



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE PSICOLOGÍA

**ELECCIÓN SUBÓPTIMA: BLOQUEO DEL VALOR
INCENTIVO PROMUEVE MEJOR TOMA DE DECISIONES**

TESIS

**QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN PSICOLOGÍA**

PRESENTA:

Julio Flores Padrón

Ciudad Universitaria, CDMX, 2020.



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.

El presente trabajo se realizó con el apoyo de los proyectos CONACYT (281548) y PAPIIT-DGAPA (IN306818).

A mi padre, gracias por todo el esfuerzo que has hecho para criarme de la mejor manera posible, espero un día ser tan buena persona como tú lo eres.

A mi madre, no tengo palabras suficientes para expresar mi agradecimiento por todo lo que me has brindado, soy afortunado por recibir siempre tu cariño y apoyo, sin ti no hubiera llegado hasta aquí, eres mi inspiración.

Les dedico este trabajo, les amo demasiado.

Agradecimientos

Agradezco a mi familia por el apoyo incondicional que me han proporcionado a lo largo de mi carrera universitaria. A mi hermana, **Sara Flores**, por tolerar los momentos en los que he estado bajo estrés, y a pesar de ello, tratar de animarme, te amo. A mi abuela **Susana**, gracias por tu cálida hospitalidad cuando lo necesité y sobre todo gracias por enseñarme el significado de la perseverancia aun cuando todo parece ser tan complicado.

A **Karen Yareni**, te agradezco por todas las pláticas, el tiempo y los consejos que me has dado, sin tu apoyo hubiera sido demasiado difícil concluir este proceso tan largo y en ocasiones demandante.

A todo el laboratorio 101, por hacerme sentir tan cómodo estos años. Agradezco al **Dr. Vladimir Orduña**, por permitirme colaborar con él, por su asesoría y por siempre procurar que desarrollemos las mejores habilidades en investigación. A mis compañeros, **Axel, Daniel, Alfredo, William, Enrique, Gaby, Fernanda, Liliana**, gracias por su compañía y hacer divertidos todos esos días que coincidimos en el laboratorio. A **Rodrigo Alba** y **Rodrigo González**, mi equipo de la mañana, por todas esas charlas, colaboraciones o reuniones, en ocasiones aprendimos juntos y en muchas otras me toco aprender de ustedes.

A mis amigos, los que me han acompañado a lo largo de mi trayectoria escolar, ustedes saben quiénes son, gracias por estar atentos en mis avances, por tantas aventuras compartidas y por ser una fuente de motivación, no tengo duda alguna de que sin ustedes este trabajo lo hubiera terminado hace mucho tiempo.

Agradezco a mi revisor, el **Rodolfo Bernal**, por siempre mostrarse atento y dispuesto a ayudarme durante el proceso de titulación. A mis sinodales, los doctores **Oscar Zamora, Gustavo**

Bachá y Ángel Tovar por sus revisiones y comentarios que fueron de gran ayuda para mejorar el presente trabajo.

Finalmente agradezco a la **UNAM**, por hacerme sentir como en casa desde Iniciación Universitaria, espero poder retribuir todo lo que me ha aportado a lo largo de los años.

ÍNDICE

RESUMEN	1
INTRODUCCIÓN	2
Programas Concurrentes	3
Elección Subóptima.....	5
Explicaciones en elección subóptima.....	10
Saliencia Incentiva	14
El sistema dopaminérgico y valor incentivo	15
Comparación entre especies	17
MATERIALES Y MÉTODO	20
Sujetos	20
Aparatos	20
Procedimiento.....	21
Habitación a la inyección	22
Habitación.....	22
Automoldeamiento	22
Pre-entrenamiento.....	23
Entrenamiento. Procedimiento de elección subóptima.....	24
Entrenamiento en reversión. Procedimiento de elección subóptima	25
RESULTADOS.....	26
Habitación	26
Automoldeamiento	26
Pre-entrenamiento	27
Proporción de elección	28
Respuestas a la tecla.....	32
DISCUSIÓN	36
REFERENCIAS.....	44

RESUMEN

Una amplia gama de estudios sobre el procedimiento de elección subóptima ha mostrado que las palomas presentan una fuerte preferencia por una alternativa con una tasa baja de reforzamiento, pero con estímulos que proporcionan información acerca de sus consecuencias, sobre otra alternativa con una tasa mayor de reforzamiento pero que carece de dichos estímulos informativos. Se han postulado diversas hipótesis, pero la explicación de este fenómeno todavía no es definitiva. Recientemente, se ha propuesto que los resultados obtenidos pueden deberse a la utilización de estímulos a los que los organismos atribuyen alta saliencia incentiva, teclas iluminadas en el procedimiento con palomas y palancas en el caso de ratas, y dichos estímulos podrían influir en la preferencia de elección. Bajo esta lógica, la hipótesis del presente estudio es que palomas disminuirán su elección subóptima al reducir la saliencia incentiva de los estímulos empleados. Para ello, se administró a dos grupos de palomas el fármaco haloperidol, un antagonista dopaminérgico, con la finalidad de disminuir el valor incentivo de las teclas iluminadas. Los resultados muestran una tendencia en la preferencia de elección por la alternativa óptima en la mayoría de las palomas pertenecientes a los grupos experimentales en comparación con el grupo control. Estos hallazgos dan sustento a la explicación de que el valor incentivo de los estímulos juega un papel relevante en el procedimiento de elección subóptima y es importante realizar investigación futura en este campo para evaluar la generalidad entre especies. Debido a que el procedimiento de elección subóptima ha sido propuesto como un modelo de juego patológico deben tomarse en cuenta las implicaciones que podría tener la saliencia incentiva para generar tratamientos eficaces hacia conductas maladaptativas en humanos.

INTRODUCCIÓN

Todos los días tomamos decisiones y las consecuencias de ellas van a dar pauta al curso de nuestra vida. Tomar malas decisiones implica invertir algún tipo de recurso y obtener como resultado una pérdida que varía en importancia dependiendo las características de la elección. Cuando hablamos de toma de decisiones debemos tener presentes ciertos aspectos los cuales valoramos para seleccionar la opción que es más conveniente sobre las demás. Por ejemplo, cuando salimos en fin de semana con amigos, solemos evaluar entre diversos lugares los beneficios que ofrecen para decidir el lugar más conveniente y que esté dentro de nuestro presupuesto. Cuando vamos retrasados a nuestro destino intentamos elegir la ruta que dada nuestra experiencia nos ha resultado más eficiente para llegar sin percances. En este sentido, aprendemos de las decisiones que tomamos y los resultados obtenidos para volver a reproducirlos en el futuro.

Estos simples ejemplos permiten visualizar mejor la idea de que en la vida cotidiana debemos llevar a cabo decisiones en todo momento. Sin embargo, en muchas ocasiones las decisiones pueden ser más difíciles de dilucidar y esto va a depender de diversos factores, especialmente cuando el número de alternativas y su modalidad aumenta. Una elección óptima o eficiente es llevada a cabo al elegir aquellas opciones que nos proporcionan una consecuencia beneficiosa o menos aversiva y para ello es necesaria una evaluación de las opciones disponibles.

Pierce y Cheney (2004), proporcionan la siguiente definición de elección:

El análisis de elección se refiere a la distribución del comportamiento operante entre fuentes alternativas de refuerzo. Cuando hay varias opciones disponibles, se puede elegir

una alternativa con más frecuencia que otras. Cuando esto ocurre, se llama preferencia por una fuente alternativa de refuerzo. (pág. 235)

Imaginemos la siguiente escena, existen dos restaurantes del mismo nombre, ambos a la misma distancia, pero en ubicaciones opuestas uno del otro, en ambos sirven un platillo reconocido por su sabor, incluso emplean la misma preparación, parece ser que ambas escenas son idénticas, sin embargo, difieren en un aspecto, el tamaño del plato suele ser más grande en uno de ellos, bajo estas condiciones, ¿en cuál de los dos lugares se debería ir a comer teniendo dicha información? A pesar de sus similitudes, la mejor decisión sería elegir el restaurante donde se obtendría una mayor ración ya que es el único aspecto que nos proporciona información acerca de una ganancia entre ambas opciones. Este simple ejemplo ayuda a comprender como los individuos pueden elegir alternativas que maximizan las ganancias y minimizan esfuerzo, supuestos explicados, por ejemplo, por teorías como la de forrajeo óptimo (Fantino & Abarca, 1985).

Esta visualización de elección suele ser más compleja en entornos reales en los que hay disponibles varias alternativas de respuesta y refuerzo, y el análisis de este proceso es importante y de gran interés para entender y predecir la conducta de elección. Su estudio se ha llevado a cabo en humanos y animales no humanos, en diversos contextos y manipulando diversas variables en busca de esclarecer los factores que tienen un papel en esta conducta.

Programas Concurrentes

Para llevar a cabo de manera sistemática este estudio dentro del laboratorio se suelen llevar a cabo procedimientos en cámaras operantes bajo programas concurrentes de reforzamiento, los

cuales consisten en un procedimiento de dos o más programas de reforzamiento, cada uno asociado con una diferente opción de respuesta, que puede ser una tecla, palanca, o *nosepoke*, por mencionar algunas, ambas disponibles al mismo tiempo y por medio de las cuales los organismos que se están estudiando pueden llevar a cabo de manera libre su elección (Shettleworth, 2010).

Una forma para medir dicha elección entre dos alternativas es calcular la tasa relativa de respuestas dadas en, digamos, la tecla izquierda por ejemplo, y para ello se divide la tasa de respuestas dadas en la tecla izquierda entre la tasa total de respuestas que resulta de la suma de las tasas de la tecla izquierda y derecha (Domjan, 2010). De forma matemática se vería de la siguiente forma:

$$\frac{CI}{CI + CD}$$

Donde CI representaría las respuestas dadas en la tecla izquierda y CD las respuestas dadas en la tecla derecha. Suponiendo que exista un número de respuestas igual entre ambas alternativas, entonces tendríamos una tasa total de 0.5. Si por otra parte existen más respuestas en la tecla izquierda, entonces la tasa total será más cercana a 1, y si por el contrario la tasa de respuestas es mayor en la tecla derecha, entonces la tasa total será más cercana a 0.

Este tipo de programas permiten llevar a cabo una medición continua de la elección proporcionada por los sujetos que puede cambiar a lo largo de las sesiones en las que se lleva a cabo el experimento, y donde el investigador en cuestión necesita tener un registro de que la preferencia mostrada sea constante y, por lo tanto, válida para el estudio.

Elección Subóptima

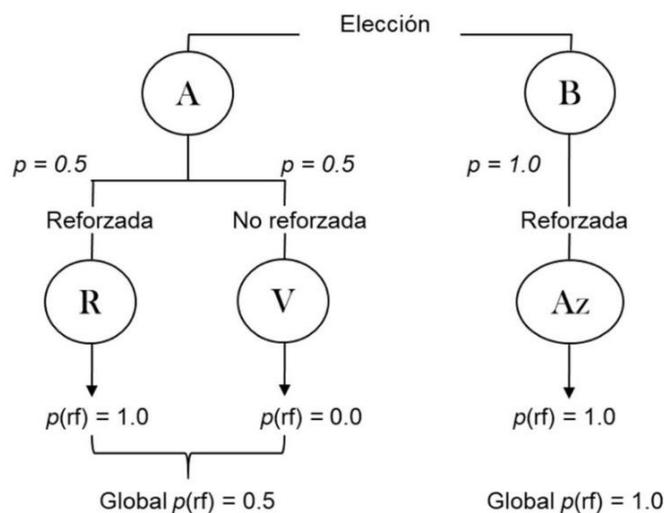
En la conducta de elección, los resultados parecen sugerir que los organismos consideran la evaluación de diferentes variables como la frecuencia (Herrnstein, 1961), magnitud (Catania, 1963) inmediatez (Chung & Herrnstein, 1967) y calidad (Hollard & Davison, 1971) de los reforzadores.

Cuando se llevan a cabo decisiones maladaptativas y que parecen estar fuera de este orden, es importante comprender por qué el organismo las lleva a cabo. He aquí lo interesante del tema, indagar sobre cómo los organismos toman decisiones, ha causado que diferentes disciplinas intenten aportar la mejor explicación. La investigación dentro del campo puede ser fundamental para comprender las causas y procesos que permiten que la mala toma de decisiones pueda llevarse a cabo. Esta comprensión, entre otras cosas tiene fines prácticos, por ejemplo, para después estructurar tratamientos e intervenciones eficientes ante un número grande de patologías, tales como impulsividad, abuso de sustancias, ludopatía, etc.

En la última década se ha estudiado un procedimiento conocido como elección subóptima, un ejemplo claro de una tarea en la que se ha reportado un gran número de veces mala toma de decisiones por parte de palomas como sujetos experimentales (Stagner, Laude, & Zentall, 2011; Zentall, 2016; Zentall & Laude, 2013). Dicho procedimiento contradice teorías como la de forrajeo óptimo, mencionado anteriormente, en el que se contempla como premisa principal que los individuos maximizan el reforzamiento gastando la menor cantidad de recursos (Fantino & Abarca, 1985). El procedimiento de elección subóptima se ha estudiado en diversas tareas y especies (Molet et al., 2012; Trujano & Orduna, 2015), además, este paradigma fue propuesto

como un modelo de juego patológico en humanos (Stagner & Zentall, 2010), lo cual ha llevado a la realización continua de una serie de estudios para tratar de definir si dicha tarea y sus hallazgos en palomas son comparables a la conducta de ludopatía existente en humanos.

El estudio actual de elección subóptima puede partir de diferentes estudios pioneros, uno de los principales en causar gran interés fue un experimento realizado por Kendall (1974), en el que se expuso a palomas a una condición de elección (ver Fig. 1) entre una alternativa A, en la que la mitad de las veces podía desencadenar un estímulo, por ejemplo una tecla roja, que siempre era seguida de alimento y la otra mitad de las veces otro estímulo, digamos una tecla verde, era seguida por ausencia de alimento (tasa global de reforzamiento de 0.5), u otra alternativa B, que desencadenaba un único estímulo, una tecla azul por ejemplo, que siempre era seguida de alimento (tasa global de reforzamiento de 1.0). Contrario a lo esperado, los resultados reportados muestran una fuerte preferencia por la alternativa A que produce menor alimento. Aunque dicho estudio pudo mostrar hallazgos sorprendentes, diversos intentos de réplica no tuvieron éxito en investigaciones subsecuentes al no poder obtener datos contundentes (Belke & Spetch, 1994; Dunn & Spetch, 1990; Fantino, Dunn, & Meck, 1979).

Figura 1**Figura 1 – Procedimiento de Kendall (1974).**

Más adelante, Roper y Zentall (1999), realizaron un experimento con el objetivo de indagar más acerca del poder que tiene la información de los estímulos sobre la preferencia de los individuos; para ello realizaron un experimento en el que palomas podían elegir entre una alternativa que al ser escogida podía proveer dos teclas iluminadas de diferentes colores, una de ellas, digamos roja, siempre era seguida de alimento, mientras que la otra tecla, digamos verde, nunca era seguida de alimento (esta alternativa proporcionaba una tasa de reforzamiento de 0.5), y otra alternativa que al ser escogida resultaba en otros dos estímulos diferentes que no proporcionaban información sobre cuando se otorgaría alimento (esta alternativa también proporcionaba una tasa de reforzamiento de 0.5). A pesar de que ambas alternativas eran igual de ricas, las palomas mostraron una fuerte preferencia por la alternativa que proporcionaba estímulos informativos. Tomando este estudio como antecedente, Gipson, Alessandri, Miller, y Zentall (2009), proponen una modificación de ese procedimiento (ver Fig 2), deciden hacer más rica la alternativa no informativa (óptima), con una tasa de reforzamiento de 0.75, mientras que la alternativa informativa (subóptima) permanece sin cambios con una tasa de reforzamiento de 0.5,

encontrando que las palomas llevan a cabo una elección por la alternativa que predice una tasa baja de reforzamiento pero que proporciona estímulos informativos, sobre otra que es más beneficiosa, pero proporciona estímulos que no son informativos.

Figura 2

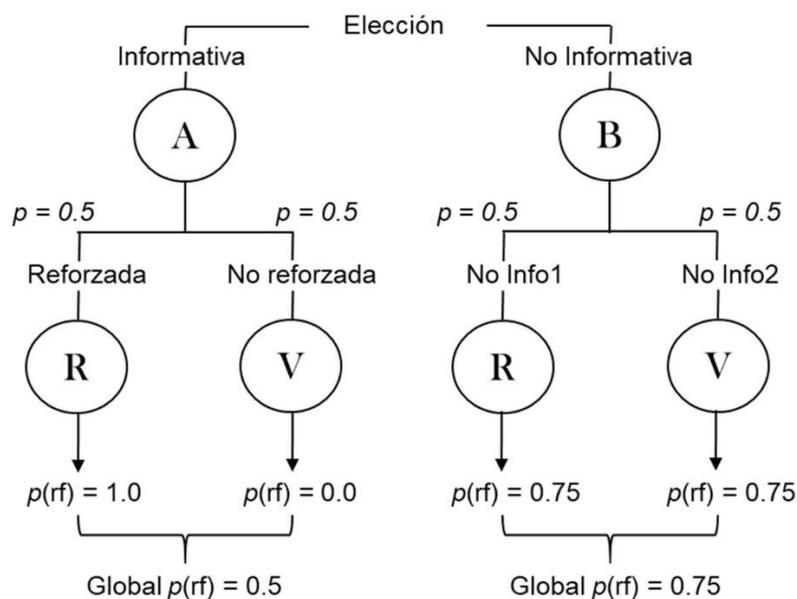


Figura 2 – Procedimiento de Gipson et al. (2009).

En un estudio posterior, Stagner y Zentall (2010), desarrollarían “la tarea clásica de elección subóptima” que se ha utilizado en mayor medida con palomas (ver Fig. 3). El procedimiento es similar al mencionado anteriormente, con la diferencia que las probabilidades de reforzamiento de ambas alternativas fueron disminuidas. Los resultados demuestran que dentro de la tarea existe una fuerte preferencia por una alternativa (alternativa subóptima) que al ser elegida puede desencadenar dos estímulos; uno de ellos el 80% de las veces señala la ausencia de reforzamiento, y el 20% de las veces puede aparecer otro estímulo diferente que siempre es seguido por comida, es decir, esta alternativa proporciona una tasa global de reforzamiento de 20%

(alternativa denominada como discriminativa debido a que los estímulos señalizan de manera precisa cuando habrá o no reforzamiento), sobre otra alternativa (alternativa óptima) que al ser elegida puede desencadenar otros dos estímulos diferentes que el 50% de las veces serán seguidos por reforzamiento, otorgando del mismo modo una tasa global de reforzamiento de 50%, pero donde dichos estímulos no proporcionan señales de cuándo será entregada la comida (razón por la que se denomina alternativa no discriminativa).

Figura 3

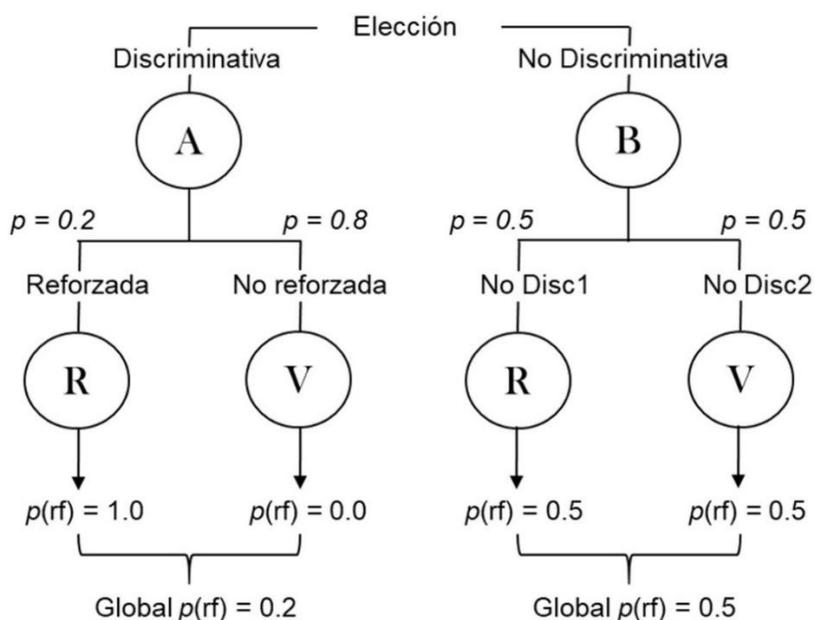


Figura 3 – Procedimiento de Stagner y Zentall (2010).

El procedimiento de elección subóptima ha sido postulado como un modelo de juego patológico en humanos sin poseer aún generalidad ni conocimiento exacto de las bases subyacentes por las que se efectúa dicha mala elección (Zentall, 2016; Zentall & Laude, 2013). Es importante recalcar que ha sido relacionada con diversos problemas sociales y poblaciones de riesgo relacionadas con impulsividad, mala toma de decisiones, juego patológico, abuso de sustancias,

etc., por lo que se requiere de investigación que explique este fenómeno y de contar con el sustento requerido se pueda validar como un modelo de toma de decisiones mal adaptativas en humanos, lo cual podría contribuir a buscar mejores intervenciones de bienestar social ante un número grande de patologías.

Explicaciones en elección subóptima

A pesar de que ha existido un amplio estudio de elección subóptima en palomas aún ha sido difícil postular una explicación definitiva que resuelva la incógnita sobre por qué las palomas realizan una mala toma de decisiones en esta tarea. Una de las primeras teorías propuestas en el campo para explicar la mala toma de decisiones sigue el curso de la idea planteada por Roper y Zentall (1999), se trata de la importancia dada a la información de los estímulos discriminativos, esta explicación aborda el hecho de que la alternativa discriminativa -compuesta por un estímulo predictor de reforzamiento, el estímulo excitatorio, y otro estímulo predictor de la ausencia de reforzamiento, el estímulo inhibitorio-, proporciona con certeza cuando será entregado el alimento, al contrario de la alternativa no discriminativa que aunque ofrece una tasa mayor de reforzamiento, está compuesta por estímulos no informativos respecto a la entrega de reforzamiento.

Sin embargo, esta explicación fue puesta a prueba en un estudio realizado por Zentall y Stagner (2011), y desechada al encontrar que cuando la alternativa no informativa es cambiada por una alternativa que siempre será seguida de alimento y aun con una tasa mayor de reforzamiento sigue siendo menos preferida por los sujetos. Para ello utilizaron una modificación de la tarea clásica, cambiando probabilidades de reforzamiento por magnitudes, es decir, esta vez en la alternativa discriminativa, el estímulo excitatorio que aparece el 20% de las veces será seguido por

10 pellets, mientras que el estímulo inhibitorio que aparece el 80% de las veces será seguido por 0 pellets, con un promedio general de 2 pellets recibidos en esta alternativa. En la otra alternativa, un único estímulo aparece el 100% de las veces y proporcionará 3 pellets en promedio. A pesar de que en esta tarea ambas alternativas son informativas, las palomas siguen prefiriendo fuertemente la alternativa subóptima que posee tanto el estímulo excitatorio como inhibitorio. Estos hallazgos parecen indicar que, aunque la información proporcionada por la alternativa discriminativa en la tarea clásica no es la causa por la que las palomas la prefieren, efectivamente los estímulos de dicha alternativa resultan relevantes para la especie. El rechazo de esta explicación inició una serie de estudios con palomas que han intentado explicar la conducta subóptima.

El interés por proporcionar una explicación se ha convertido en un importante reto al encontrar optimalidad con ratas, el modelo animal más utilizado para estudiar conducta en investigación básica (Martínez, Alba, Rodríguez, & Orduña, 2017; Orduna & Alba, 2019; Trujano, López, Rojas-Leguizamón, & Orduña, 2016; Trujano & Orduna, 2015). Una de las principales hipótesis planteadas para intentar explicar las diferencias existentes entre palomas y ratas dentro de la tarea, es la valoración que cada especie le da al estímulo inhibitorio, es decir, un estímulo que siempre predice la ausencia de reforzamiento y que, de acuerdo con esta explicación, podría ser poco relevante para palomas, pero muy valorado por ratas. Lo que intenta explicar esta postura es que al presentarse la alternativa discriminativa (subóptima), el estímulo inhibitorio aparecerá el 80% de las veces, por lo que se propone que palomas en presencia de dicho estímulo lo pueden llegar a ignorar e invertir tiempo en otras tareas como podrían ser asearse, voltearse, explorar la caja experimental, etc., y de este modo, solamente estarían contemplando en su elección al otro estímulo perteneciente a esa alternativa, el estímulo excitatorio, cuyo valor entonces sería de 100%

de las veces otorgando reforzamiento, sobre otra alternativa (no discriminativa) que el 50% de las veces ofrece reforzamiento. Para observar si esta explicación era cierta, en un experimento, Stagner et al. (2011), modificaron la presentación del estímulo inhibitorio que en el procedimiento clásico era una tecla iluminada de color, digamos verde, por una luz ambiental del mismo color. La lógica del experimento era proporcionar un estímulo difuso que el sujeto fuera incapaz de ignorar, de este modo, tendría que evitar dicha opción debido a sus contingencias y, por lo tanto, tener que ejecutar óptimamente dentro de la tarea, sin embargo, dicha modificación no pudo decrementar el nivel de elección subóptima tal como era esperado. Es posible entonces que existan diferencias en el efecto de inhibición condicionada causado por el estímulo que predice ausencia de reforzamiento, mientras que en palomas el efecto se disipa rápidamente a lo largo de la tarea (Laude, Stagner, & Zentall, 2014), en ratas parece no disminuir (Trujano et al., 2016).

Además, se ha intentado determinar que la impulsividad puede ser un factor que está teniendo un papel en la elección subóptima por parte de palomas. La impulsividad, es una característica de aquellos organismos que llevan a cabo conductas de riesgo y toma de decisiones inadecuadas; se puede definir como llevar a cabo acciones sin tener un adecuado autocontrol (Baumann & Odum, 2012). Dentro del contexto de impulsividad, se desglosa otro término, la elección impulsiva. De acuerdo con Broos et al. (2012), “Elección impulsiva es operacionalizada por decisiones impulsivas que resultan de una evaluación distorsionada de las consecuencias retrasadas de la conducta y una mayor preferencia por las recompensas inmediatas (más pequeñas) que por recompensas más tardías y más beneficiosas” (p. 1). La elección impulsiva es un elemento de un gran número de conductas patológicas, como juego patológico, desorden por déficit de atención e hiperactividad, vulnerabilidad hacia abuso de sustancias, y disfunción prefrontal

(Kalenscher et al., 2005). Laude, Beckmann, Daniels, y Zentall (2014), implementaron una tarea de descuento temporal (función que es considerada una medida efectiva de impulsividad) antes de llevar a cabo el procedimiento de elección subóptima, encontrando una correlación positiva entre el grado de elección subóptima y la inclinación de las funciones de descuento temporal, indicando que existe una tendencia en la preferencia de palomas a elegir de manera subóptima relacionado con el grado de impulsividad que manifiestan.

Por otra parte, recientemente ha surgido otra hipótesis que podría explicar las diferencias existentes en ratas y palomas dentro del procedimiento de elección subóptima definida como el valor incentivo atribuido a los estímulos. En los primeros estudios realizados en elección subóptima con ratas como sujeto experimental se emplearon luces como estímulos (Trujano et al., 2016; Trujano & Orduna, 2015), sin embargo, las luces o los tonos son estímulos atribuidos con baja saliencia incentiva para esta especie (Beckmann & Chow, 2015). En un estudio reciente, Chow, Smith, Wilson, Zentall, y Beckmann (2017), han propuesto que la conducta de elección por la alternativa de menor probabilidad de reforzamiento puede deberse a los estímulos utilizados y la atribución de saliencia incentiva que poseen. Al utilizar palancas que parecen tener mayor valor incentivo para ratas se encontró una tendencia a realizar una elección subóptima y, aunque ha sido difícil replicar dichos resultados (Alba, Rodríguez, Martínez, & Orduña, 2018; López, Alba, & Orduña, 2018; Martínez et al., 2017; Orduna & Alba, 2019), es el primer estudio en encontrar suboptimalidad en ratas bajo la hipótesis de saliencia incentiva y que podría explicar la elección subóptima en palomas al utilizar teclas iluminadas que parecen tener alto valor incentivo para esta especie, esto basado en que las palomas responden a ellas aunque sus respuestas sean independientes de la entrega de reforzador (Brown & Jenkins, 1968).

Saliencia Incentiva

Existen estímulos que poseen saliencia incentiva y que los convierten en fuertes estímulos condicionados para los organismos. Se han descrito operacionalmente las características que deben tener estos estímulos:

1. Los estímulos incentivos sesgan la atención hacia ellos y son tan atractivos que los sujetos se aproximan a ellos.
2. Los estímulos incentivos por sí mismos comienzan a ser objeto de deseo, en el sentido que los individuos ahora trabajarán para obtenerlos, y ellos pueden también reforzar el aprender una nueva respuesta instrumental para obtenerlos (es decir, fungen como reforzadores condicionados o secundarios).
3. Estímulos con saliencia incentiva pueden generar un estado motivacional condicionado que puede incitar o estimular una búsqueda renovada para la recompensa asociada.
(Meyer et al., 2012, p. 1)

Entonces, aquellos estímulos que cumplan con estos requisitos podrían provocar que los animales se acerquen a ellos (seguimiento de señales), un ejemplo de ello es visto en el procedimiento de automoldeamiento. En dicho procedimiento, el 100% de las palomas realizan respuestas de picoteo de tecla al presentar repetidamente una tecla iluminada seguida de la presentación de comida, pero no condicionada al picoteo (Brown & Jenkins, 1968) y esta conducta persiste incluso si los picotazos anulan la entrega de alimento (Williams & Williams, 1969); por otra parte, en ratas se observa la aparición de la respuesta al presentar una palanca seguida de

comida (Powell, Saunders, & Thompson, 1968), en aproximadamente el 35% de los sujetos (Meyer et al., 2012).

Además, es importante señalar que de acuerdo a la teoría de saliencia incentiva encontramos algunos factores de riesgo comórbidos de la importancia que representan estas señales para los individuos, como lo es la dependencia en adicciones que se desarrolla de manera más eficaz por señales asociadas con drogas que motivan y representan un factor para el abuso de sustancias. Dentro de esta postura, entra en juego un fenómeno que es llamado Seguimiento de Señales, que describe como de manera preferencial los individuos se aproximan a estímulos condicionados asociados con recompensas (con alta saliencia incentiva) exhibiendo un rango de cambios neurobiológicos y conductuales relacionados con adicciones (Langdon, Hathaway, Zorowitz, Harris, & Winstanley, 2019).

Tomando en cuenta esta idea, es posible que, si la elección por la alternativa discriminativa por parte de palomas es debida al alto valor incentivo atribuido a las teclas iluminadas, entonces debería esperarse conducta óptima si dicha saliencia incentiva no existiera en los estímulos utilizados en el procedimiento.

El sistema dopaminérgico y valor incentivo

Al hablar de saliencia incentiva atribuida a los estímulos hay evidencia que explica que el sistema dopaminérgico es uno de los sistemas de neurotransmisores involucrado en este proceso (Berridge, 2007; Berridge & Robinson, 1998; DiFeliceantonio & Berridge, 2016). Anteriormente se ha encontrado que el haloperidol, un antagonista dopaminérgico, disminuye expresiones

conductuales anticipatorias de pollos ante un estímulo condicionado (EC) que predice una recompensa (Acerbo, Godoy, & Delius, 2003), dichas expresiones anticipatorias son consideradas como un índice cuantificable del nivel de atribución de saliencia incentiva que dan los pollos a un estímulo predictor de reforzamiento. Estos resultados aportan evidencia de que áreas cerebrales, específicamente el Núcleo Accumbens y el Área Tegmental Ventral (relacionadas como el sistema de recompensa), con señalización dopaminérgica tienen una implicación en el valor de las recompensas y que éste puede ser bloqueado por antagonistas dopaminérgicos.

La recompensa es un elemento que los individuos toman como punto de partida para llevar a cabo decisiones del comportamiento que deben efectuar ante diversas situaciones. El procesamiento de recompensas conlleva una serie de modulaciones neuronales en las que se ven involucrados diversos neurotransmisores y que hacen posible que podamos identificar dichas recompensas como elementos placenteros o de desagrado.

Esta serie de conexiones neuronales es mejor conocida como Sistema de Recompensa. Reuter et al. (2005), puntualiza de manera más concreta cómo se conectan de manera específica las diferentes zonas cerebrales en el sistema de recompensa:

El sistema de recompensa se proyecta en todo el cerebro medio, las áreas de los ganglios basales, límbicos y frontales y puede subdividirse en numerosos aspectos anatómicos, bioquímicos y de comportamiento. Está involucrado desde instintos básicos como comer y la sexualidad hasta aspectos sociales altamente complejos.

Las proyecciones dopaminérgicas desde el Área Tegmental Ventral (ATV) al Estriado, la Corteza Cingulada Anterior (CCA), las estructuras Amígdala e Hipocampo y Prefrontal forman un circuito de procesamiento de recompensa. Dentro de este circuito, el Estriado Ventral y su Núcleo Accumbens desempeñan un papel central. Está bien conectado a las Cortezas Prefrontal y Frontal, incluida la Corteza Orbitofrontal (COF). Las estructuras límbicas como el CórTEX Cingulado Anterior, la Amígdala y el Hipocampo también participan en el procesamiento de recompensas y muestran una alta conectividad con otras áreas relacionadas con recompensas. Los problemas relacionados con las recompensas, como la toma de decisiones, el control de los impulsos, el comportamiento dirigido al objetivo y el aprendizaje, están integrados en el sistema de recompensas y no pueden verse aislados del procesamiento de recompensas. (pág. 5)

Actualmente, técnicas, como la inmunohistoquímica, nos permiten conocer las vías dopaminérgicas, es decir, dónde se localizan las neuronas de dopamina y dónde se proyectan. Las neuronas dopaminérgicas se originan en varios núcleos del mesencéfalo vecinos, siendo la sustancia nigra y el área tegmental ventral las que se proyectan al prosencéfalo (Puig, Rose, Schmidt, & Freund, 2014).

Comparación entre especies

Si bien es cierto que existen diferencias entre el cerebro de mamíferos y aves, separadas por una larga línea evolutiva, es imposible negar que hay similitudes en una gran diversidad de capacidades cognitivas, lo cual ha permitido que las palomas formen parte de las especies experimentales más empleadas dentro de las ciencias de comportamiento. Del mismo modo,

neuroanatómicamente se sabe que las aves poseen estructuras cerebrales que comparten similitudes con las de los mamíferos, tal es el caso del Nidopallio Caudolateral considerado equivalente a la corteza prefrontal de los mamíferos (Güntürkün, 2005). De acuerdo con Puig et al. (2014), el Nidopallio Caudolateral es el principal objetivo cortical de la inervación dopaminérgica. Como en los mamíferos, estas proyecciones surgen en el Área Tegmental Ventral y en la Substancia Nigra. El papel de dopamina en el cerebro de aves es bastante comparable al del cerebro mamífero, está involucrada en control motor y aprendizaje.

Los agonistas y antagonistas de dopamina se han utilizado para comprender mejor el papel del sistema dopaminérgico en cuestiones de aprendizaje. Los agonistas y antagonistas de dopamina tienen afinidad a dos superfamilias de receptores de dopamina (DA), los receptores D1 y D2. De manera general, Martelle y Nader (2008) exponen las localizaciones de estos receptores en el Sistema Nervioso Central:

Los receptores tipo DA D1 se ubican de manera postsináptica en mayor medida en el cuerpo estriado, una estructura del mesencéfalo que incluye el núcleo caudado, el putamen y el núcleo accumbens. Por otro lado, los receptores tipo D2 se localizan pre y postsinápticamente en las regiones cerebrales estriadas. Además, los sitios de unión al receptor de tipo D2 también están presentes en toda la corteza cerebral, aunque las densidades son menores que en el cuerpo estriado. (pág 248)

El presente estudio se centró en los efectos bloqueadores del haloperidol, un antagonista dopaminérgico selectivo del receptor de dopamina tipo D2, el cual actúa directamente uniéndose

a receptores D2 manteniendo al neurotransmisor endógeno alejado de los sitios de unión y, en consecuencia, bloqueando su acción (Acerbo, 2001). El haloperidol es uno de los fármacos más empleados para bloquear las propiedades motivacionales incentivas que poseen ciertos estímulos condicionados tanto en ratas (Danna & Elmer, 2010) como en aves (Acerbo et al., 2003; Danna & Elmer, 2010; Moe et al., 2011); también hay otros antagonistas que, aunque se han utilizado en menor medida y únicamente en roedores, aportan evidencia a favor del papel de dopamina en aprendizaje asociativo ante estímulos reforzantes, como es el caso de eticlopride, antagonista específico del receptor D2 (Grakalic, Panlilio, Thorndike, & Schindler, 2007; Martelle & Nader, 2008); SCH-23390, antagonista del receptor D1 (Chow, Nickell, Darna, & Beckmann, 2016; Zarrindast, Hajian-Heydari, & Hoseini-Nia, 1992); α -flupenthixol, antagonista de receptores D1/D2 (Di Ciano, Cardinal, Cowell, Little, & Everitt, 2001).

A partir de la literatura mencionada anteriormente, surge la idea de que al utilizar un antagonista dopaminérgico podría bloquearse de manera neurobiológica las propiedades incentivas que poseen los estímulos condicionados y de este modo, si dicha hipótesis es correcta, se encontraría dentro del procedimiento clásico de elección subóptima en palomas en el que se utilizan teclas iluminadas con alto valor incentivo para esta especie, una toma de decisión óptima por la alternativa que provee una mayor tasa de reforzamiento, aportando evidencia favorable a la explicación de saliencia incentiva.

MATERIALES Y MÉTODO

Sujetos

Con el fin de evaluar el papel del antagonista dopaminérgico, haloperidol, en el procedimiento de elección subóptima se utilizó un diseño entre grupos. Se utilizaron 18 palomas híbridas adultas, de sexo indistinto y mantenidas privadas al 80% de su peso en alimentación libre, con agua disponible ad libitum, al final de cada sesión experimental fueron alimentadas para mantener ese peso estable. Los sujetos habitaban de manera individual en jaulas dentro de un cuarto a 26 °C con ciclos de 12 horas de luz y oscuridad. Todos los protocolos experimentales siguieron la Norma Oficial Mexicana de Especificaciones Técnicas para la Producción, Cuidado y Uso de Animales de Laboratorio (NOM-062-ZOO-1999).

Se formaron tres grupos de palomas que fueron inyectadas de manera crónica:

- 1) Haloperidol Dosis Alta (n = 6), conformado por palomas que fueron inyectadas a lo largo de todo el experimento con haloperidol a una dosis de 0.2 mg/kg.
- 2) Haloperidol Dosis Baja (n = 6), conformado por palomas que fueron inyectadas bajo las mismas condiciones que el grupo de Dosis Alta con haloperidol a una dosis de 0.1 mg/kg.
- 3) Salina grupo Control (n = 6), conformado por palomas que fueron inyectadas con salina como vehículo durante el experimento.

Aparatos

El experimento fue llevado a cabo en 5 cajas de condicionamiento operante (MED Associates, Inc., Modelo ENV-007). Cada cámara operante medía 35 cm (largo) x 30 cm (ancho)

x 37 cm (alto) y encerrada en un cubículo atenuante de sonido con un ventilador para enmascarar sonidos externos. La cámara operante estaba equipada de 3 teclas de respuesta de 2.5 cm de diámetro, con 7.5 cm de distancia entre ellas, colocadas 23 cm por encima del suelo de la cámara. Las teclas podían iluminarse de 3 colores diferentes, blanco, rojo y verde. Un comedero (MED Associates, Inc., Modelo ENV-205) fue instalado 9.5 cm por debajo de la tecla central que se iluminaba al activarse por una lámpara de 28-V, 0.04-A. El reforzador otorgado consistía en una combinación de diferentes granos de comida para aves. Además, una webcam Logitech C170 fue añadida en uno de los muros de la caja para monitorear las sesiones. La presentación de estímulos y la recopilación de datos fueron controlados por computadoras personales que usan el lenguaje de programación Medstate (Med-PC-IV, MED Associates, Inc.).

Procedimiento

Se llevó a cabo una manipulación farmacológica de forma crónica, administrando el fármaco haloperidol en los grupos experimentales y salina en el grupo control durante todas las fases del procedimiento. La aplicación de los tratamientos fue llevada a cabo una vez por día. Después de haber sido inyectadas las palomas permanecieron en sus jaulas habitación durante 30 minutos, antes de ingresarlas a las cámaras operantes.

Dado el conocimiento de estudios previos utilizando haloperidol con aves (Moe et al., 2011), se buscó utilizar únicamente dosis bajas y moderadas para evitar efectos sedativos, es decir, dosis menores a 0.5 mg/kg, ya que a dosis elevadas a esta se han encontrado resultados atribuidos a sedación y no tanto al efecto que pudiera tener el fármaco sobre el valor incentivo de los estímulos.

Habitación a la inyección

Debido a que se llevó a cabo un programa de administración crónica de los tratamientos por inyección intramuscular (pectoralmente) cuando todos los sujetos alcanzaron el 80% de su peso libre y antes de que estuvieran en contacto con el fármaco, se les inyectó durante 3 días seguidos salina vía intrapectoral, en un volumen de 1 ml/kg. proporcional al volumen de fármaco que fue inyectado en el período experimental.

Habitación

Se llevaron a cabo diversas fases experimentales para desarrollar adecuadamente el procedimiento de elección subóptima. En esta primera fase las palomas recibieron de 1 a 3 sesiones de habitación a la cámara experimental, que consistía en un programa de tiempo fijo (TF) 25 s en el que una vez transcurrido este tiempo los animales tenían acceso al comedero durante 3 s, todos los demás estímulos siempre estuvieron inactivos. El programa finalizó después de 20 minutos o después de 20 entradas al comedero. Los sujetos podían avanzar a la siguiente fase del experimento al día siguiente una vez que registraran las 20 entradas al comedero, por otra parte, si un sujeto no realizaba ninguna respuesta al comedero a lo largo de 3 sesiones era excluido del experimento.

Automoldeamiento

Al día siguiente después de realizar exitosamente la habitación, los sujetos ingresaron a un programa de automoldeamiento (Brown & Jenkins, 1968), con el objetivo de desarrollar la respuesta de picoteo hacia una tecla, que sería la respuesta operante requerida en la tarea de elección subóptima.

El programa consistía en la iluminación de la tecla central de color blanco en cada ensayo por 10 s y seguida por la activación del comedero durante 3 s, todos los estímulos permanecieron inactivos y el siguiente ensayo comenzó después de 52 s. Las respuestas a la tecla iluminada no eran necesarias para acceder al comedero. El programa finalizó después de 60 ensayos y se empleó a lo largo de 3 sesiones. Los sujetos que habían realizado más del 50% de ensayos con respuestas de picoteo a la tecla iluminada avanzaban a la siguiente fase del experimento. Los sujetos que no habían realizado exitosamente las respuestas operantes en esta condición eran moldeados manualmente a través de acercamientos sucesivos por un máximo de 5 sesiones, una vez concluidas dicho número de sesiones, los sujetos que permanecían sin registro de respuestas fueron excluidos del experimento.

Pre-entrenamiento

Después de que las palomas aprendieron la respuesta operante de picoteo se les entrenó a realizar dicha respuesta a las teclas derecha e izquierda con diferentes colores (rojo, blanco y verde) presentadas de manera aleatoria bajo un programa de razón fija 1 (RF1) con la finalidad de evitar algún sesgo de posición por alguna de las teclas. La misma tecla no podía presentarse más de 3 veces de manera consecutiva. La respuesta a la tecla activa era seguida por 3 s de acceso al comedero, la tecla opuesta se mantenía inactiva. El siguiente ensayo iniciaba después de un intervalo entre ensayos (IEE) de 7 s de duración. El programa finalizó después de 1 hora o al completar 72 ensayos, 36 en cada tecla, 12 con la tecla iluminada blanca, 12 verde y 12 roja. El criterio para avanzar a la siguiente fase era que las medianas de las latencias de respuesta entre ambas teclas fuera menor a 1 s de diferencia. Esta condición constaba de 1 a 3 sesiones aproximadamente.

Entrenamiento. Procedimiento de elección subóptima

En el inicio de la tarea puede aparecer de manera aleatoria un ensayo forzado o un ensayo de elección. En ensayos forzados, una única tecla se iluminaría de blanco, en ya sea el sitio izquierdo o derecho, mientras la tecla opuesta se mantendría oscura. Un picotazo en la tecla blanca llevaría a la aparición de un estímulo coloreado en esa misma tecla. En una tecla, digamos izquierda, el 20% de los ensayos, un picotazo a la tecla blanca sería seguida por la iluminación de algún color, por ejemplo, rojo, durante un TF 10 s que siempre será seguido de reforzamiento (3 segundos de acceso al comedero) y 80% de los ensayos, una respuesta a esa misma tecla podía iluminarla con un estímulo diferente, por ejemplo, verde, durante un TF10 s que siempre era seguida por no reforzamiento. En esta alternativa la probabilidad global de reforzamiento era del 20% y debido a que ambos estímulos señalan diferencialmente el resultado de comida o no comida esta alternativa ha sido denominada en la literatura de elección subóptima como “alternativa discriminativa”.

En la presencia de la tecla opuesta, digamos la tecla derecha, el 20% de los ensayos, una respuesta iluminaba la tecla con un estímulo, por ejemplo, roja, durante un TF10 s que el 50% de las veces sería seguido por reforzamiento y, por otra parte, en el 80% de los ensayos en esa alternativa, una respuesta a la tecla blanca la iluminaría con un estímulo diferente, por ejemplo, verde, durante un TF10s que el 50% de las veces era seguido por reforzamiento. Entonces, para esta alternativa, las palomas obtendrían una probabilidad global de reforzamiento del 50% y debido a que ambos estímulos no señalan diferencialmente el resultado de comida o no comida esta alternativa ha sido denominada como “alternativa no discriminativa”. Los estímulos utilizados fueron contrabalanceados entre los sujetos.

Durante ensayos de elección, ambas alternativas fueron simultáneamente presentadas (ambas teclas estaban iluminadas de blanco) y una respuesta activaría alguno de los estímulos asociados con las mismas contingencias de reforzamiento que en los ensayos forzados. La opción no elegida sería cancelada.

Debido a que los sujetos presentaron una extinción de respuesta en el eslabón inicial ante ensayos forzados (especialmente en la tecla asociada con la alternativa discriminativa), se añadió a la tarea la posibilidad de que pasados 5 s sin respuesta a la tecla blanca del eslabón inicial se pasaría automáticamente al eslabón final, y se registraron los ensayos omitidos en cada sesión. Durante los ensayos de elección no había esta posibilidad de omisión, los sujetos debían responder ante alguna de las alternativas para acceder al eslabón final.

Cada sesión consistió en 60 ensayos que fueron divididos en 20 ensayos forzados para cada alternativa y 20 ensayos de elección, todos presentados de forma aleatoria. Los ensayos serían separados por un intervalo entre ensayos de 10 s, durante el cual todos los estímulos estarían inactivos. El criterio para que los animales pasarán a la siguiente condición era cumplir un mínimo de 20 sesiones y que los datos mantuvieran una proporción de elección estable en las últimas 5 sesiones.

Entrenamiento en reversión. Procedimiento de elección subóptima

En esta última fase, que sería conducida de igual manera por un mínimo de 20 sesiones y hasta encontrar estabilidad en la proporción de elección de los sujetos en sus últimas 5 sesiones, la locación de la alternativa discriminativa y no discriminativa sería cambiada al sitio opuesto, con

el objetivo de comprobar que la preferencia por una alternativa fue debida a las contingencias asociadas a cada alternativa y descartar la posibilidad de un sesgo de posición.

RESULTADOS

Habitación

Los sujetos que cumplieron 20 entradas al comedero durante la condición de habitación podían pasar a la condición de automoldeamiento al día siguiente, únicamente un sujeto, SN2, no cumplió el criterio por lo que fue excluido del experimento.

Automoldeamiento

Con el propósito de conocer si el fármaco haloperidol tenía el efecto de disminuir el valor incentivo de la tecla para las palomas, se administró de manera crónica desde el inicio del experimento a dos grupos, Dosis Alta (0.2 mg/kg) y Dosis Baja (0.1 mg/kg). Bajo esta hipótesis, en la fase de automoldeamiento los resultados esperados serían que el grupo control pudiera desarrollar la respuesta a la tecla de manera independiente a la entrega de alimento y, por otra parte, los grupos experimentales no presentaran este fenómeno del mismo modo. La tabla 1, muestra los resultados individuales de cada grupo obtenidos durante las 3 sesiones llevadas a cabo en la condición de automoldeamiento.

Tabla 1

Respuestas Individuales Durante las 3 Sesiones en la Condición de Automoldeamiento.

Grupo Control (salina)				Grupo Dosis Baja (0.1 mg/kg)				Grupo Dosis Alta (0.2 mg/kg)			
Sujetos	Sesión 1	Sesión 2	Sesión 3	Sujetos	Sesión 1	Sesión 2	Sesión 3	Sujetos	Sesión 1	Sesión 2	Sesión 3
SN1	0	59	59	SN13	0	9	12	SN7	0	0	0
SN3	0	4	53	SN14	0	0	0	SN9	0	1	10
SN4	37	58	55	SN15	0	0	0	SN10	0	0	0
SN5	18	59	51	SN16	0	0	13	SN11	0	0	0
SN6	0	0	0	SN17	0	0	0	SN12	0	28	54
SN8	4	0	3	SN18	0	0	0				

Los sujetos que no cumplieron por lo menos 30 respuestas en alguna de las 3 sesiones, equivalente al 50% de las respuestas en una sesión, pasaron a una condición de moldeamiento manual para asegurar que pudieran realizar la respuesta operante necesaria para la tarea de elección subóptima.

Los resultados muestran que para el grupo Control sólo el sujeto SN6 no realizó ninguna respuesta y el sujeto SN8 no cumplió el criterio de realizar al menos 30 respuestas. Para el grupo de Dosis Alta únicamente el sujeto SN12, cumplió el criterio al completar 54 ensayos en la sesión 3. Finalmente, en el grupo Dosis Baja, ningún sujeto pudo completar el criterio establecido.

Pre-entrenamiento

En esta condición los sujetos que lograron completar 1 sesión con 60 ensayos y las respuestas dadas en las dos teclas de respuesta, izquierda y derecha, con diferencia en latencias menor a 1 s, podían pasar a la tarea de elección subóptima, esto con el propósito de intentar evitar que los sujetos tuvieran un sesgo de posición por alguna de las teclas. El sujeto SN15, no cumplió el criterio, por lo que fue excluido del experimento.

Proporción de elección

Se calculó la proporción de elección por la alternativa subóptima (discriminativa) durante los 20 ensayos libres del procedimiento a lo largo de las sesiones desarrolladas en las fases de entrenamiento y reversión. Los resultados de los sujetos SN3 y SN14 no fueron considerados debido a que mostraron un sesgo por una de las teclas reflejado durante la condición de reversión y, por lo tanto, sus resultados no mostraban una preferencia real por alguna de las alternativas.

La proporción de elección está definida por un valor que puede ir de 0 a 1, donde un valor de 0 indicaría una proporción óptima; 0.5 una proporción intermedia o de indiferencia por alguna alternativa; y 1, una proporción subóptima. Debido a que los sujetos pertenecientes a los grupos de Dosis Alta y Baja corrieron diferente número de sesiones con el propósito de encontrar estabilidad en su preferencia de elección, la Figura 4 muestra únicamente el promedio de la proporción de elección obtenida por cada grupo, Dosis Baja, Alta y Control durante las primeras 20 sesiones de entrenamiento y las primeras 20 sesiones de reversión con la finalidad de observar el desarrollo de la preferencia entre los grupos.

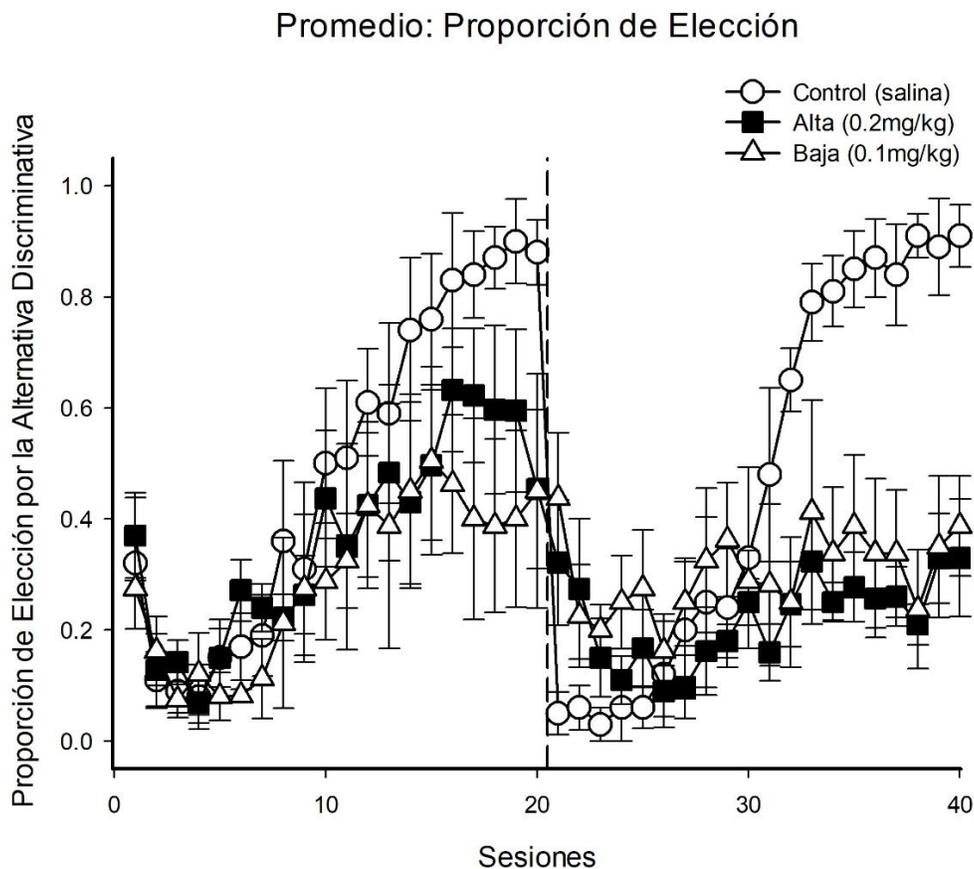


Figura 4 – Promedio entre grupos de elección por la alternativa discriminativa durante las primeras 20 sesiones realizadas en las fases de entrenamiento y reversión. La línea vertical discontinua de color negro indica el final de la fase de entrenamiento y el inicio de la fase de reversión. Las barras representan el error estándar con respecto a la media.

En las figuras 5 a 7 se muestra la proporción de elección obtenida por cada sujeto.

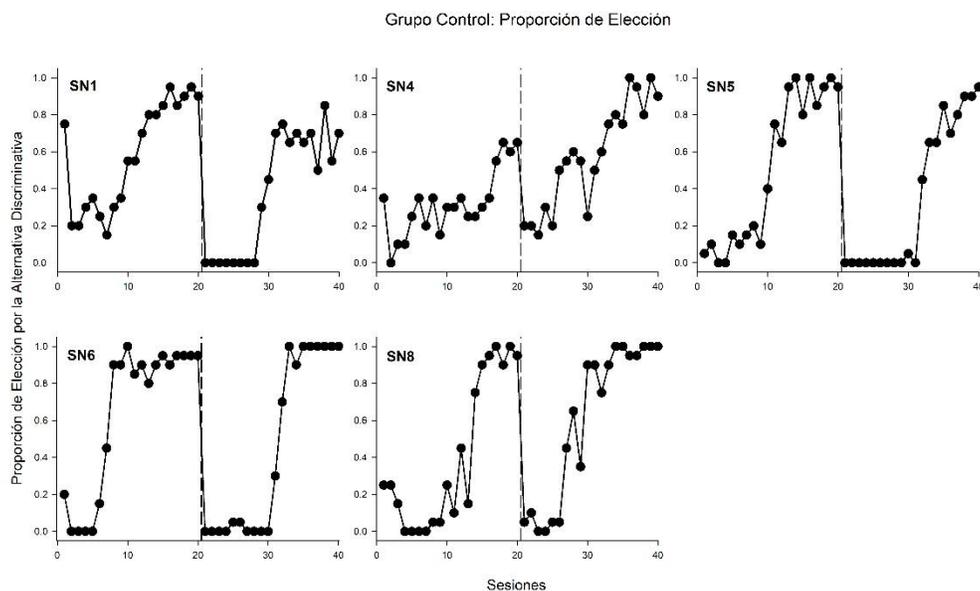


Figura 5 – Proporción de elección individual en el grupo Control por la alternativa discriminativa durante cada sesión realizada en las fases de entrenamiento y reversión. La línea vertical discontinua de color negro indica el final de la fase de entrenamiento y el inicio de la fase de reversión.

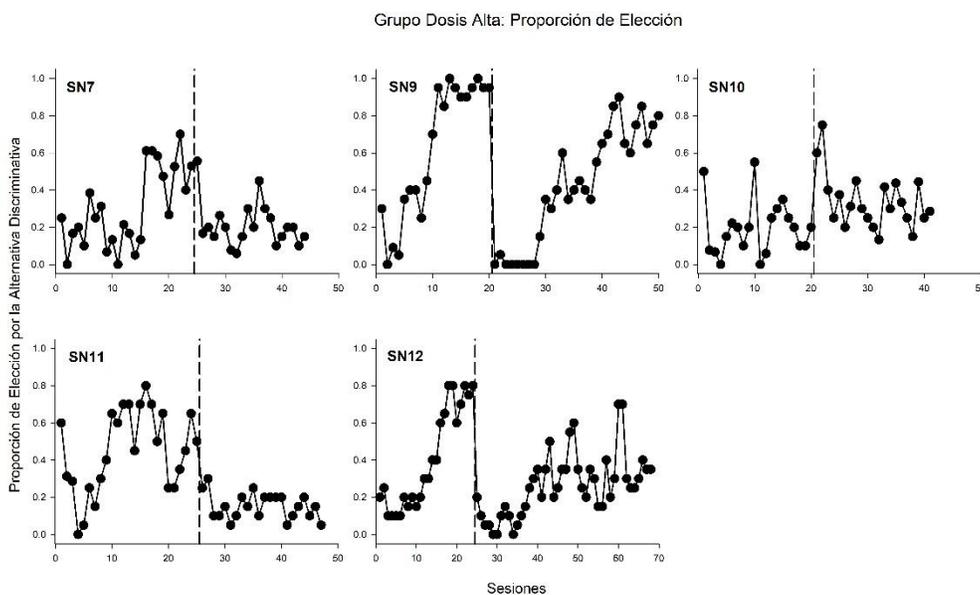


Figura 6 – Proporción de elección individual en el grupo Dosis Alta por la alternativa discriminativa durante cada sesión realizada en las fases de entrenamiento y reversión. La línea vertical discontinua de color negro indica el final de la fase de entrenamiento y el inicio de la fase de reversión.

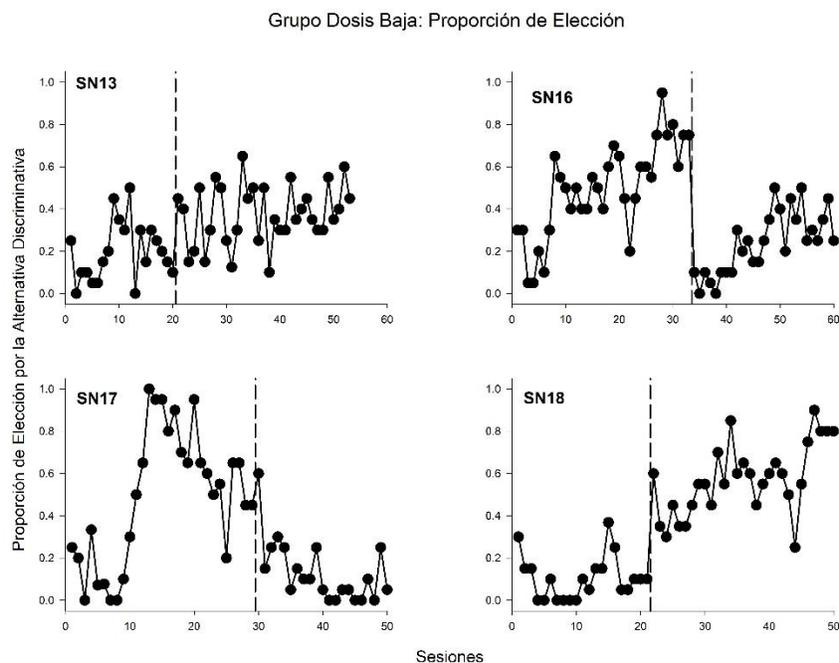


Figura 7 – Proporción de elección individual en el grupo Dosis Baja por la alternativa discriminativa durante cada sesión realizada en las fases de entrenamiento y reversión. La línea vertical discontinua de color negro indica el final de la fase de entrenamiento y el inicio de la fase de reversión.

Con la finalidad de conocer si existen diferencias en el nivel de preferencia por la alternativa discriminativa entre los tres grupos, se llevó a cabo un análisis de varianza de las últimas 5 sesiones de cada condición (entrenamiento y reversión). El análisis determinó que existen diferencias estadísticamente significativas entre los grupos, $F(2,11) = 10.55$, $p < 0.01$, los efectos de condición e interacción Grupos*Condición no mostraron significancia $F(2,11) = 0.94$, $p = 0.41$; $F(1,11) = 0.32$, $p = 0.57$, respectivamente. Un análisis Post Hoc Tukey reveló que las diferencias existen entre el grupo Control y el grupo de Dosis Alta en la condición de reversión ($p = 0.01$).

Para conocer si la preferencia obtenida por cada grupo era significativamente diferente de una preferencia indiferente, se llevó a cabo una Prueba *T* de Student entre las últimas 5 sesiones

de cada grupo contra una proporción de 0.5 (índice de indiferencia), tanto para la condición de entrenamiento como de reversión. Los resultados indicaron que durante la condición de entrenamiento el grupo Control obtuvo proporción de 0.86 ± 0.17 (Media \pm SEM), que fue estadísticamente diferente de indiferencia $t(4) = 4.76$, $p < 0.01$. Durante la condición de reversión, la preferencia obtenida en promedio por el grupo Control se mantuvo por la misma alternativa preferida en la condición de entrenamiento. En esta condición el promedio de la proporción de elección por la alternativa discriminativa fue de 0.88 ± 0.13 , que fue estadísticamente diferente de indiferencia $t(4) = 6.22$, $p < 0.01$.

Se llevó a cabo el mismo análisis para los grupos de Dosis Baja y Alta. Para el grupo de Dosis Alta, se obtuvo un promedio de proporción de elección de 0.55 ± 0.29 en la condición de entrenamiento, que no fue estadísticamente diferente de indiferencia $t(4) = 0.41$, $p = 0.7$. En la condición de reversión se obtuvo un promedio de 0.32 ± 0.25 , no diferente estadísticamente de indiferencia $t(4) = -1.52$, $p = 0.20$.

En cuanto al grupo de Dosis Baja, se obtuvo un promedio de proporción de elección de 0.37 ± 0.29 en la condición de entrenamiento, que no fue estadísticamente diferente de indiferencia $t(3) = -0.87$, $p = 0.44$. En la condición de reversión se obtuvo un promedio de 0.42 ± 0.30 , que tampoco fue estadísticamente diferente de indiferencia. $t(3) = -0.5$, $p = 0.64$.

Respuestas a la tecla

Para lograr identificar que los resultados obtenidos por los sujetos estaban asociados a la discriminación de las contingencias de los estímulos utilizados, se analizaron las respuestas dadas

en presencia de los estímulos, con la expectativa de que en la alternativa discriminativa habría un mayor número de respuestas para el estímulo predictor de reforzador en comparación con el estímulo predictor de no reforzador. Por otra parte, dentro de la alternativa no discriminativa se esperaría tener un número similar de respuestas entre los estímulos que la componen. Para ello se realizó un análisis del índice de discriminación entre alternativas, que consiste para el caso de la alternativa discriminativa, en la división de las respuestas dadas en el estímulo reforzado entre la suma total de respuestas en esa alternativa, un índice cercano a 1, indicaría una alta discriminación entre los estímulos que la componen. En cuanto a la alternativa no discriminativa, consistiría en la división de uno de los estímulos sobre el total de respuestas en esa alternativa, un índice cercano a 0.5 indicaría que no existe discriminación entre ambas alternativas.

Los promedios del índice de discriminación de las últimas 5 sesiones de cada grupo muestran similitudes para la alternativa discriminativa y no discriminativa tanto en la condición de entrenamiento como de reversión. Los resultados indicaron que para el grupo Control durante la condición de entrenamiento se obtuvo una media de 0.89 ± 0.13 . Durante la condición de reversión, el índice obtenido en promedio fue cercano al obtenido en la condición de entrenamiento, 0.87 ± 0.13 . Para la alternativa no discriminativa en la condición de entrenamiento se obtuvo una media de 0.51 ± 0.03 , y para la condición de reversión el promedio fue de 0.48 ± 0.03 .

Para el grupo de Dosis Alta, se obtuvo un promedio del índice de discriminación de 0.96 ± 0.01 en la condición de entrenamiento. En la condición de reversión se obtuvo un promedio de

0.92 ± 0.09 . Para la alternativa no discriminativa se obtuvo una media de 0.52 ± 0.1 en la condición de entrenamiento, y para la condición de reversión el promedio fue de 0.49 ± 0.02 .

Para el grupo de Dosis Baja, el índice promediado en la condición de entrenamiento fue de 0.85 ± 0.08 , mientras que en la condición de reversión se obtuvo un promedio de 0.88 ± 0.04 . Para la alternativa no discriminativa en la condición de entrenamiento se obtuvo una media de 0.47 ± 0.03 , y para la condición de reversión el promedio fue de 0.53 ± 0.002 .

Las figuras 8 a 10, muestran los promedios de los índices de respuestas realizadas durante las primeras 20 sesiones de entrenamiento y las primeras 20 sesiones después de reversión.

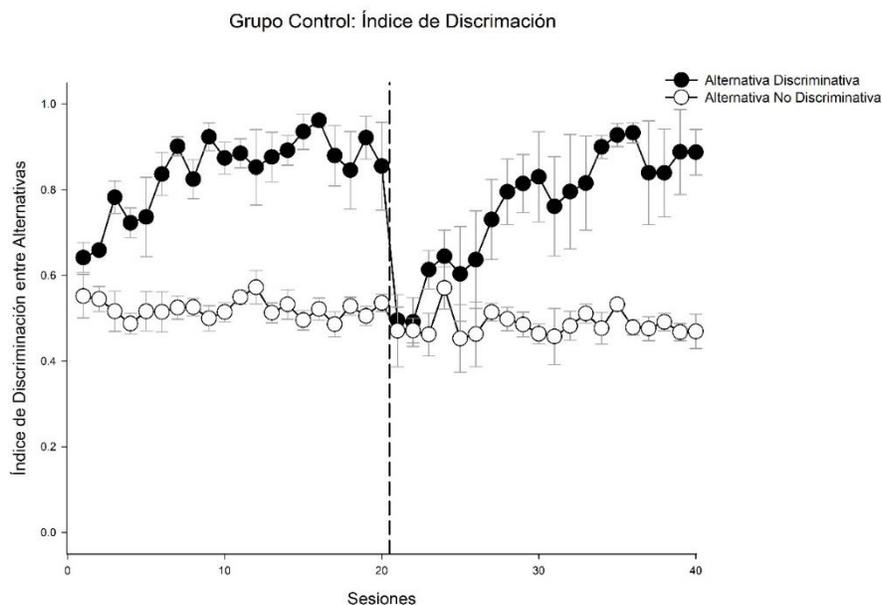


Figura 8 – Índice de Discriminación en el grupo Control para las alternativas discriminativa (círculos cerrados) y no discriminativa (círculos abiertos) durante cada sesión realizada en las fases de entrenamiento y reversión. La línea vertical discontinua de color negro indica el final de la fase de entrenamiento y el inicio de la fase de reversión.

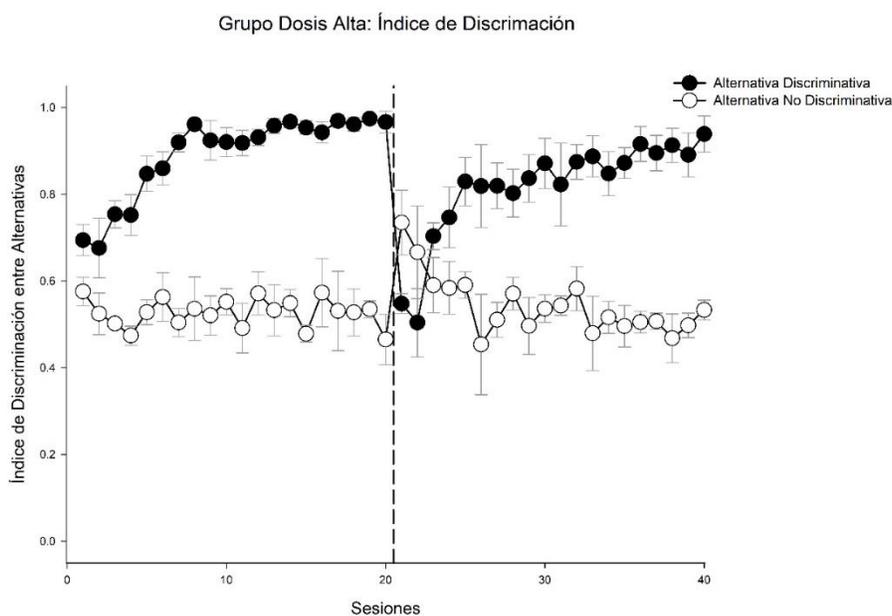


Figura 9 – Índice de Discriminación en el grupo Dosis Alta para las alternativas discriminativa (círculos cerrados) y no discriminativa (círculos abiertos) durante cada sesión realizada en las fases de entrenamiento y reversión. La línea vertical discontinua de color negro indica el final de la fase de entrenamiento y el inicio de la fase de reversión.

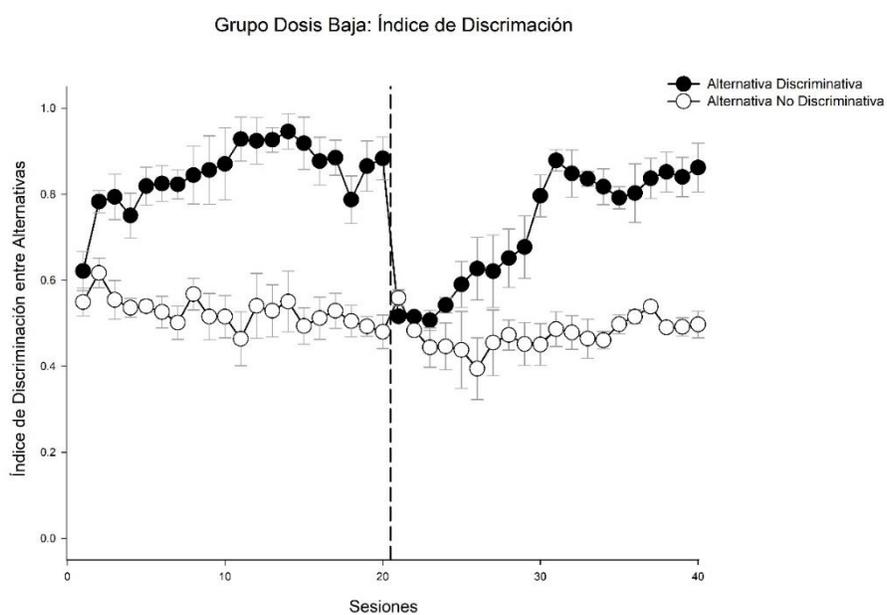


Figura 10 – Índice de Discriminación en el grupo Dosis Baja para las alternativas discriminativa (círculos cerrados) y no discriminativa (círculos abiertos) durante cada sesión realizada en las fases de entrenamiento y reversión. La línea vertical discontinua de color negro indica el final de la fase de entrenamiento y el inicio de la fase de reversión.

Se realizó un análisis de varianza para determinar si existían diferencias estadísticamente significantes en los índices de discriminación tanto para la condición de entrenamiento como de reversión considerando las últimas 5 sesiones de cada sujeto. Los resultados indicaron que para la condición de entrenamiento, existen diferencias significativas entre las alternativas $F(1,11) = 192.87$, $p < 0.01$, pero no las hay entre los grupos $F(2,11) = 1.84$, $p = 0.20$, ni en la interacción Grupos*Alternativas $F(2,11) = 0.49$, $p = 0.62$. Un análisis Post Hoc Tukey, indicó que existían diferencias entre los índices de ambas alternativas en todos los grupos, Control, Dosis Alta y Baja $p < 0.01$; $p < 0.01$; $p < 0.01$, respectivamente. Para la condición de reversión, del mismo modo se encontraron diferencias significativas entre las alternativas discriminativa y no discriminativa $F(1,11) = 240.21$, $p < 0.001$, pero no entre los grupos $F(2,11) = 0.39$, $p = 0.68$, ni en la interacción $F(2,11) = 0.86$, $p = 0.44$. Un análisis Post Hoc Tukey, indicó que existían diferencias entre los índices de ambas alternativas en todos los grupos, Control, Dosis Alta y Baja $p < 0.01$; $p < 0.01$; < 0.01 , respectivamente.

DISCUSIÓN

El estudio de elección subóptima ha causado gran interés en la última década debido a que muestra como palomas llevan a cabo una elección por una alternativa que les otorga poco alimento pero que proporciona estímulos informativos, sobre otra alternativa que es más rica en alimento, pero carece de dicha información (Stagner & Zentall, 2010). Por este motivo, se ha planteado que los resultados de elección subóptima en palomas podrían extrapolarse a un modelo de juego patológico en humanos, ya que, en un contexto de apuesta, apostadores suelen llevar a cabo su elección por estímulos predictores de ganancias, aunque estos ocurran con poca frecuencia.

Dentro del estudio de elección subóptima se han propuesto diversas teorías explicativas sobre el porqué palomas llevan a cabo una elección por la alternativa que ofrece menor tasa de reforzamiento global (Chow et al., 2017; Stagner & Zentall, 2010; Zentall, 2016). Sin embargo, la mayoría de ellas al ponerse a prueba empíricamente muestran resultados contradictorios que han impedido establecer una única explicación del fenómeno (Martínez et al., 2017; Trujano et al., 2016) y, por lo tanto, la necesidad de crear nuevas vías de investigación. El presente estudio tuvo como objetivo evaluar si el valor incentivo de los estímulos, una de las aproximaciones más novedosas en el tema, juega un papel importante en la tarea de elección subóptima.

En el juego patológico, las personas prefieren una opción que, al señalar una ganancia, viene acompañada de estímulos contextuales con alto valor incentivo, como lo pueden ser diferentes luces, música resonante o refuerzo social, que pueden opacar las pérdidas obtenidas que carecen de dicha saliencia incentiva. Se ha propuesto, que dentro de la tarea de elección subóptima, las teclas iluminadas con palomas, poseen alto valor incentivo para esta especie ya que generan el efecto de seguimiento de señales, una característica principal de la saliencia incentiva (Meyer et al., 2012). Es por ello que analizar el valor incentivo puede ser un buen acercamiento hacia la explicación de los resultados encontrados con palomas, dado que prefieren una alternativa discriminativa compuesta por una opción que señala ganancias una vez que aparece, y otra opción que, aunque señala que no habrá ganancias, puede ser ignorada o poco valorada.

Para evaluar si el valor incentivo es parte clave en la elección de palomas, la propuesta del experimento fue afectar el procesamiento de dicho valor, para ello se llevó a cabo el procedimiento

clásico de elección subóptima, administrando a dos grupos, diferentes dosis del fármaco haloperidol, un antagonista dopaminérgico del receptor DA2, el cual se ha visto puede producir un bloqueo en el valor incentivo de los estímulos reforzantes (Acerbo et al., 2003). La hipótesis de este trabajo consiste en que al utilizar un antagonista dopaminérgico, podría bloquearse la saliencia incentiva perteneciente al estímulo condicionado, predictor de recompensa y, por lo tanto, se encontraría dentro del procedimiento clásico de elección subóptima en palomas, una toma de decisión óptima basada exclusivamente en la mayor tasa de reforzamiento.

Los resultados de este estudio sustentan la hipótesis mencionada anteriormente, se puede observar, por una parte, que para el grupo Control al que le fue administrado salina, presentó el mismo patrón reportado en todos los estudios realizados con palomas, es decir, una fuerte preferencia por la alternativa discriminativa o subóptima. Por otro lado, se pueden observar resultados diferentes en los grupos experimentales, tanto en el grupo de Dosis Alta como el grupo de Dosis Baja existe una preferencia menor por la alternativa discriminativa, encontrando una diferencia significativa entre el grupo Control y de Dosis Alta para la condición de reversión. A pesar de no haber diferencias estadísticamente significativas entre el grupo Control y Dosis Baja, se puede notar que este último muestra en el promedio grupal una clara tendencia hacia la optimalidad, tal como el grupo de Dosis Alta.

Aunque las proporciones de los grupos experimentales no son lo suficientemente bajas como para ser estadísticamente diferentes de indiferencia y ser consideradas estrictamente óptimas, sin duda, se puede hablar de la existencia de una disminución en la proporción de elección

debida a un efecto por parte del fármaco, es decir, de manera más específica, haloperidol pudo disminuir la saliencia incentiva de las teclas iluminadas y, por lo tanto, evitar una preferencia por la alternativa discriminativa. Además, el análisis de los índices de discriminación de los tres grupos muestran una alta discriminación entre los estímulos pertenecientes a la alternativa discriminativa en las condiciones de entrenamiento y reversión, lo que permite descartar dos cosas, en primer lugar, el hecho de que la disminución en la preferencia de elección encontrada en los grupos experimentales pueda ser debida a una falla en la discriminación de las alternativas, y en segundo lugar, descartar cuestiones de sedación por parte del fármaco al encontrar una correcta ejecución de respuestas entre alternativas por parte de los grupos experimentales en comparación con el grupo Control.

Además, es importante retomar la dificultad que se ha tenido en el estudio de elección subóptima por no encontrar generalidad entre especies en la tarea, ya que las ratas al contrario de palomas, presentan una fuerte preferencia por la alternativa óptima. Recientemente se ha propuesto un modelo que busca explicar de manera cuantitativa el fenómeno de elección subóptima y las características en las que dirigen su elección ratas y palomas, se trata del modelo teórico de información temporal (Cunningham & Shahan, 2018). En este modelo se predice que la preferencia por la alternativa discriminativa está dada por una competencia entre la información temporal y la tasa de reforzamiento:

$$p_{Sub} = w \frac{H^{a_{sub}}}{H^{a_{sub}} + H^{a_{opt}}} + (1 - w) \frac{R^{b_{sub}}}{R^{b_{sub}} + R^{b_{opt}}}$$

Donde H es la información temporal transmitida por estímulos del enlace terminal (ET) para cada alternativa, a se refiere a la sensibilidad de elección hacia la información temporal. Esta primera parte de la ecuación es capaz de ejemplificar que palomas prefieren la alternativa que provee los estímulos del ET con mayor información temporal. En la segunda parte, R es la tasa de reforzamiento asociada con cada una de las alternativas, b se refiere a la sensibilidad de elección hacia la tasa de reforzamiento. Las ratas basarían su elección contemplando en mayor medida la tasa de reforzamiento. Una tercera característica del modelo es w , que es un mecanismo de ponderación que determina el grado en que la elección es gobernada por la información temporal y la tasa de reforzamiento y se desglosa de manera más específica del siguiente modo:

$$w = \frac{1}{1 + e^{-\beta \left(\frac{Df}{Ds} - m \right)}}$$

Donde Df es el promedio del retraso del alimento en el punto de elección, Ds el promedio del retraso de las señales informativas temporales en el punto de elección, β se refiere a la sensibilidad de w a variaciones en Df/Ds , y m es un sesgo del uso de la información temporal transmitida por las señales para tomar decisiones.

La ecuación en general menciona que cuando hay un retraso largo para la comida (ET) relativo al retraso para la alternativa informativa temporalmente (enlace inicial o EI) en el punto de elección, daría un indicativo de que el valor de Df/Ds sería relativamente largo, entonces, w se aproxima a 1 y la elección debería ser gobernada por la información temporal proporcionada por las señales que son más cercanas en el tiempo que el alimento por sí mismo, y como resultado, se tendría una elección subóptima. Cuando w se aproxima a 0, la elección es gobernada por la tasa de

reforzamiento y esto originaría una elección óptima. El modelo indicaría que el problema por no encontrar elección subóptima en ratas es que el valor de Df/Ds no ha sido lo suficientemente grande para incrementar w , es decir, se debería buscar aumentar el ET (Cunningham & Shahan, 2019), o en su defecto, incrementar el EI y disminuir el ET.

Adicionalmente, el presente estudio da sustento a que el valor incentivo atribuido a los estímulos efectivamente juega un papel importante en el procedimiento de elección subóptima, el cuál sesga la preferencia hacia la alternativa discriminativa que posee el estímulo predictor de recompensa en lugar de aquella que resulta ser más rica en alimento. Los resultados de este estudio sustentan lo encontrado por Chow et al. (2017), quienes al modificar la tarea clásica con ratas en la cual se solían utilizar luces (cuyo valor incentivo es bajo), por palancas (cuyo valor incentivo es mayor), encontraron que ratas tienden a elegir en mayor medida la alternativa discriminativa. Esos argumentos están enlazados con la literatura de seguimiento de señales y seguimiento de objetivos, de modo puntual se puede decir que al utilizar luces o tonos con ratas como estímulos asociados a una recompensa se puede generar un efecto de seguimiento de objetivos, es decir, las ratas dirigen su respuesta hacia la fuente de alimento y, por otra parte, al utilizar palancas se puede generar en mayor medida un efecto de seguimiento de señales, es decir, las ratas dirigen su respuesta hacia el estímulo predictor de recompensa, en este caso, la palanca (Chow et al., 2016; Flagel, Watson, Robinson, & Akil, 2007).

Cabe mencionar que, aunque los resultados con palancas obtenidos por Chow et al. (2017) no han sido replicados (Alba et al., 2018; Martínez et al., 2017; Orduna & Alba, 2019), se ha propuesto también que los estímulos utilizados con alta saliencia incentiva que se utilizan en los

procedimientos, teclas para palomas y palancas para ratas, tienen un valor no equitativo en estas especies, lo que podría ser un factor importante para generar seguimiento de señales, es decir, prácticamente todas las palomas son capaces de responder a una tecla en automoldeamiento aun cuando la entrega de comida sea independiente de la respuesta (Brown & Jenkins, 1968) -tal como se observó de igual manera en el grupo Control de este experimento-, pero dicho automoldeamiento ocurre solamente el 35% de las veces al utilizar palancas con ratas (Meyer et al., 2012).

Recientemente, Orduña y Alba (2020), han encontrado que cuando la alternativa discriminativa tiene alto valor incentivo (palanca) y la alternativa no discriminativa tiene bajo valor incentivo (luces), las ratas presentan una tendencia a elegir en mayor grado la alternativa con alto valor incentivo, aunque esto signifique menor alimento obtenido en la tarea, y que dichos resultados son sostenidos con ratas seguidoras de señales, en comparación con ratas seguidoras de objetivos. En ese sentido, sería necesario encontrar para ratas un estímulo que posea la misma saliencia incentiva que la tecla representa en palomas, de ese modo podría generarse una comparación homóloga y así, esperar encontrar en ratas suboptimalidad en un nivel semejante al que se desarrolla en palomas.

Del mismo modo que en este experimento, una propuesta para futuras investigaciones es el acercamiento hacia la farmacología, ya que podría manipularse de manera farmacológica el valor incentivo en ratas. Dados los resultados encontrados múltiples veces en el procedimiento de elección subóptima con palomas, en este estudio se buscó disminuir el valor incentivo de los

estímulos empleados a través de la administración de un antagonista dopaminérgico y, si el objetivo fuera aumentar dicho valor incentivo en los estímulos utilizados con ratas (Robinson, Anselme, Suchomel, & Berridge, 2015), entonces se podrían utilizar agonistas dopaminérgicos y estudiar su efecto en la tarea de elección subóptima en espera de encontrar resultados que apunten a la subóptimalidad, acorde a lo encontrado con palomas y haloperidol.

En conclusión, los resultados obtenidos en este estudio dan sustento a la existencia de un efecto por parte del valor incentivo en la tarea de elección subóptima, que, por otra parte, también está relacionado con el efecto que la saliencia incentiva puede tener en el juego patológico. Por ello, es importante desarrollar mayor investigación dentro del área de elección subóptima y valor incentivo para poder reafirmar su papel en el procedimiento y poder consolidar los hallazgos entre las diversas especies, no sólo ratas y palomas. Además, mayor investigación podría dar sustento a tratamientos eficaces que puedan disminuir padecimientos patológicos que engloban una mala toma de decisiones, como lo es, la impulsividad, el juego patológico y abuso de sustancias.

REFERENCIAS

- Acerbo, M. J. (2001). *The Role of Dopamine and Glutamate in Associative Learning by the Pigeon (Columba livia)*. (Doctoral dissertation).
- Acerbo, M. J., Godoy, A. M., & Delius, J. D. (2003). Haloperidol blocks the acquisition but not the retrieval of a conditioned sensitization to apomorphine. *Behavioural Pharmacology*, *14*(8), 631-640. doi:<https://doi.org/10.1097/00008877-200312000-00007>
- Alba, R., Rodríguez, W., Martínez, M., & Orduña, V. (2018). Rats' preferences in the suboptimal choice procedure: Evaluating the impact of reinforcement probability and conditioned inhibitors. *Behavioural Processes*, *157*, 574-582. doi:<https://doi.org/10.1016/j.beproc.2018.04.013>
- Baumann, A. A., & Odum, A. L. (2012). Impulsivity, risk taking, and timing. *Behavioural Processes*, *90*(3), 408-414. doi:<https://doi.org/10.1016/j.beproc.2012.04.005>
- 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. doi:<https://doi.org/10.1101/lm.037382.114>
- 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. doi:<https://doi.org/10.1901/jeab.1994.62-353>
- Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology*, *191*(3), 391-431. doi:<https://doi.org/10.1007/s00213-006-0578-x>
- Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28*(3), 309-369. doi:[https://doi.org/10.1016/S0165-0173\(98\)00019-8](https://doi.org/10.1016/S0165-0173(98)00019-8)
- Broos, N., Schmaal, L., Wiskerke, J., Kosteljik, L., Lam, T., Stoop, N., Goudriaan, A. E. (2012). The Relationship between Impulsive Choice and Impulsive Action: A Cross-Species Translational Study. *PLoS ONE*, *7*, e36781.
- Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, *11*(1), 1-8. doi:<https://doi.org/10.1901/jeab.1968.11-1>
- 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. doi:<https://doi.org/10.1901/jeab.1963.6-299>
- Chow, J. J., Nickell, J. R., Darna, M., & Beckmann, J. S. (2016). Toward isolating the role of dopamine in the acquisition of incentive salience attribution. *Neuropharmacology*, *109*, 320-331. doi:<https://doi.org/10.1016/j.neuropharm.2016.06.028>
- 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. doi:<https://doi.org/10.1016/j.bbr.2016.12.013>
- Chung, S. H., & Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, *10*(1), 67-74. doi:<https://doi.org/10.1901/jeab.1967.10-67>

- 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), 1-22. doi: <https://doi.org/10.1037/xan0000160>
- 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), 301-310. doi: <https://doi.org/10.1037/xan0000211>
- Danna, C. L., & Elmer, G. I. (2010). Disruption of conditioned reward association by typical and atypical antipsychotics. *Pharmacology Biochemistry and Behavior*, 96(1), 40-47. doi: <https://doi.org/10.1016/j.pbb.2010.04.004>
- Di Ciano, P., Cardinal, R. N., Cowell, R. A., Little, S. J., & Everitt, B. J. (2001). Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of Pavlovian approach behavior. *The Journal of Neuroscience*, 21(23), 9471-9477.
- DiFeliceantonio, A. G., & Berridge, K. C. (2016). Dorsolateral neostriatum contribution to incentive salience: opioid or dopamine stimulation makes one reward cue more motivationally attractive than another. *European Journal of Neuroscience*, 43(9), 1203-1218. doi: <https://doi.org/10.1111/ejn.13220>
- Domjan, M. (2010). *The principles of learning and behavior*: Cengage Learning.
- Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior*, 53(2), 201-218. doi: <https://doi.org/10.1901/jeab.1990.53-201>
- Fantino, E., & Abarca, N. (1985). Choice, optimal foraging, and the delay-reduction hypothesis. *Behavioral and Brain Sciences*, 8(2), 315-362. doi: <https://doi.org/10.1017/S0140525X00020847>
- Fantino, E., Dunn, R., & Meck, W. (1979). Percentage reinforcement and choice. *Journal of the Experimental Analysis of Behavior*, 32(3), 335-340. doi: <https://doi.org/10.1901/jeab.1979.32-335>
- Flagel, S. B., Watson, S. J., Robinson, T. E., & Akil, H. (2007). Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. *Psychopharmacology*, 191(3), 599-607. doi: <https://doi.org/10.1007/s00213-006-0535-8>
- Gipson, C. D., Alessandri, J. J. D., Miller, H. C., & Zentall, T. R. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning & Behavior*, 37(4), 289-298. doi: <https://doi.org/10.3758/LB.37.4.289>
- Grakalic, I., Panlilio, L. V., Thorndike, E. B., & Schindler, C. W. (2007). Differential involvement of dopamine receptors in conditioned suppression induced by cocaine. *European Journal of Pharmacology*, 573(1-3), 116-123. doi: <https://doi.org/10.1016/j.ejphar.2007.06.037>
- Güntürkün, O. (2005). The avian 'prefrontal cortex' and cognition. *Current Opinion in Neurobiology*, 15(6), 686-693. doi: <https://doi.org/10.1016/j.conb.2005.10.003>

- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4(3), 267-272. doi:<https://doi.org/10.1901/jeab.1961.4-267>
- Hollard, V., & Davison, M. C. (1971). Preference for qualitatively different reinforcers. *Journal of the Experimental Analysis of Behavior*, 16(3), 375-380. doi:<https://doi.org/10.1901/jeab.1971.16-375>
- Kalenscher, T., Windmann, S., Diekamp, B., Rose, J., Gunturkun, O., & Colombo, M. (2005). Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. *Current Biology*, 15(7), 594-602. doi:<https://doi.org/10.1016/j.cub.2005.02.052>
- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, 21(3), 463-473. doi: <https://doi.org/10.1901/jeab.1974.21-463>
- Langdon, A. J., Hathaway, B. A., Zorowitz, S., Harris, C. B. W., & Winstanley, C. A. (2019). Relative insensitivity to time-out punishments induced by win-paired cues in a rat gambling task. *Psychopharmacology (Berl)*, 236(8), 2543-2556. doi:<https://doi.org/10.1007/s00213-019-05308-x>
- 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), 2-11. doi:<https://doi.org/10.1037/xan0000001>
- 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), 12-21. doi:<https://doi.org/10.1037/xan0000010>
- 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. doi:<https://doi.org/10.1016/j.bbr.2017.12.028>
- Martelle, J. L., & Nader, M. A. (2008). A Review of the Discovery, Pharmacological Characterization, and Behavioral Effects of the Dopamine D2-Like Receptor Antagonist Eticlopride. *CNS Neuroscience & Therapeutics*, 14(3), 248-262. doi:<https://doi.org/10.1111/j.1755-5949.2008.00047.x>
- 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. doi:<https://doi.org/10.1016/j.beproc.2017.06.012>
- Meyer, 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, e38987. Retrieved from <https://ui.adsabs.harvard.edu/abs/2012PLoS...738987M>
- Moe, R. O., Nordgreen, J., Janczak, A. M., Spruijt, B. M., Kostal, L., Skjerve, E., Bakken, M. (2011). Effects of haloperidol, a dopamine D2-like receptor antagonist, on reward-related behaviors in laying hens. *Physiology & Behavior*, 102(3-4), 400-405. doi:<https://doi.org/10.1016/j.physbeh.2010.12.008>

- 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. doi:<https://doi.org/10.3758/s13420-012-0065-7>
- Orduna, V., & Alba, R. (2019). Rats' optimal choice behavior in a gambling-like task. *Behavioural Processes*, *162*, 104-111. doi: <https://doi.org/10.1016/j.beproc.2019.02.002>
- Orduña, V., & Alba, R. (2020). The incentive salience of the stimuli biases rats' preferences in the "suboptimal choice" procedure. *Behavioural Processes*, *172*, 104057. doi:<https://doi.org/10.1016/j.beproc.2020.104057>
- Pierce, W. D., & Cheney, C. D. (2004). *Behavior analysis and learning (3rd ed)*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Powell, R. W., Saunders, D., & Thompson, W. (1968). Shaping, auto-shaping and observational learning with rats. *Psychonomic Science*, *13*(3), 167-168. doi:<https://doi.org/10.3758/BF03342463>
- Puig, M. V., Rose, J., Schmidt, R., & Freund, N. (2014). Dopamine modulation of learning and memory in the prefrontal cortex: insights from studies in primates, rodents, and birds. *Frontiers in Neural Circuits*, *8*(93). doi: <https://doi.org/10.3389/fncir.2014.00093>
- Reuter, J., Raedler, T., Rose, M., Hand, I., Gläscher, J., & Büchel, C. (2005). Pathological gambling is linked to reduced activation of the mesolimbic reward system. *Nature Neuroscience*, *8*(2), 147-148. doi: <https://doi.org/10.1038/nn1378>
- Robinson, M. J. F., Anselme, P., Suchomel, K., & Berridge, K. C. (2015). Amphetamine-induced sensitization and reward uncertainty similarly enhance incentive salience for conditioned cues. *Behavioral Neuroscience*, *129*(4), 502-511. doi: <https://doi.org/10.1037/bne0000064>
- 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. doi: <https://doi.org/10.1006/lmot.1999.1030>
- Shettleworth, S. J. (2010). *Cognition, evolution, and behavior (2nd ed)*. New York, NY, US: Oxford University Press.
- 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. doi:<https://doi.org/10.1016/j.lmot.2011.09.001>
- Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review*, *17*(3), 412-416. doi:<https://doi.org/10.3758/pbr.17.3.412>
- 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. doi:<https://doi.org/10.1016/j.beproc.2016.07.005>
- Trujano, R. E., & Orduna, V. (2015). Rats are optimal in a choice task in which pigeons are not. *Behavioural Processes*, *119*, 22-27. doi:<https://doi.org/10.1016/j.beproc.2015.07.010>
- Williams, D. R., & Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior*, *12*(4), 511-520. doi:<https://doi.org/10.1901/jeab.1969.12-511>

- Zarrindast, M. R., Hajian-Heydari, A., & Hoseini-Nia, T. (1992). Characterization of dopamine receptors involved in apomorphine-induced pecking in pigeons. *General Pharmacology: The Vascular System*, 23(3), 427-430. doi:[https://doi.org/10.1016/0306-3623\(92\)90106-t](https://doi.org/10.1016/0306-3623(92)90106-t)
- Zentall. (2016). An Animal Model of Human Gambling. *International Journal of Psychological Research*, 9, 96-112. doi:<http://dx.doi.org/10.21500/20112084.2284>
- Zentall, & Laude, J. R. (2013). Do Pigeons Gamble? I Wouldn't Bet Against It. *Current Directions in Psychological Science*, 22(4), 271-277. doi:<https://doi.org/10.1177/0963721413480173>
- Zentall, & 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. doi:<https://doi.org/doi:10.1098/rspb.2010.1607>