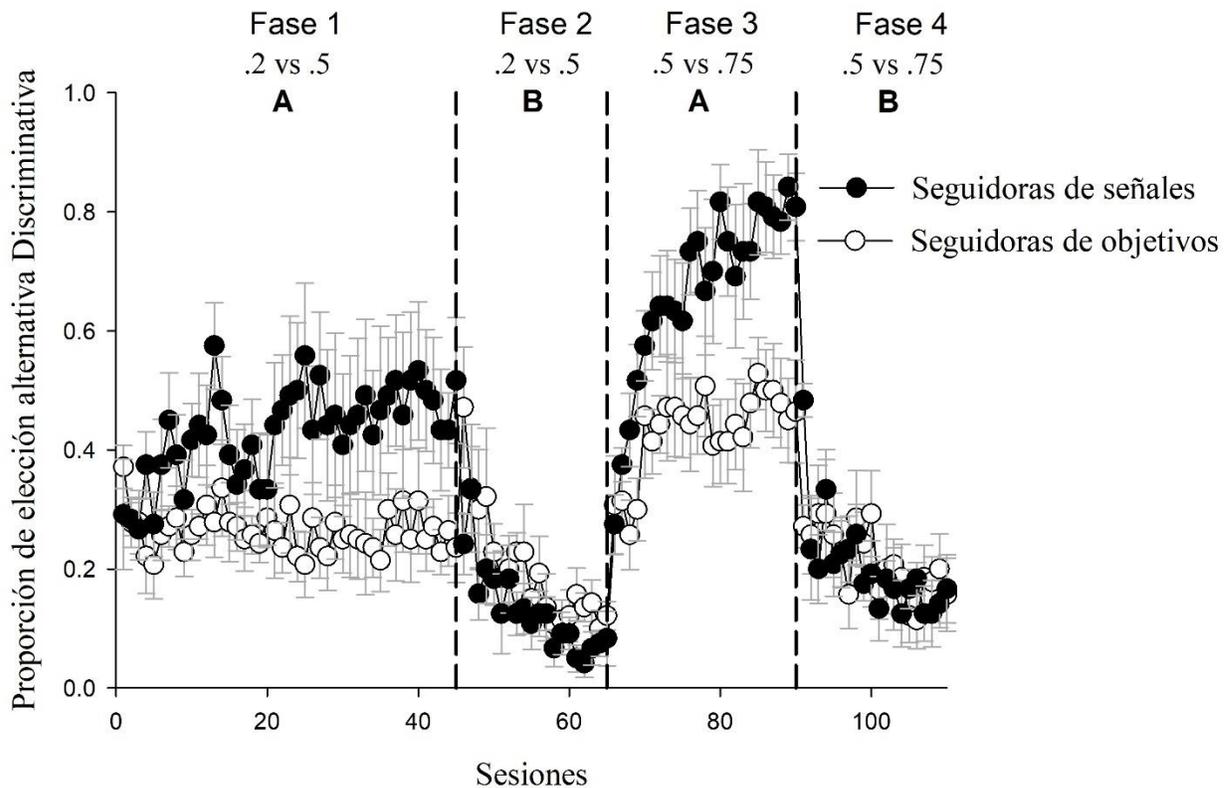


Figura 33. Experimento 7: Promedio de la proporción de elección de la alternativa Discriminativa a lo largo de las 4 fases del presente experimento. Las líneas verticales punteadas señalan la transición de fases; durante las fases 1 y 2, la probabilidad global de reforzamiento fue de .20 y .50 en las alternativas Discriminativas y No Discriminativas, respectivamente. Las condiciones A y B se presentaron sucesivamente a todos los sujetos en el mismo orden. En las fases 3 y 4, las condiciones A y B se volvieron a presentar sucesivamente, pero las probabilidades globales de reforzamiento se modificaron (.50 y .75 para las alternativas Discriminativas y No Discriminativas, respectivamente). Los círculos negros representan las preferencias de las ratas clasificadas como seguidoras de señales y los círculos blancos de las seguidoras de objetivos.

La figura 34 muestra, para cada grupo en cada una de las cuatro fases, el promedio del número de respuestas a la palanca y entradas al comedero en presencia del estímulo con alto valor incentivo. Se puede observar que el número de respuestas a la palanca fue mayor para los seguidores de señales que para los seguidores de objetivos, en el caso de las entradas al comedero esta relación se invirtió. Estas diferencias que son características de la clasificación de acuerdo a su atribución de incentivo se mantuvieron después de un extenso entrenamiento en el procedimiento de elección subóptima.

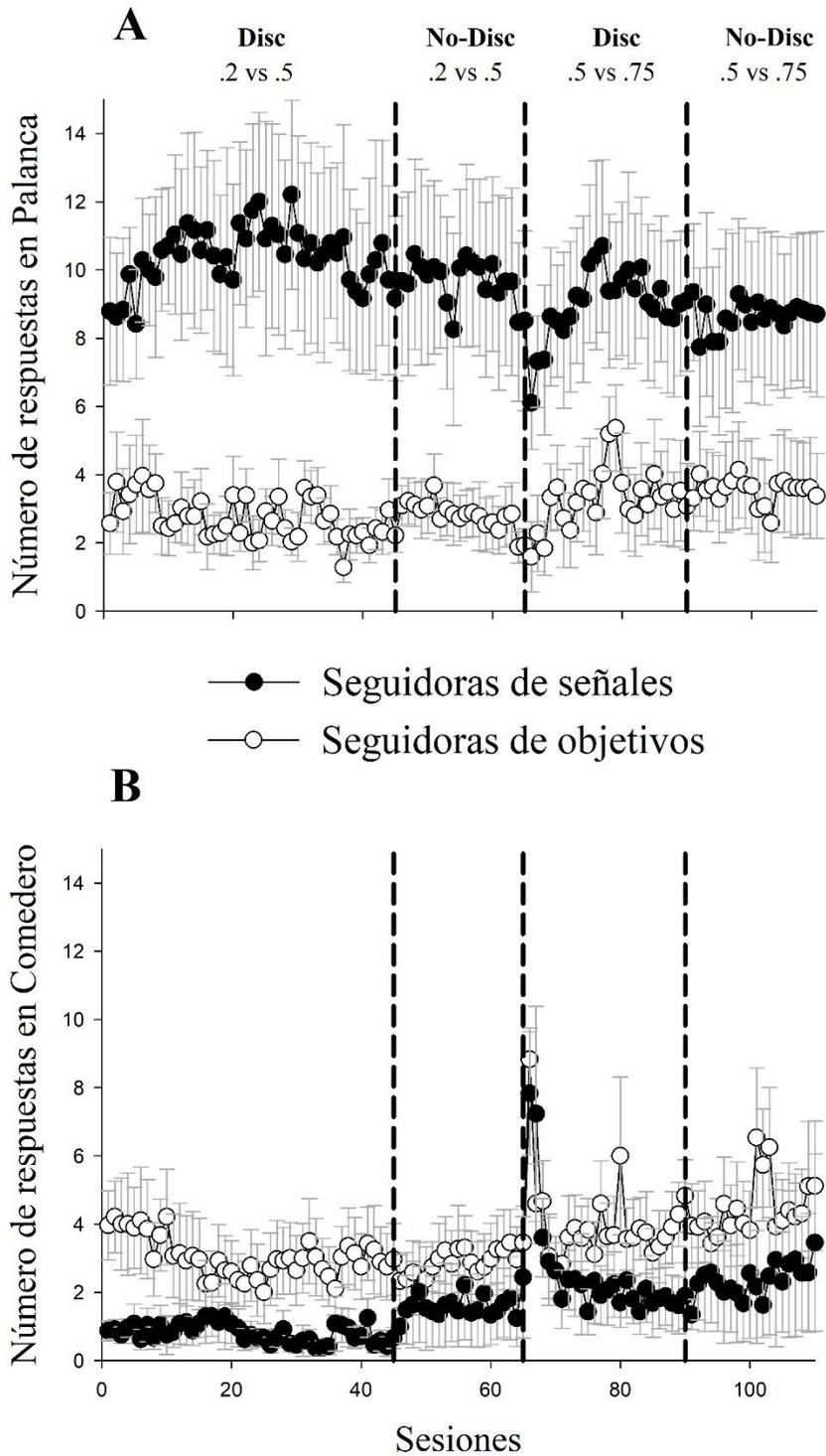


Figura 34. Experimento 7: Número de repuestas a la palanca y entradas al comedero en presencia del estímulo con alto valor incentivo en los ensayos forzados en todas las sesiones durante las 4 fases del presente experimento. El panel A muestra el promedio de respuestas a la palanca y el panel B el promedio de respuestas al comedero. Los círculos negros representan a las seguidoras de señales y los círculos blancos representan a las seguidoras de objetivos. Las líneas verticales punteadas señalan la transición de fases.

Un Anova del número de respuestas a la palanca informó diferencias estadísticamente significativas del factor Grupo ($F(1, 11) = 7.20, p = 0.02$), pero no en el factor Condición ($F(3, 33) = 0.14, p = 0.94$) ni la interacción grupo-condición ($F(3, 33) = 0.91, p = 0.45$). Un Anova separado informó que la diferencia en el número de entradas al comedero no fue estadísticamente significativa entre los grupos ($F(1, 11) = 1.89, p = 0.20$) o entre las condiciones ($F(3, 33) = 2.34, p = 0.09$). La interacción tampoco fue significativa ($F(3, 33) = .14, p = 0.93$).

Discusión

Los resultados del experimento 6B fueron consistentes con los resultados del Experimento 6A en el sentido de que la preferencia por una alternativa aumenta cuando sus estímulo asociado tenía alto valor incentivo. Además, aunque el efecto se observó en ambos grupos, fue más fuerte para los seguidores de señales.

De acuerdo con nuestra hipótesis, el efecto de la preferencia por la alternativa Discriminativa dependía de la presentación del estímulo Discriminativo con mayor valor incentivo: cuando se presentó en el 50% de los ensayos en los que se eligió la alternativa Discriminativa, el sesgo para elegir esta alternativa fue más fuerte que cuando el estímulo Discriminativo se presentó solo en el 20% de los ensayos. La mayor tasa de presentación de este estímulo fue suficiente para generar una elección subóptima en los seguidores de señales durante la fase 3, como lo demuestra su proporción de elección para la alternativa Discriminativa que es significativamente superior al 50%.

Finalmente, la diferencia en el número de respuestas a la palanca y entradas al comedero en presencia de los estímulos con alto valor incentivo en ensayos forzados en ambos grupos mostró que las diferencias individuales en la capacidad de atribución de incentivos (que definieron entre los grupos) se mantuvieron durante todo el experimento.

Discusión general

El procedimiento de elección subóptima se ha propuesto como un modelo animal de algunos aspectos del comportamiento de la adicción al juego en humanos (Molet et al., 2012) y parece ideal para proporcionar información sobre la relevancia que tienen las señales que predicen un reforzador, o su ausencia (para una revisión, ver T. R. Zentall, 2016). Apoyando la validez aparente del procedimiento, tanto las palomas como los humanos están más influenciados por los estímulos que predicen las ganancias que por los que predicen las pérdidas. Se ha propuesto que el comportamiento de los jugadores patológicos está influenciado por una atribución diferencial del valor de incentivo a los estímulos asociados con las ganancias y pérdidas, porque los estímulos presentados cuando se obtienen victorias son más importantes que cuando se experimentan pérdidas (Dixon et al., 2014). Aunque impulsadas por un mecanismo diferente (Vasconcelos, Machado, & Pandeirada, 2018), las palomas asignan un valor desproporcionadamente grande al estímulo que predice un reforzador e ignoran el estímulo que predice confiablemente la ausencia de un reforzador, una combinación que resulta en una elección subóptima. El hecho de que las ratas, otra especie que se ha estudiado extensamente con este procedimiento, generalmente no participan en una elección subóptima mientras están influenciadas notablemente por el estímulo predictor de no reforzamiento, sugiere que las terapias para disminuir la adicción al juego deberían centrarse en aumentar la atención en este estímulo y al mismo tiempo disminuir la atención a las señales relacionadas con las ganancias (es decir, cambio en el sesgo de atención; Hønsi, Mentzoni, Molde, & Pallesen, 2013).

Además del valor predictivo de los estímulos que indican la presentación de un reforzador o su ausencia, se ha sugerido que el valor incentivo de esos estímulos juega un papel importante en la determinación de las preferencias tanto en el procedimiento de

elección subóptima (Chow et al., 2017; Zentall, Smith, & Beckmann, 2019) como en otros protocolos de elección (Barrus & Winstanley, 2016; Beckmann & Chow, 2015; Smith, Hofford, Zentall, & Beckmann, 2018). Como se revisó anteriormente, se ha demostrado que cuando las ratas eligen entre un estímulo asociado con la entrega un reforzador con $p = 1$, pero que tiene bajo valor incentivo, y otro estímulo asociado con la entrega de un reforzador con $p = .5$, pero que tiene alto valor incentivo, prefieren este último, por lo tanto, las ratas muestran elecciones subóptimas (Beckmann y Chow 2015). En este contexto, el presente conjunto de experimentos fue diseñado para analizar la relevancia de esta variable en el procedimiento de elección subóptima. Si bien los estudios previos han aumentado el valor incentivo de los estímulos en el procedimiento de elección subóptima (Martínez et al., 2017; López, Alba y Orduña 2018, Exp 1, 2, 3 y 4 presente tesis), la particularidad de la presente propuesta es una manipulación asimétrica del valor incentivo entre las alternativas Discriminativas y No Discriminativas, similar a la investigación previa con otros protocolos de elección (Beckmann y Chow 2015).

El principal hallazgo del experimento 6A fue que la preferencia de las ratas estaba sesgada hacia la alternativa que presentaba estímulos con alto valor de incentivo. Específicamente, las ratas tendieron a la subóptimalidad cuando el estímulo con alto valor incentivo se presentó en la alternativa Discriminativa y tendieron a la optimalidad cuando este estímulo se asoció con la alternativa No Discriminativa. Estos efectos se mostraron intra y entre grupos. Es posible que los estudios previos que dispusieron de un estímulo con alto valor incentivo en todos los estímulos discriminativos no lograron encontrar un comportamiento subóptimo porque los sujetos atribuyeron el mismo valor de incentivo a los estímulos tanto en las alternativas Discriminativas como las No Discriminativas.

Nuestros resultados también están relacionados con la línea de investigación que explora la relación entre las diferencias individuales en la capacidad de atribución de valor de incentivo y la ejecución de conductas desadaptativas (Flagel et al., 2011; Flagel et al., 2010). El principal hallazgo del experimento 6B fue que los sujetos clasificados como seguidores de señales –los sujetos con mayor capacidad de atribución de incentivo– mostraron un mayor nivel de comportamiento subóptimo en las condiciones que se presentaban los estímulos con alto valor de incentivo en la alternativa Discriminativa y que el nivel de esta preferencia se vio afectada por la tasa de presentación del estímulo Discriminativo asociado con el reforzador. Este resultado se suma a una gran cantidad de investigaciones en las que se ha demostrado que los sujetos clasificados como seguidores de señales presentan niveles más altos de conducta desadaptativa que los sujetos clasificados como seguidores de objetivos (para revisión, ver, Meyer, Tripi, Tomie, & Morrow, 2018). También es notable el hecho de que en una de las condiciones en el que el alto valor incentivo se asoció con la alternativa No Discriminativa, el comportamiento de los sujetos de clasificados como seguidores de señales estaba significativamente más cerca de la optimalidad que los sujetos clasificados como seguidores de objetivos. Estos resultados en conjunto sugieren que, en el procedimiento de elección subóptima, el valor incentivo asociado con una alternativa es un factor determinante más fuerte de preferencia, al menos en ratas, que la información proporcionada por esa alternativa.

El procedimiento de elección subóptima, como se empleó aquí, parece capturar un aspecto esencial del comportamiento real de la adicción al juego en humanos que no está presente en el procedimiento tradicional, en el sentido de que el valor incentivo diferencial de los estímulos asociados con los diferentes comportamientos disponibles (es decir, apostar o no apostar en el caso de humanos); por ejemplo, las señales contextuales

asociadas con ir al casino son claramente más salientes (Finlay, Kanetkar, Londerville, & Marmurek, 2006) que las señales asociadas con comportamientos alternativos. La modificación del procedimiento presentado aquí permite la posibilidad de estudiar el impacto de las señales contextuales, además del impacto de las señales específicas de resultados o anticipatorias (Barrus, Cherkasova, & Winstanley, 2015).

En resumen, los resultados de ambos experimentos proporcionan evidencia de una fuerte relación entre el valor incentivo asociado con una alternativa y la preferencia en el procedimiento de elección subóptima, similar a lo observado en otros protocolos de elección (Beckmann y Chow 2015; Barrus y Winstanley 2016). Dado que esta relación es consistente con el vínculo que se ha demostrado entre el juego patológico y la atribución de valor incentivo en humanos (Robinson y Berridge 2008), la posibilidad de complementar las terapias farmacológicas para la adicción al juego (Dagher & Robbins, 2009) con terapias basadas en la modificación conductual debe considerar el valor incentivo de los diferentes resultados y las diferentes alternativas.

CAPITULO VIII

DISCUSIÓN GENERAL

En una amplia revisión de la literatura sobre el procedimiento de elección subóptima se llegó a las siguientes conclusiones: a) una gran cantidad de datos respalda la noción de que las palomas se desempeñan de manera subóptima en este procedimiento (McDevitt et al., 2016; Zentall et al., 2019); b) dicha subóptimalidad es resultado de la sobrevaloración del estímulo que predice la entrega de un reforzador y el poco efecto que tiene el estímulo que predice la ausencia de reforzamiento (Vasconcelos et al., 2018; Zentall, 2016); c) dicho comportamiento ha adquirido gran relevancia porque ha sido considerado como un modelo animal que permite aproximarse al estudio de la adicción al juego que es padecida por una proporción significativa de la población humana (Molet et al., 2012; Zentall & Stagner, 2011); d) la generalidad entre especies ha sido cuestionada en una serie de estudios en los cuales se mostró que las ratas se comportan de manera óptima en el procedimiento de elección subóptima (Martínez et al., 2017; Trujano & Orduña, 2015); e) se han propuesto dos hipótesis para explicar estas diferencias entre especies, la primera describe que estas diferencias se deben a una mayor sensibilidad a la inhibición condicionada en ratas que en palomas. Específicamente, las palomas en un inicio son sensibles al inhibidor condicionado, sin embargo, al transcurso de las sesiones no consideran a este estímulo en su valoración de la alternativa discriminativa (subóptima) (Laude, Stagner, et al., 2014). En contraste, las ratas lo consideran con gran fuerza en todas las sesiones del procedimiento de elección subóptima (Trujano et al., 2016). La segunda hipótesis sugiere que estas diferencias se deben al valor incentivo de los estímulos empleados con las dos especies. Específicamente, para las palomas las teclas iluminadas son poderosos reforzadores condicionados con alto valor incentivo y para las ratas las luces y los tonos no presentan tal característica (Chow et al., 2017; Zentall et al., 2019); f) Para la primera hipótesis antes mencionada, se han presentado datos contundentes que respaldan

esta hipótesis (Stagner et al., 2011; Trujano et al., 2016). Contrario, a la evidencia presentada con relación a la hipótesis del valor incentivo de los estímulos, específicamente, en un primer estudio (Chow et al., 2017), se reportó que cuando los estímulos predictores de reforzamiento tuvieron valor incentivo, las ratas se comportaron de manera subóptima. Sin embargo, a pesar de que este resultado apoyó la hipótesis del valor incentivo, el procedimiento utilizado no empleó un inhibidor condicionado. Por tal motivo en un estudio posterior con el fin de contrastar estas hipótesis se agregó a la opción discriminativa una estímulo con valor incentivo (palanca) predictor de la ausencia de reforzamiento. El resultado contundente fue un alto nivel de optimalidad en las ratas (Martínez et al., 2017).

Como se pudo observar en las conclusiones, el fenómeno de elección subóptima es muy robusto en palomas, pero no ha sido replicado en ratas. Y a pesar de que en la literatura se han planteado hipótesis para explicar estas diferencias entre especies a la fecha no es clara la razón de estas diferencias.

A pesar de los resultados mixtos reportados en los estudios que usaron estímulos con valor incentivo, antes de aceptar la idea de que el inhibidor condicionado tiene un impacto mayor sobre la preferencia que el valor incentivo de los estímulos, era crucial evaluar la generalidad de este resultado en ratas y palomas. Además de evaluar la generalidad de la conducta óptima en ratas y subóptima en palomas ante estímulos con alto y bajo valor incentivo, era importante evaluar la generalidad de la optimalidad de las ratas en el procedimiento que ha sido propuesto como un modelo animal que permite aproximarse al estudio de la adicción al juego.

En la primera parte de la presente tesis en los primeros tres experimentos se derivaron tres variables de los estudios previos que evaluaron el impacto del valor incentivo de los estímulos en el procedimiento de elección subóptima y se probaron sus efectos sobre

el desarrollo de la conducta subóptima en ratas. Las variables fueron la probabilidad global de reforzamiento asociada a cada una de las alternativas, la presencia de un estímulo predictor de no reforzamiento (inhibidor condicionado) y las características de las alternativas (Discriminativa y No Discriminativa). En el experimento cuatro se evaluó la generalidad de la conducta óptima de las ratas al utilizar el procedimiento de elección subóptima que ha sido considerado como un modelo animal que permite aproximarse al estudio de la adicción al juego.

En la segunda parte de la presente tesis se manipulo el valor incentivo de los estímulos de las alternativas del procedimiento de elección subóptima. Específicamente, en el experimento cinco se emplearon tonos (estímulos con bajo valor incentivo) en el procedimiento de elección subóptima usando palomas como sujetos experimentales. En todos los estudios que han manipulado el valor incentivo de los estímulos en el procedimiento de elección subóptima en ratas han mantenido el mismo valor en todos los estímulos y dada la evidencia en otros protocolos de elección donde se ha demostrado que el valor incentivo en una alternativa sesga la preferencia de las ratas por esa alternativa (Beckmann & Chow, 2015), en el experimento 6A se manipuló el valor incentivo de los estímulos de las alternativas Discriminativa y No Discriminativa presentando palancas (estímulos con alto valor de incentivo) en una de ellas y luces (estímulos con bajo valor de incentivo). Adicionalmente, se evaluó a ratas clasificadas como Seguidoras de señales con esta manipulación (experimento 6B).

Este capítulo es un intento de resumir e integrar los resultados de los seis experimentos de la presente tesis con respecto a los supuestos de las hipótesis que se han propuesto para explicar las diferencias entre especies en el procedimiento de elección subóptima.

Probabilidad global de reforzamiento asociada con la Alternativa Discriminativa y No Discriminativa

Como se describió en el capítulo 1 en los experimentos de elección subóptima que usaron palomas como sujetos experimentales se ha reportado una relación importante entre la probabilidad global de reforzamiento de las alternativas Discriminativa y No Discriminativa y el desarrollo de la conducta subóptima. En particular, la preferencia por la alternativa subóptima ha sido mayor cuando las probabilidades globales de reforzamiento en las alternativas Discriminativas vs No Discriminativas fueron .2 vs .5 (Stagner & Zentall, 2010) que cuando fueron .5 vs .75. (Gipson et al., 2009). En los estudios con ratas que han evaluado la hipótesis del valor incentivo de los estímulos en el procedimiento de elección subóptima se han empleado ambas probabilidades, en particular en el único estudio donde se ha reportado conducta subóptima en ratas en la evaluación de esta hipótesis emplearon las probabilidades .25 vs .50 (Chow et al., 2017), en contraste, en el estudio que muestra evidencia en contra de esta hipótesis se emplearon las probabilidades globales de reforzamiento de .5 vs .75 (Martínez et al., 2017).

A pesar de que las diferencias encontradas en estos estudios se atribuyeron a la ausencia de un inhibidor condicionado, es importante evaluar la posibilidad de que las diferencias en las probabilidades globales de reforzamiento no fuera la variable responsable de estos resultados contradictorios. En los experimentos 1 y 3 de la presente tesis las ratas se evaluaron en el procedimiento de “elección subóptima” usando las probabilidades globales de reforzamiento de .2 vs .5 en la alternativa Discriminativa frente a la No Discriminativa respectivamente. Los resultados indicaron que las ratas se comportaron de manera óptima, eligiendo la alternativa asociada con la mayor probabilidad global de

reforzamiento, tal vez incluso más que con la proporción de .5 vs .75 que se empleó en investigaciones previas (López, Alba y Orduña 2018; Martínez et al.2017).

De esta manera se puede atribuir que la diferencia importante en estos estudios fue la ausencia del inhibidor condicionado y no las diferencias en las probabilidades globales de reforzamiento asociadas a las alternativas Discriminativa y No Discriminativa. Por último, es importante resaltar que estos resultados aportan evidencia de la generalidad de la conducta óptima en ratas en el procedimiento de elección subóptima usando estímulos con valor incentivo y diferentes probabilidades globales de reforzamiento.

Presentación de un estímulo predictor de la ausencia de reforzamiento y el desarrollo de la conducta subóptima en ratas.

En la mayoría de los experimentos de elección subóptima con palomas se presentan dos estímulos, uno predictor de la entrega de un reforzador y otro estímulo que predice la ausencia de un reforzador en la alternativa Discriminativa y dos estímulos diferentes los dos asociados a la misma probabilidad de reforzamiento en la alternativa No Discriminativa, bajo estas condiciones las palomas muestran una fuerte preferencia por la alternativa Discriminativa (subóptima) (Zentall et al., 2019). Dada la evidencia mostrada en el capítulo 2, se puede concluir que las palomas ignoran el estímulo predictor de la ausencia de una recompensa, en la valoración de la alternativa subóptima, por esta razón, Smith et al., (2016) modificaron el procedimiento clásico de elección subóptima y propusieron un diseño simplificado de este procedimiento que omitió la presentación del estímulo predictor de no reforzamiento. Con esta modificación del procedimiento clásico se ha reportado que

las palomas se comportan de manera subóptima igual que en el procedimiento clásico (Smith et al., 2017; Smith & Zentall, 2016; Zentall et al., 2017).

El único experimento que reporta conducta subóptima en ratas al utilizar estímulos con valor incentivo utilizó la versión modificada el procedimiento clásico de elección subóptima y presentó un blackout en lugar de un estímulo predictor de la ausencia de reforzamiento. Dada la relevancia que presumiblemente tiene el inhibidor condicionado en la conducta de las ratas en el procedimiento de elección subóptima, Martínez et al., utilizó el procedimiento clásico de elección subóptima y una palanca como estímulo predictor de la ausencia de reforzamiento. Bajo estas condiciones las ratas mostraron una fuerte preferencia por la alternativa óptima. Martínez et al., explicaron que las ratas en el estudio de Chow et al., se comportaron de manera subóptima debido a la ausencia de un estímulo que predijera de manera confiable la ausencia de una recompensa. Sin embargo, en estos dos estudios además de la diferencia en la presencia o ausencia de un inhibidor condicionado, también difieren en la versión del procedimiento de elección subóptima empleado, por tal motivo, es importante evaluar la posibilidad de que las diferentes versiones del procedimiento no sean la variable responsable del desarrollo de la conducta subóptima en ratas.

En el experimento 2 y 3 de la presente tesis, el estímulo asociado con ausencia de reforzamiento fue sustituido por un blackout. En el experimento 2 se utilizó la versión clásica del procedimiento de elección subóptima y en el experimento 3 se utilizó la versión simplificada. Estas condiciones proporcionaron una aproximación más cercana al procedimiento empleado por Chow et al. (2017). A pesar de esta manipulación, las ratas en los dos experimentos mostraron una fuerte preferencia por la alternativa óptima. Este resultado sugiere que la ausencia de un estímulo que prediga confiablemente la ausencia de

una recompensa no fue la variable relevante para generar un comportamiento subóptimo en el estudio de Chow et al. (2017). Los resultados de los presentes experimentos aportan evidencia de la generalidad de la optimalidad de las ratas en el procedimiento de elección subóptima bajo diversas condiciones, en particular, las ratas se comportan de manera óptima en las dos versiones del procedimiento de elección subóptima (con 2 y 4 estímulos) y la ausencia de un estímulo asociado explícitamente con la ausencia de reforzamiento no tiene un impacto en la preferencia de las ratas por la alternativa óptima.

Debido a los resultados de los primeros tres experimentos de la presente tesis, los hallazgos reportados por Chow et al., donde muestran una clara preferencia por la alternativa Discriminativa (subóptima) cuando se emplean estímulos con valor incentivo, pueden atribuirse a una explicación diferente a lo propuesto originalmente por Martínez et al., que atribuían este resultado a la ausencia de un inhibidor condicionado.

En el experimento reportado por Chow et al., 2017, en la primera condición se emplearon las probabilidades globales de reforzamiento de .5 vs .5 en la alternativa Discriminativa y No Discriminativa respectivamente. Los autores reportaron que las ratas desarrollaron una fuerte preferencia por la alternativa Discriminativa. Luego redujeron la probabilidad global de reforzamiento para la alternativa Discriminativa de .5 a .25. y bajo estas condiciones las ratas mantuvieron su preferencia por la alternativa Discriminativa. De esta manera los autores interpretaron sus hallazgos en términos de elección subóptima debido a que las ratas prefirieron la alternativa que presentaba un estímulo que predecía la presencia de una recompensa o un blackout y una baja probabilidad global de reforzamiento sobre otra alternativa que presentaba un estímulo asociado con una probabilidad global de reforzamiento mayor. Sin embargo, debido a que en el estudio de Chow et al., las ratas en la primera condición ya habían desarrollado una fuerte preferencia

por la alternativa Discriminativa cuando las probabilidades globales de reforzamiento entre las dos alternativas eran las mismas, no se puede atribuir al desarrollo de la conducta subóptima, de tal manera que Chow et al., solo mostraron la persistencia en la preferencia por la alternativa Discriminativa a pesar de que esta alternativa ahora estaba asociada con una menor probabilidad global de reforzamiento.

Generalidad de la conducta óptima de las ratas en el procedimiento de elección subóptima propuesto como un modelo animal de adicción al juego

La mayoría de los estudios de elección subóptima utilizan la versión clásica del procedimiento que involucra alternativas que difieren en la información que dan los estímulos y las probabilidades globales de reforzamiento. Sin embargo, existe otro procedimiento que además de presentar estímulos que predicen confiablemente la entrega o ausencia de una recompensa, también están asociados a distintas magnitudes de reforzamiento; esto debido a que se ha propuesto que el procedimiento clásico de elección subóptima no es el mejor análogo del juego patológico y que un mejor procedimiento debería considerar que en el juego patológico, la diferencia más relevante entre las alternativas (jugar versus no jugar) no son las probabilidades globales de reforzamiento, si no la diferencia en las magnitudes de reforzamiento.

Zentall y Stagner (2011) propusieron que este procedimiento es análogo a la conducta de juego patológico debido a que los jugadores prefieren escoger una alternativa con una expectativa de una victoria bastante pequeña, y persisten en esa alternativa a pesar

de perder su dinero, debido a que los jugadores patológicos tienden a sobrevalorar sus ganancias y subvalorar sus pérdidas.

Como se describió en el capítulo 1 las palomas al ser evaluadas en este procedimiento desarrollan una clara preferencia por la alternativa subóptima. Sin embargo, a pesar de que este resultado es contundente, no se ha evaluado la generalidad de este resultado en ratas, que toma especial relevancia dadas las diferencias reportadas en la otra versión del procedimiento de elección subóptima. Los resultados del experimento 4 de la presente tesis muestran que las ratas a diferencia de las palomas se comportan de manera óptima en este procedimiento eligiendo la alternativa que a la larga las llevan a obtener una mayor cantidad de reforzamiento. Este resultado es importante porque muestra que el comportamiento óptimo presentado por las ratas en las dos versiones del procedimiento de elección subóptima es tan general y contundente como el comportamiento subóptimo presentado por las palomas en este procedimiento.

Un hallazgo importante de destacar es que las ratas mostraron una preferencia por la alternativa óptima a pesar de que los estímulos empleados en este experimento tenían valor incentivo. Este resultado al igual que los datos obtenidos en los experimentos 1, 2 y 3 de la presente tesis muestran que el valor incentivo de los estímulos no es la variable que promueve el desarrollo de la conducta subóptima en ratas.

Impacto del valor incentivo de los estímulos y el desarrollo de la conducta óptima en palomas

La hipótesis del valor incentivo de los estímulos propone que la diferencia entre ratas y palomas en el procedimiento de elección subóptima se debe a la relación que cada especie tiene con los estímulos utilizados, especialmente con su valor incentivo. Mientras que para las palomas las teclas iluminadas tienen valor incentivo, para las ratas éste no es el caso. A pesar de que esta hipótesis originalmente parecía explicar las diferencias entre especies en este procedimiento, los datos reportados en los experimentos 1,2,3 y 4 de la presente tesis muestran que al menos en ratas, el valor incentivo de los estímulos no es la variable que promueve el desarrollo de la conducta subóptima en este procedimiento.

A pesar de que la conducta subóptima desarrollada por las palomas se ha demostrado en más de una decena de estudios, la hipótesis del valor incentivo de los estímulos no ha sido evaluada, ya que en todos estos estudios se han empleado teclas iluminadas como estímulos. Los resultados del experimento 5 mostraron que usar tonos (estímulos con bajo valor incentivo) en el procedimiento de elección subóptima tiene un impacto importante en la preferencia de las palomas en este procedimiento. Específicamente, el 70% de los sujetos evaluados mostraron una clara preferencia por la alternativa óptima, un resultado contrario a todos los datos reportados previamente en la literatura que mostraban una fuerte preferencia de las palomas por la alternativa subóptima en este procedimiento.

En conclusión, los resultados del Experimento 5 muestran que modificar el valor incentivo de los estímulos resulta en la preferencia de la alternativa óptima y apoya la hipótesis de que el valor incentivo de los estímulos es la variable que promueve el desarrollo de la conducta subóptima en este caso en palomas.

Valor incentivo de los estímulos en las alternativas del procedimiento de elección subóptima

La revisión de la literatura presentada en el capítulo 6 muestra que la evidencia respecto a la hipótesis del valor incentivo de los estímulos como la variable responsable del desarrollo de la conducta subóptima en ratas en el procedimiento de elección subóptima hasta antes de los experimentos de la presente tesis era contradictoria.

Los resultados de los Experimentos 1,2,3 y 4 muestran de manera contundente que la relación entre el valor incentivo de los estímulos y el procedimiento de “elección subóptima” en ratas no es la variable que modula este comportamiento. Sin embargo, una explicación de estos resultados es que el valor incentivo de los estímulos fue el mismo en ambas alternativas, por lo que los sujetos le atribuyeron el mismo valor tanto a la alternativa Discriminativa como a la No Discriminativa. Se ha demostrado en otros procedimientos de elección (Beckmann & Chow, 2015; Smith et al., 2018) que los estímulos con alto valor incentivo asociados a una alternativa sesgan la preferencia por ella a pesar de no ser la alternativa que los lleve a los mejores resultados.

En el experimento 6A de la presente tesis se manipuló el valor incentivo de los estímulos de las alternativas Discriminativa y No Discriminativa presentando palancas (estímulos con alto valor de incentivo) en una de ellas y luces (estímulos con bajo valor de incentivo) en la otra, para evaluar su impacto en la elección subóptima. Adicionalmente, en se evaluó si las ratas clasificadas como Seguidoras de señales muestran una mayor tendencia a la elección subóptima en el "procedimiento de elección subóptima" con esta manipulación. Los resultados de estos experimentos muestran que la preferencia de las ratas estaba sesgada hacia la alternativa que presentaba estímulos con valor de incentivo.

Específicamente, las ratas tendieron a la subóptimalidad cuando el estímulo con valor incentivo se presentó en la alternativa Discriminativa y eligieron la alternativa óptima cuando este estímulo se asoció con la alternativa No Discriminativa. Es importante mencionar que a pesar de la preferencia por la alternativa Discriminativa (subóptima) fue mayor a lo reportado en estudios previos en este procedimiento y en los primeros experimentos de la presente tesis, esta preferencia no fue realmente diferente a la indiferencia. Sin embargo, los datos del experimento 6B mostraron que el sesgo por la alternativa Discriminativa cuando esta presentaba un estímulo con valor incentivo, fueron más fuertes en las ratas clasificadas como seguidoras de señales, que mostraron un nivel de comportamiento subóptimo similar al reportado extensamente con palomas en el procedimiento de elección subóptima.

Inhibición condicionada y valor incentivo de los estímulos como variables que explican la diferencia entre especies en el procedimiento de elección subóptima

Los resultados de la presente tesis nos permiten sugerir que las razones de las fuertes diferencias entre el comportamiento de las palomas y las ratas en el procedimiento de elección subóptima son consistentes con la idea de que las ratas son más sensibles al inhibidor condicionado que las palomas y que el valor incentivo de los estímulos tiene una influencia determinante en la conducta subóptima de palomas, pero en el caso de las ratas tiene un papel menos influyente.

La generalidad de la conducta óptima de las ratas en los experimentos de la presente tesis son contradictorios con la mayoría de las teorías de elección subóptima que se han

propuesto para explicar el comportamiento de los organismos en este procedimiento [SiGN (McDevitt et al., 2016), Contraste (Zentall, 2016), el Modelo Ecológico (Fortes, Machado y Vasconcelos, 2017; Vasconcelos, Monteiro, & Kacelnik, 2015)]. Sin embargo, es importante mencionar que estas teorías se desarrollaron considerando exclusivamente los datos de las palomas.

La nula influencia que estas teorías atribuyen al estímulo que predice la ausencia de una recompensa es comprensible, ya que, hasta hace poco, los únicos datos que las teorías de elección subóptima tenían que explicar eran los obtenidos con palomas. La creciente evidencia de que las ratas son óptimas y la demostración de que las ratas son persistentemente sensibles al inhibidor condicionado (Trujano, López, et al., 2016), plantean un desafío a este aspecto de las teorías, pero también sugieren algunas formas posibles de integrar esta diferencia entre especies en una teoría más general.

Las teorías más recientes de elección subóptima, desarrolladas después de que se informara la diferencia entre especies, han comenzado esta integración. Por ejemplo, el modelo teórico de información temporal (Cunningham & Shahan, 2018) propone que el valor de cada una de las alternativas en el procedimiento de elección subóptima está determinado conjuntamente por la información temporal que cada uno de los estímulos proporciona sobre la entrega de alimentos, por la tasa de reforzamiento primario asociado con cada alternativa y por un mecanismo de ponderación que atribuye un peso relativo a las dos primeras variables. El modelo explica la diferencia entre especies asumiendo que las palomas y las ratas asignan un peso diferente a esas variables; en concreto, se propuso que las palomas le dan más peso a la información temporal, mientras que las ratas le dan más peso a la tasa de reforzamiento primario. Por tal motivo los incrementos en la longitud de los enlaces terminales deberían aumentar el peso que los organismos atribuyen a la

información temporal y, como consecuencia, debería aumentar el comportamiento subóptimo.

Aunque las ratas y las palomas han mostrado preferencias opuestas bajo los mismos parámetros temporales en el procedimiento de elección subóptima, es posible que dada la mayor sensibilidad a la demora que se ha informado para las palomas que para las ratas en los procedimientos de descuento temporal (Green, Myerson, Holt, Slevin, & Estle, 2004), las ratas necesiten enlaces terminales más largos que las palomas para mostrar un comportamiento subóptimo. Para evaluar esta hipótesis, Cunningham y Shahan (2019) expusieron ratas a un procedimiento de elección subóptima en el que los enlaces terminales tenían diferentes longitudes en todas las condiciones, desde 10 s a 50 s. Los resultados indicaron que aunque las ratas eran óptimas en la condición en la que los enlaces terminales tenían una duración de 10 s, la mayoría de las ratas cambiaron a la preferencia por la alternativa subóptima cuando la longitud de los enlaces terminales se incrementó a 30 s o más (Cunningham & Shahan, 2019). Curiosamente, bajo conjuntos de suposiciones completamente independientes, el modelo de utilidad anticipatoria (Iigaya, Story, Kurth-Nelson, Dolan, & Dayan, 2016) y el modelo SiGN (McDevitt et al., 2016) también predicen que la elección subóptima aumentará a medida que los enlaces terminales se alarguen. Esta predicción se ha confirmado tanto con palomas (Dunn & Spetch, 1990; McDevitt, Pisklak, Spetch, & Dunn, 2018) como con humanos (Iigaya et al., 2016). Sin embargo, la evidencia presentada en el experimento de Trujano et al. (2015) que mostraron que las ratas seguían comportándose de manera óptima cuando alargaron la duración de los enlaces terminales a 30 s, sugieren que alargar los eslabones terminales no promueven un comportamiento subóptimo en ratas, por lo cual, es apropiado hacer un análisis de las diferencias en los procedimientos empleados por Cunningham & Shahan y Trujano et al.,

para determinar cuál fue la variable diferente en estos procedimientos que generó el comportamiento subóptimo reportado por Cunningham & Shahan. En particular, en el estudio de Cunningham & Shahan los autores presentaron un estímulo predictor de reforzamiento y un blackout como predictor de la ausencia de una recompensa en la alternativa Discriminativa, en contraste, Trujano et al., presentaron los dos estímulos en esta alternativa. La ausencia del estímulo predictor de la ausencia de una recompensa en el estudio de Cunningham & Shahan y dada la importancia que las ratas atribuyen a este estímulo (Trujano et al., 2016) sugiere que esta fue la variable responsable de promover el comportamiento subóptimo en su estudio, por lo tanto, el efecto de la duración de alargar los estímulos de los eslabones terminales parece depender de presentar solo un estímulo predictor de la recompensa en la alternativa Discriminativa para promover el comportamiento subóptimo en ratas.

Los datos de la presente tesis están relacionados con esta teoría. Específicamente, la conducta óptima reportada por las palomas en este experimento, sugieren que el cambio en el valor incentivo de los estímulos en palomas tiene un impacto en la valoración de la información temporal que predicen los estímulos. Los datos sugieren entonces que disminuir el valor incentivo de los estímulos en el procedimiento de elección subóptima en palomas, resulta en una disminución del valor información temporal que se les atribuye a estos estímulos. Este resultado sugiere que el valor incentivo de los estímulos modula el valor que le atribuyen las especies a la información temporal y por tal motivo en el Experimento 5 de la presente tesis el valor que le atribuyen las palomas a los estímulos de la alternativa Discriminativa respecto a la información temporal de la entrega de los reforzadores es menor que el valor que le asignan a la alternativa No Discriminativa, y

como originalmente esta teoría propuso, las palomas eligen las alternativas que les proporcionen mayor información temporal de la entrega de los reforzadores.

Por otro lado, el modelo del decaimiento de asociabilidad (Daniels & Sanabria, 2018) supone que cada estímulo del enlace terminal y su resultado influyen en el valor del enlace inicial de acuerdo con un parámetro de asociabilidad, que a su vez depende de la certeza de la relación entre el enlace terminal y presentación de resultados. Aunque en los procedimientos de elección subóptima, ambos estímulos de la alternativa Discriminativa predicen con certeza su resultado (reforzador y no reforzador, respectivamente) y, por lo tanto, la asociabilidad de ambos será cero en estado estacionario, la dinámica de esta pérdida, junto con el valor del estímulo positivo, hace que se le atribuya un valor más alto a la alternativa subóptima. Al ajustar el modelo a los datos de palomas y ratas, el modelo encontró una notable diferencia en el parámetro de asociabilidad para estas especies: mientras que para las palomas el estímulo que predijo la ausencia de una recompensa perdió la asociabilidad más rápido que el estímulo que predijo la entrega de una recompensa, para las ratas no hubo pérdida de asociabilidad para ninguno de estos estímulos. Por tal motivo, la diferencia entre especies en el procedimiento de elección subóptima es explicada por esta teoría en términos un impacto más fuerte del inhibidor condicionado en ratas que en palomas.

Los datos de la presente tesis apoyan la suposición de esta teoría y muestran que el valor incentivo de los estímulos mantiene el mismo valor de la asociabilidad de los estímulos con su resultado en las alternativas Discriminativa y No Discriminativa y por tal motivo el impacto del inhibidor condicionado se mantiene y las ratas prefieren la alternativa óptima. Es importante mencionar, que los datos del experimento 2 y 3 muestran que presentar un estímulo o un blackout no disminuye el valor de la asociabilidad del

resultado y por tal motivo, una palanca o un blackout funcionan de manera confiable como un inhibidor condicionado.

Los datos del experimento 5 sugieren que disminuir el valor incentivo de los estímulos en el procedimiento de elección subóptima en palomas, resulta en un incremento en la asociabilidad del resultado del estímulo que predice la ausencia de una recompensa y a diferencia del nulo impacto de una tecla que señala la ausencia de una recompensa, un tono que predijo la ausencia de una recompensa mantiene la asociabilidad y funciona como un inhibidor condicionado que promueve la elección de la alternativa óptima.

Conclusiones

En la presente tesis evaluamos la generalidad del comportamiento óptimo de las ratas y el comportamiento subóptimo de las palomas en el procedimiento de elección subóptima evaluando la hipótesis del valor incentivo de los estímulos. Si partimos de la premisa postulada por esta hipótesis, las teclas iluminadas para palomas y las palancas para ratas tienen valor incentivo y por tal motivo ambas especies se comportan de manera subóptima cuando son evaluadas con estos estímulos. Sin embargo, los resultados de la presente tesis sugieren que el valor incentivo de los estímulos no es la variable que explica las diferencias en el desempeño de las palomas y ratas en el procedimiento de "elección subóptima". La figura 35 presenta la comparación del desempeño de ratas y palomas en el procedimiento de elección subóptima cuando se utilizaron estímulos con alto y bajo valor incentivo.

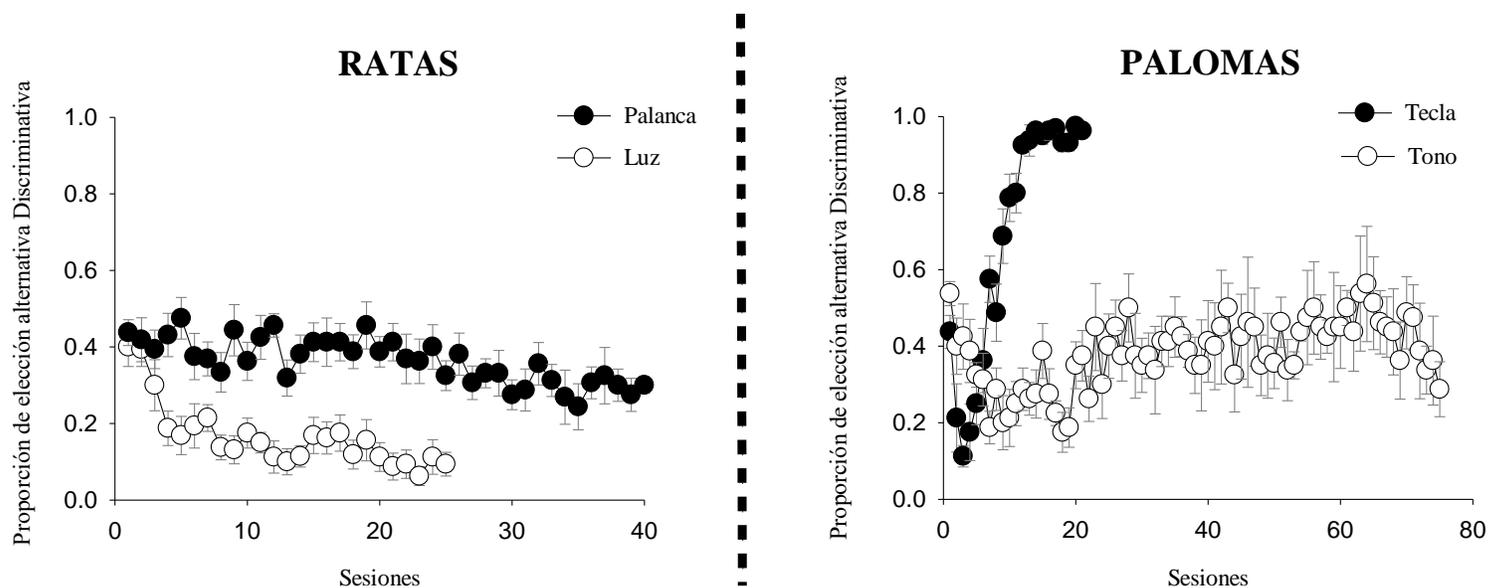


Figura 35 Proporción de elección de la alternativa Discriminativa de ratas y palomas con estímulos con alto y bajo valor incentivo. En el panel izquierdo se presentan los datos con ratas cuando se utilizaron palancas (alto valor incentivo) círculos negros y luces (bajo valor incentivo) círculos blancos. En el panel derecho se presentan los datos con palomas con estímulos con alto valor incentivo (tecla) y bajo valor incentivo (tono). Los datos del panel izquierdo corresponden a los datos del Experimento 2 de la presente tesis y los datos de Trujano y Orduña (2015). Los datos del panel derecho corresponden a datos inéditos del laboratorio y los datos del Experimento 5 de la presente tesis.

Como se puede observar en la Figura 35 las ratas desarrollaron una fuerte preferencia por la alternativa No Discriminativa (óptima) bajo condiciones donde se presentaron palancas (alto valor incentivo) y luces (bajo valor incentivo) en el procedimiento de elección subóptima. A pesar de que la proporción de elección es mayor cuando se utilizan palancas, al final del entrenamiento las ratas se comportaron de manera óptima al igual que cuando se utilizan luces. En contraste, las palomas desarrollaron una fuerte preferencia por la alternativa subóptima cuando se utilizan teclas iluminadas (alto valor incentivo) que cuando se utilizaron tonos (bajo valor incentivo).

Los resultados de la presente tesis muestran que a pesar de que las ratas atribuyen mayor valor incentivo a las palancas que a las luces, el nivel del valor incentivo es posiblemente más bajo que el nivel que las palomas le atribuyen a las teclas iluminadas, y posiblemente no supera el umbral para generar una elección subóptima. Basamos esta hipótesis en los resultados de la presente tesis, y en particular a los datos encontrados en los Experimentos 6 A y B que mostraron que solo bajo manipulaciones asimétricas del valor incentivo de los estímulos de las alternativas del procedimiento de elección subóptima las ratas aumentan su preferencia por la alternativa Discriminativa (subóptima), sin embargo, el nivel en la preferencia de esta alternativa no se compara con el nivel de preferencia mostrado por las palomas en este procedimiento. Es importante destacar el resultado encontrado con esta manipulación con ratas clasificadas como seguidores de señales quienes mostraron un mayor nivel de comportamiento subóptimo en las condiciones que se presentaban los estímulos con alto valor de incentivo en la alternativa Discriminativa (Ver Figura 36).

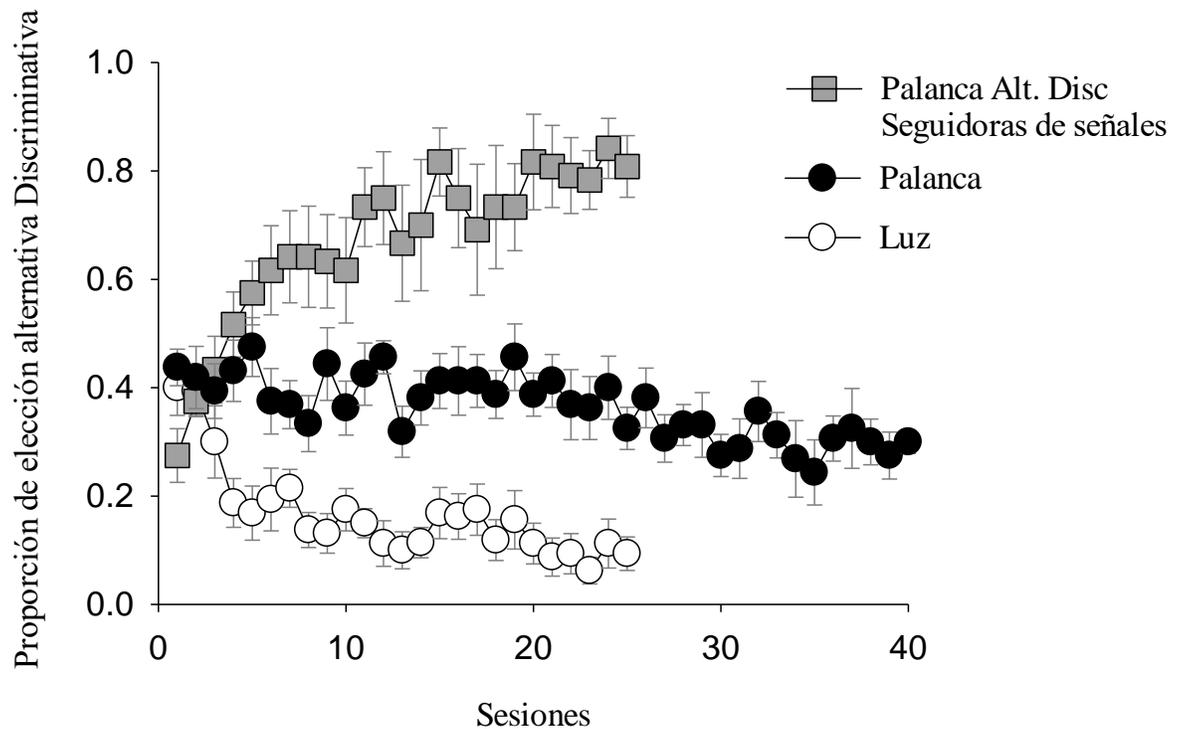


Figura 36 Proporción de elección de la alternativa Discriminativa en condiciones con estímulos con alto valor incentivo (Palanca), bajo valor incentivo (Luz) y Alto valor incentivo en la alternativa Discriminativa (con ratas clasificadas como seguidoras de señales).

Este resultado es crucial en el planteamiento de esta hipótesis debido a que este grupo de ratas muestran una mayor capacidad de atribuir valor incentivo a las palancas que sus grupos de comparación (ratas clasificadas como seguidoras de objetivos), y representan aproximadamente el 35% del total de la población (Flagel et al., 2011; Gillis & Morrison, 2019; Meyer et al., 2012; Roughley & Killcross, 2019). Dentro de este contexto, es razonable que solo este grupo de ratas desarrollara una fuerte preferencia por la alternativa Discriminativa debido a que el nivel del valor incentivo que le atribuyen a los estímulos podría compararse con el nivel del valor incentivo que las palomas les atribuyen a las teclas iluminadas y por tal motivo solo este grupo de ratas logró desarrollar un comportamiento subóptimo. Por lo tanto, futuras investigaciones de este tema tienen que buscar aquellos

estímulos que generen los mismos niveles de valor incentivo que generan las teclas en el caso de las palomas para poder buscar la generalidad entre especies en el procedimiento de elección subóptima.

Encontrar estos estímulos apoyaría una propuesta teórica reciente que explica las diferencias entre palomas y ratas en el procedimiento de elección subóptima. En particular esta propuesta se basa en la idea general de la teoría de los sistemas de comportamiento (Timberlake 1993), Zentall, Smith y Beckmann (2019) sugirieron que el picoteo a las teclas está relacionado con el comportamiento consumatorio en las palomas y por lo tanto estos estímulos activan un mecanismo de búsqueda focal de alimento, se atribuyen con valor incentivo y promueven el comportamiento subóptimo. En el caso de las ratas, presionar una palanca debería estar relacionado con el comportamiento consumatorio para atribuirle valor incentivo y generar comportamiento subóptimo, sin embargo, los resultados de la presente tesis muestran que las palancas en el caso de las ratas pueden estar generando dos escenarios posibles, en un primer escenario estos estímulos activa la búsqueda general (en lugar de la búsqueda focal), no se atribuyen valor incentivo y, por lo tanto, no promueven una elección subóptima o en un segundo escenario las palancas si activan una búsqueda focal de alimento pero no se le atribuye el mismo valor incentivo que generan las teclas en el caso de las palomas por lo cual las ratas no muestran un comportamiento subóptimo. Con respecto al hallazgo de la conducta optima en palomas cuando se emplean estímulos con bajo valor incentivo (Experimento 5) este resultado aporta evidencia a esta propuesta ya que los estímulos empleados activan la búsqueda focal de alimento (teclas) y búsqueda general (tonos).

Adicionalmente, nuestros resultados sugieren que las palomas no consideran a los estímulos predictores de no reforzamiento (inhibidores condicionados) en su valoración de la alternativa discriminativa cuando se presentan estímulos con alto valor incentivo y si ponderan este estímulo cuando tienen bajo valor incentivo. La evidencia de la presente tesis sugiere que la diferencia entre ratas y palomas en el procedimiento de elección subóptima está mediada por una respuesta diferencial a los estímulos predictores de no reforzamiento (inhibidores condicionados) que las ratas lo consideran con gran fuerza no importando el valor incentivo de los estímulos y las palomas parecen considerarlo solo cuando el estímulo predictor de no reforzamiento tiene bajo valor incentivo.

El procedimiento de elección subóptima ha traído en los últimos años un debate teórico que ha promovido una cantidad importante de publicaciones y posturas que pretenden explicar las diferencias entre especies encontradas con este procedimiento, sin embargo, actualmente el debate persiste y aunque parece resolverse con el papel que tienen los estímulos predictores de no reforzamiento en la elección de ratas y palomas, no debemos olvidar que en el mayoría de las investigaciones solemos partir de la premisa que todos los comportamientos deberían ser flexibles y perfectamente adaptables a las demandas ambientales que se le presentan a los organismos y es por ello que con frecuencia se asume que el comportamiento en sí mismo debe ser óptimo. Esta conclusión tiene que ser tomada con cautela y permitírnos mirar al comportamiento en términos de optimización bajo ciertas circunstancias que reflejen realmente las principales propiedades de los entornos que configuraron estas conductas. Por ejemplo, la premisa propuesta por Vasconcelos et al. (2015) que describe que en ambientes naturales en conducta de forrajeo las palomas deben prestar atención a los estímulos asociados con reforzadores e ignorar los

estímulos asociados con la ausencia de reforzadores es un ejemplo claro de estas circunstancias, ya que, dado que este mecanismo se desarrolló en el entorno natural, en el laboratorio las palomas también intentan escapar de ese estímulo y bajo estas circunstancias se reproducen las mismas condiciones que en el ambiente natural, en consecuencia, el comportamiento puede ser cercano al óptimo en el entorno en el cual una paloma probablemente encontrará en su ecología típica, pero subóptimo cuando el entorno actual no refleja las características principales del entorno en el que evolucionó. Por lo tanto, investigaciones futuras deben considerar esta premisa y buscar en el caso de las ratas las condiciones naturales en su ecología típica bajo las cuales se comporten de manera similar a lo encontrado con palomas.

En resumen, en la presente serie de experimentos encontramos que el valor incentivo de los estímulos no promueve el desarrollo de la conducta subóptima en ratas, debido a que los sujetos en la presente tesis eligieron las alternativas óptimas a pesar del uso de estímulos con alto valor incentivo. Por lo tanto, la evidencia actual sugiere que el valor incentivo de los estímulos no es necesario ni suficiente para desarrollar una elección subóptima en ratas. Caso contrario a las palomas en las cuales el valor incentivo de los estímulos modula su preferencia en este procedimiento, de tal manera que presentar estímulos con alto valor incentivo promueve el desarrollo de la conducta subóptima y presentar estímulos con bajo valor incentivo resulta en la elección de las alternativas óptimas. Finalmente, el reciente interés que se ha mostrado en esta línea de investigación en buscar modelos generales que integren las variables propuestas para explicar las diferencias entre especies en el procedimiento de elección subóptima, deben agregar un parámetro que permita cuantificar la contribución del valor incentivo de los estímulos para generar un

comportamiento subóptimo, ya que, los datos de la presente tesis sugieren que el valor incentivo de los estímulos podrían estar medidos a través de un continuo donde en un extremo estén aquellos estímulos que solo aportan información predictiva pero no generan ningún valor de incentivo y en el otro extremo estarían estos estímulos que aportan información predictiva y además tiene el valor de incentivo necesario para generar un comportamiento subóptimo.

REFERENCIAS

- Ahrens, A. M., Singer, B. F., Fitzpatrick, C. J., Morrow, J. D., & Robinson, T. E. (2016). Rats that sign-track are resistant to Pavlovian but not instrumental extinction. *Behavioural Brain Research*, 296, 418-430.
- Anselme, P. (2016). Motivational control of sign-tracking behaviour: A theoretical framework. *Neuroscience & Biobehavioral Reviews*, 65, 1-20.
- Autor, S. M. (1960). *The Strength of Conditioned Reinforcers as a Function of Frequency and Probability of Reinforcement: A Thesis Presented to the Department of Psychology in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Subject of Psychology*: Harvard University.
- Barrus, M. M., Cherkasova, M., & Winstanley, C. A. (2015). Skewed by cues? The motivational role of audiovisual stimuli in modelling substance use and gambling disorders. In *Behavioral Neuroscience of Motivation* (pp. 507-529): Springer.
- Barrus, M. M., & Winstanley, C. A. (2016). Dopamine D3 receptors modulate the ability of win-paired cues to increase risky choice in a rat gambling task. *Journal of Neuroscience*, 36(3), 785-794.
- Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation *Journal of the experimental analysis of behavior*, 12(6), 861-874.
- Beckmann, J. S., & Chow, J. J. (2015). Isolating the incentive salience of reward-associated stimuli: value, choice, and persistence. *Learning & memory*, 22(2), 116-127.
- Belke, T. W., & Spetch, M. L. (1994). Choice between reliable and unreliable reinforcement alternatives revisited: Preference for unreliable reinforcement. *Journal of the experimental analysis of behavior*, 62(3), 353-366.
- Breen, R. B., & Zuckerman, M. (1999). Chasing'in gambling behavior: Personality and cognitive determinants. *Personality and individual differences*, 27(6), 1097-1111.
- Brown, P. L., & Jenkins, H. M. (1968). AUTO-SHAPING OF THE PIGEON'S KEY-PECK 1. *Journal of the experimental analysis of behavior*, 11(1), 1-8.
- Case, J. P., & Zentall, T. R. (2018). Suboptimal choice in pigeons: Does the predictive value of the conditioned reinforcer alone determine choice? *Behavioural Processes*, 157, 320-326.
- Catania, A. C. (1963). Concurrent performances: a baseline for the study of reinforcement magnitude *Journal of the experimental analysis of behavior*, 6(2), 299-300.
- Chow, J. J., Smith, A. P., Wilson, A. G., Zentall, T. R., & Beckmann, J. S. (2017). Suboptimal choice in rats: Incentive salience attribution promotes maladaptive decision-making. *Behavioural Brain Research*, 320, 244-254.
- Chung, S. H., & Herrnstein, R. J. (1967). Choice and delay of reinforcement *Journal of the experimental analysis of behavior*, 10(1), 67-74.
- Cleland, G. G., & Davey, G. C. (1983). Autoshaping in the rat: The effects of localizable visual and auditory signals for food. *Journal of the experimental analysis of behavior*, 40(1), 47-56.
- Cunningham, P. J., & Shahan, T. A. (2018). Suboptimal choice, reward-predictive signals, and temporal information. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44(1), 1.
- Cunningham, P. J., & Shahan, T. A. (2019). Rats engage in suboptimal choice when the delay to food is sufficiently long. *Journal of Experimental Psychology: Animal Learning and Cognition*, 45(3), 301.
- Dagher, A., & Robbins, T. W. (2009). Personality, addiction, dopamine: insights from Parkinson's disease. *Neuron*, 61(4), 502-510.

- Daniels, C. W., & Sanabria, F. (2018). An associability decay model of paradoxical choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(3), 258.
- Dixon, M. J., Harrigan, K. A., Santesso, D. L., Graydon, C., Fugelsang, J. A., & Collins, K. (2014). The impact of sound in modern multiline video slot machine play. *Journal of gambling studies*, *30*(4), 913-929.
- Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the experimental analysis of behavior*, *53*(2), 201-218.
- Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the experimental analysis of behavior*, *12*(5), 723-730.
- Fantino, E., & Abarca, N. (1985). Choice, optimal foraging, and the delay-reduction hypothesis. *Behavioral and Brain Sciences*, *8*(2), 315-330.
- Fantino, E., Dunn, R., & Meck, W. (1979). Percentage reinforcement and choice. *Journal of the experimental analysis of behavior*, *32*(3), 335-340.
- Finlay, K., Kanetkar, V., Londerville, J., & Marmurek, H. H. (2006). The physical and psychological measurement of gambling environments. *Environment and Behavior*, *38*(4), 570-581.
- Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., . . . Akil, H. (2011). A selective role for dopamine in stimulus–reward learning. *Nature*, *469*(7328), 53.
- Flagel, S. B., Robinson, T. E., Clark, J. J., Clinton, S. M., Watson, S. J., Seeman, P., . . . Akil, H. (2010). An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: implications for addiction. *Neuropsychopharmacology*, *35*(2), 388.
- Fortes, I., Machado, A., & Vasconcelos, M. (2017). Do pigeons (*Columba livia*) use information about the absence of food appropriately? A further look into suboptimal choice. *Journal of Comparative Psychology*, *131*(4), 277.
- Fortes, I., Vasconcelos, M., & Machado, A. (2016). Testing the boundaries of “paradoxical” predictions: Pigeons do disregard bad news. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*(4), 336.
- Gillis, Z. S., & Morrison, S. E. (2019). Sign tracking and goal tracking are characterized by distinct patterns of nucleus accumbens activity. *eNeuro*, *6*(2).
- Gipson, C. D., Alessandri, J. J., Miller, H. C., & Zentall, T. R. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning & behavior*, *37*(4), 289-298.
- Goudriaan, A. E., Oosterlaan, J., de Beurs, E., & Van den Brink, W. (2004). Pathological gambling: a comprehensive review of biobehavioral findings. *Neuroscience & Biobehavioral Reviews*, *28*(2), 123-141.
- Green, L., Myerson, J., Holt, D. D., Slevin, J. R., & Estle, S. J. (2004). Discounting of delayed food rewards in pigeons and rats: is there a magnitude effect? *Journal of the experimental analysis of behavior*, *81*(1), 39-50.
- Gupta, R., & Derevensky, J. (1997). Familial and social influences on juvenile gambling behavior. *Journal of gambling studies*, *13*(3), 179-192.
- Hearst, E., Besley, S., & Farthing, G. W. (1970). Inhibition and the stimulus control of operant behavior. *Journal of the experimental analysis of behavior*, *14*(3S2), 373-409.
- Heffner, H. E., & Heffner, R. S. (1985). Hearing in two cricetid rodents: Wood rat (*Neotoma floridana*) and grasshopper mouse (*Onychomys leucogaster*). *Journal of Comparative Psychology*, *99*(3), 275.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement 1, 2. *Journal of the experimental analysis of behavior*, *4*(3), 267-272.
- Herrnstein, R. J. (1964). Secondary reinforcement and rate of primary reinforcement 1. *Journal of the experimental analysis of behavior*, *7*(1), 27-36.

- Hønsi, A., Mentzoni, R. A., Molde, H., & Pallesen, S. (2013). Attentional bias in problem gambling: a systematic review. *Journal of gambling studies*, 29(3), 359-375.
- Iigaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J., & Dayan, P. (2016). The modulation of savouring by prediction error and its effects on choice. *Elife*, 5, e13747.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 363-391.
- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the experimental analysis of behavior*, 21(3), 463-473.
- Kendall, S. B. (1985). A further study of choice and percentage reinforcement. *Behavioural Processes*, 10(4), 399-413.
- Laude, J. R., Beckmann, J. S., Daniels, C. W., & Zentall, T. R. (2014). Impulsivity affects suboptimal gambling-like choice by pigeons. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(1), 2.
- Laude, J. R., Stagner, J. P., & Zentall, T. R. (2014). Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(1), 12.
- López, P., Alba, R., & Orduña, V. (2018). Individual differences in incentive salience attribution are not related to suboptimal choice in rats. *Behavioural Brain Research*, 341, 71-78.
- Lovic, V., Saunders, B. T., Yager, L. M., & Robinson, T. E. (2011). Rats prone to attribute incentive salience to reward cues are also prone to impulsive action. *Behavioural Brain Research*, 223(2), 255-261.
- Martínez, M., Alba, R., Rodríguez, W., & Orduña, V. (2017). Incentive salience attribution is not the sole determinant of suboptimal choice in rats: Conditioned inhibition matters. *Behavioural Processes*, 142, 99-105.
- Mazur, J. E., & Fantino, E. (2014). Choice. *The Wiley Blackwell handbook of operant and classical conditioning*, 195-220.
- McDevitt, M. A., Diller, J. W., & Pietrzykowski, M. O. (2019). Human and pigeon suboptimal choice. *Learning & behavior*, 1-10.
- McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices. *Journal of the experimental analysis of behavior*, 105(1), 23-40.
- McDevitt, M. A., Pisklak, J. M., Spetch, M., & Dunn, R. (2018). The influence of outcome delay on suboptimal choice. *Behavioural processes*, 157, 279-285.
- McDevitt, M. A., Spetch, M. L., & Dunn, R. (1997). Contiguity and conditioned reinforcement in probabilistic choice. *Journal of the experimental analysis of behavior*, 68(3), 317-327.
- Meyer, P. J., Lovic, V., Saunders, B. T., Yager, L. M., Flagel, S. B., Morrow, J. D., & Robinson, T. E. (2012). Quantifying individual variation in the propensity to attribute incentive salience to reward cues. *PloS one*, 7(6), e38987.
- Meyer, P. J., Tripi, J. A., Tomie, A., & Morrow, J. (2018). Sign-tracking, response inhibition, and drug-induced vocalizations. *Sign-tracking and drug addiction*.
- Molet, M., Miller, H. C., Laude, J. R., Kirk, C., Manning, B., & Zentall, T. R. (2012). Decision making by humans in a behavioral task: Do humans, like pigeons, show suboptimal choice? *Learning & behavior*, 40(4), 439-447.
- Ojeda, A., Murphy, R. A., & Kacelnik, A. (2018). Paradoxical choice in rats: Subjective valuation and mechanism of choice. *Behavioural Processes*, 152, 73-80.
- Pitchers, K. K., Flagel, S. B., O'Donnell, E. G., Woods, L. C. S., Sarter, M., & Robinson, T. E. (2015). Individual variation in the propensity to attribute incentive salience to a food cue: influence of sex. *Behavioural Brain Research*, 278, 462-469.

- Potenza, M. N. (2008). The neurobiology of pathological gambling and drug addiction: an overview and new findings. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1507), 3181-3189.
- Raylu, N., & Oei, T. P. (2002). Pathological gambling: A comprehensive review. *Clinical psychology review*, 22(7), 1009-1061.
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, 72(2), 77.
- Rivalan, M., Ahmed, S. H., & Dellu-Hagedorn, F. (2009). Risk-prone individuals prefer the wrong options on a rat version of the Iowa Gambling Task. *Biological psychiatry*, 66(8), 743-749.
- Roberts, W. A. (1972). Short-term memory in the pigeon: Effects of repetition and spacing. *Journal of Experimental Psychology*, 94(1), 74.
- Robinson, T. E., & Berridge, K. C. (2008). The incentive sensitization theory of addiction: some current issues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1507), 3137-3146.
- Robinson, T. E., Yager, L. M., Cogan, E. S., & Saunders, B. T. (2014). On the motivational properties of reward cues: individual differences. *Neuropharmacology*, 76, 450-459.
- Roper, K. L., & Zentall, T. R. (1999). Observing behavior in pigeons: The effect of reinforcement probability and response cost using a symmetrical choice procedure. *Learning and Motivation*, 30(3), 201-220.
- Roughley, S., & Killcross, S. (2019). Differential involvement of dopamine receptor subtypes in the acquisition of Pavlovian sign-tracking and goal-tracking responses. *Psychopharmacology*, 1-10.
- Skinner, B. (1938). The behavior of organisms: an experimental analysis. Appleton-Century. In: Oxford, England.
- Smith, A. P., Bailey, A. R., Chow, J. J., Beckmann, J. S., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Stimulus value predicts choice over frequencies. *PloS one*, 11(7), e0159336.
- Smith, A. P., Beckmann, J. S., & Zentall, T. R. (2017). Gambling-like behavior in pigeons: 'jackpot' signals promote maladaptive risky choice. *Scientific reports*, 7(1), 6625.
- Smith, A. P., Hofford, R. S., Zentall, T. R., & Beckmann, J. S. (2018). The role of 'jackpot' stimuli in maladaptive decision-making: Dissociable effects of D1/D2 receptor agonists and antagonists. *Psychopharmacology*, 235(5), 1427-1437.
- Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42(2), 212.
- Stagner, J. P., Laude, J. R., & Zentall, T. R. (2011). Sub-optimal choice in pigeons does not depend on avoidance of the stimulus associated with the absence of reinforcement. *Learning and Motivation*, 42(4), 282-287.
- Stagner, J. P., Laude, J. R., & Zentall, T. R. (2012). Pigeons prefer discriminative stimuli independently of the overall probability of reinforcement and of the number of presentations of the conditioned reinforcer. *Journal of Experimental Psychology: Animal Behavior Processes*, 38(4), 446.
- Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review*, 17(3), 412-416.
- Stephens, D. W., Brown, J. S., & Ydenberg, R. C. (2007). *Foraging: behavior and ecology*: University of Chicago Press.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*: Princeton University Press.
- Thorndike, E. (1911). *Animal Intelligence: Experimental Studies* Macmillan. New York.

- Trujano, R. E., López, P., Rojas-Leguizamón, M., & Orduña, V. (2016). Optimal behavior by rats in a choice task is associated to a persistent conditioned inhibition effect. *Behavioural Processes*, *130*, 65-70.
- Trujano, R. E., & Orduña, V. (2015). Rats are optimal in a choice task in which pigeons are not. *Behavioural Processes*, *119*, 22-27.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.
- Vasconcelos, M., Machado, A., & Pandeirada, J. N. (2018). Ultimate explanations and suboptimal choice. *Behavioural Processes*, *152*, 63-72.
- Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information. *Scientific reports*, *5*, 13874.
- Wagenaar, W. A. (2016). *Paradoxes of gambling behaviour*: Routledge.
- Yager, L. M., & Robinson, T. E. (2015). Individual variation in the motivational properties of a nicotine cue: sign-trackers vs. goal-trackers. *Psychopharmacology*, *232*(17), 3149-3160.
- Zentall, Smith, A. P., & Beckmann, J. S. (2019). Gambling behavior: An animal model. *Translational Issues in Psychological Science*, *5*(3), 276.
- Zentall, T. R. (2014). Suboptimal choice by pigeons: An analog of human gambling behavior. *Behavioural Processes*, *103*, 156-164.
- Zentall, T. R. (2016). Resolving the paradox of suboptimal choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*(1), 1.
- Zentall, T. R., Andrews, D. M., & Case, J. P. (2017). Prior commitment: Its effect on suboptimal choice in a gambling-like task. *Behavioural Processes*, *145*, 1-9.
- Zentall, T. R., & Stagner, J. (2011). Maladaptive choice behaviour by pigeons: an animal analogue and possible mechanism for gambling (sub-optimal human decision-making behaviour). *Proceedings of the Royal Society B: Biological Sciences*, *278*(1709), 1203-1208.