



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE PSICOLOGÍA

EVALUACIÓN DEL EFECTO DEL VALOR INCENTIVO
DE LOS ESTÍMULOS EN EL PROCEDIMIENTO DE
ELECCIÓN SUBÓPTIMA EN PALOMAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN PSICOLOGÍA

P R E S E N T A :

RODRIGO URIEL GONZÁLEZ TORRES

DIRECTOR: DR. ÓSCAR VLADIMIR ORDUÑA TRUJILLO

REVISORA: LIC. CRISTINA AVELAR RODRÍGUEZ

SINODALES: DR. ROGELIO ESCOBAR HERNÁNDEZ

DRA. ALICIA ROCA COGORDAN

DR. LUIS RODOLFO BERNAL GAMBOA



CIUDAD UNIVERSITARIA, CDMX, 2020



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

El presente trabajo se realizó con el apoyo del Consejo Nacional de Ciencia y Tecnología (CONACYT) **281548** y del Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) **IN306818**

Para Mariana

Por ser mi luz. Me motivas

a ser una mejor persona

Para Gonzalo y Martha
Por su apoyo incondicional y
por enseñarme a creer en mi

Agradecimientos

A mis padres, Gonzalo González y Martha Torres quienes me apoyaron a lo largo de toda mi trayectoria académica, quienes durante todos estos años creyeron en mí. Mi logro es también suyo.

A mi padre, por apoyarme siempre, incluso más de lo debido o lo pedido, por ser mi ejemplo y por enseñarme a ser valiente.

A mi madre, por su incondicional amor, por ayudarme a creer que los anhelos no siempre son imposibles y por enseñarme a ser fuerte.

A mi hermano Gonzalo, quien siempre se mostró dispuesto a ayudarme, fuiste mi primer aliado y mi primer amigo.

A la memoria de mis abuelos, Rafael Torres y Enriqueta Moreno, quienes tenían la esperanza de que mi hermano y yo termináramos una carrera universitaria.

A mi tía Elena González, por su cuidado y apoyo durante las situaciones más difíciles.

A Mariana, cuyo apoyo y paciencia no podría medir. Mariana es la persona que se queda, cuando parece que todos los demás se han ido. Gracias por recorrer todo este camino a mi lado.

A mi director de tesis el Dr. Vladimir Orduña, gracias por la confianza depositada en mí y por la oportunidad de formar parte de tu laboratorio.

A mis compañeros del laboratorio 101. A Julio quien me apoyo durante todo este trabajo a pesar de todo el tiempo que ello consumió. A mi primer *sensei* Enrique, por instruirme en el ámbito de la investigación, cuyas enseñanzas situaron la base de mis técnicas. A mi segundo *sensei* Rodrigo, quien vio por mí en más de una ocasión y a cuyas enseñanzas atribuyo mi interés en la electrónica. A William, por su script de resumen y por enseñarme como la programación puede hacerte la vida más fácil. También, al resto de mis compañeros quienes, aunque no contribuyeron directamente en mi trabajo, su compañía hizo mi estancia en el laboratorio más agradable.

A mi revisora la Lic. Cristina Avelar, quien, en su trabajo y dedicación como docente, sembró en mí el interés por esta área de la psicología.

A todo mi comité, por sus valiosas aportaciones, las cuales ayudaron a enriquecer y llevar este trabajo a su forma final.

A la UNAM por la calidad de la educación brindada y por volverse un espacio de oportunidades y experiencias.

Índice

Resumen	8
Introducción	9
Elección en el análisis experimental de la conducta	10
Elección subóptima	12
La atribución de saliencia incentiva a los estímulos	23
Las operantes y su relación con la atribución de saliencia incentiva	28
Experimento 1. Condición de bajo valor incentivo	30
Método	30
Sujetos	30
Aparatos	30
Procedimiento	33
Habituaación	33
Moldeamiento en presión a la palanca	33
Entrenamiento en presión a la palanca	34
Pre-exposición a los estímulos	34
Entrenamiento. Procedimiento de elección subóptima	35
Entrenamiento en reversión	36
Resultados	37
Experimento 2. Condición de alto valor incentivo	42
Método	42
Sujetos	42
Aparatos	42
Procedimiento	45

Habituaación	45
Automoldeamiento	45
Moldeamiento en picoteo a la tecla	45
Entrenamiento en picoteo a la tecla	45
Entrenamiento. Procedimiento clásico de elección subóptima	46
Entrenamiento en reversión	47
Resultados	48
Comparación entre condiciones	52
Discusión	55
La elección óptima dependió del uso de luces ambientales	57
La elección óptima dependió del cambio de operante	58
La interacción entre el uso de luces ambientales y el cambio en la operante, como responsable de la elección óptima	59
Referencias	61
Apéndice	69

Resumen

En el fenómeno de elección subóptima, las palomas eligen una alternativa con estímulos que indican la presencia de alimento pero que provee una menor tasa global de reforzamiento, frente a otra que provee más comida, pero sin estímulos señalizadores. Aunque las causas involucradas en esta elección aún no se han establecido de forma concluyente, recientes estudios con ratas sugieren que este fenómeno puede deberse a la atribución de saliencia incentiva de los estímulos. En el primer experimento del presente trabajo, se modificaron dos aspectos básicos del paradigma original, para reducir la atribución de saliencia incentiva, se sustituyeron los estímulos discriminativos por luces ambientales, y la operante de respuesta por pisadas a una palanca, logrando mostrar, por primera vez en esta tarea, conducta óptima en palomas. Por lo anterior, se realizó un segundo experimento como condición control, con las condiciones del procedimiento clásico, durante el cual los sujetos mostraron una preferencia subóptima, demostrando así la influencia de la manipulación del valor incentivo de los estímulos en la preferencia de los animales durante la tarea de elección subóptima.

Palabras clave: Elección subóptima, palomas, pisar palanca, seguidores de señales, valor incentivo

Introducción

Cada decisión que tomamos nos guía hacia diferentes destinos, acarrea distintas consecuencias y, a veces, optar por una alternativa impide disfrutar de otras opciones. A esto nos enfrentamos y seguiremos haciéndolo todos los días de nuestra vida.

Nuestras decisiones pueden resolver cuestionamientos sobre necesidades básicas como ¿qué comer?, o ¿a qué hora dormir?, y ya que casi siempre estas necesidades están cubiertas, tenemos tiempo de plantearnos otros problemas como ¿qué carrera elegir?. y ¿qué videojuego jugaré hoy? Pero decidir no es una cualidad exclusivamente humana, otros animales también tienen que tomar decisiones todo el tiempo, de hecho, se considera que si un organismo es capaz de emitir una conducta entonces es capaz de elegir, en palabras de William Baum (1981, p. 387) “...toda conducta constituye elección, en el sentido de que, no importa cuán limitada sea la situación, un organismo siempre tiene más de una actividad en la que puede participar”.

Mi perro, por ejemplo, todas las mañanas tiene que decidir entre comer su alimento recién servido o acercarse a la mesa y esperar a que le dé algo de lo que estoy comiendo, si decide comer su alimento mientras yo desayuno, al levantarme habrá perdido la posibilidad de obtener una botana que, de otra forma, difícilmente podría conseguir. Por ello, es casi seguro que, al verme sentado en la mesa, se dirija hacia mí, se siente y me mire directo a la cara poniendo su pata sobre mi pierna. Una vez que me levante y que él haya disfrutado de algunos bocadillos, se dirigirá a su tazón y continuará comiendo.

Si mi perro eligiera sólo comer el alimento que le sirvo en su plato, digamos que obtendría una unidad de comida (+1), en cambio, si se acerca a la mesa mientras yo estoy comiendo, puede obtener una unidad de comida (+1) y después continuar con lo que hay en su plato (+1) obteniendo un total de dos unidades de comida. De esta forma, mi perro estaría comportándose de forma óptima, en el sentido de que elige la alternativa con el mayor valor

(Shimp, 1969) o bien, de acuerdo con la teoría del forrajeo óptimo, mi perro maximiza, es decir opta por la alternativa que le proveen mayores ganancias, con los menores costos (Pyke, 1984).

En el enfoque en el cual está enmarcado este trabajo, que es el análisis experimental de la conducta, se estudia el comportamiento de los organismos, mediante técnicas ampliamente replicadas y confiables. Este trabajo se centrará en la conducta de elección, particularmente, en una tarea denominada elección subóptima (Stagner & Zentall, 2010), por lo que es importante describir, primero, la forma típica en que se suele estudiar la elección y posteriormente se abordará la tarea; sus explicaciones e implicaciones.

Elección en el análisis experimental de la conducta

Aun en una situación experimental muy simple, como podría ser poner a un animal en una caja de condicionamiento operante (ver Figura 1) y hacerlo presionar un botón, éste puede realizar diversas actividades sin estar limitado a aquello que espera el investigador. Por ejemplo, una paloma puede picar la única tecla de respuestas en la caja, acicalarse o moverse alrededor de ella. Entender los mecanismos de la elección y estudiarlos adecuadamente es fundamental para comprender la conducta, sin embargo, ya que la mayoría de nuestras acciones son el resultado de elegir una actividad sobre otra, analizar todos los factores que controlan las elecciones de un individuo puede ser una tarea complicada (Domjan, 2010c).

Por ello, los analistas de la conducta estudiamos los mecanismos de la elección utilizando situaciones experimentales lo más simples posible. En una típica situación de elección, a los organismos se les presentan dos alternativas de respuesta. Cada alternativa está asociada a un programa de reforzamiento, es decir, una regla que establece las condiciones para que la consecuencia sea entregada (Mazur, 2016), de esta manera, cuando la regla se cumple, el organismo puede recibir una recompensa -si la consecuencia es favorable-, o un evento desagradable -si la consecuencia es aversiva-.



Figura 1. Fotografía de una caja de condicionamiento operante para palomas dentro de un cubículo sonoamortiguador

Existen diversos programas de reforzamiento (Ferster & Skinner, 1957), a continuación se definirán sólo aquellos que serán utilizados en secciones posteriores del texto.

Entre los programas básicos que se utilizarán aquí se encuentran el programa de reforzamiento de razón fija (RF) y el programa de intervalo fijo (IF). El programa de RF consiste en que cada que se produce un número determinado de respuestas se accede a una consecuencia, por lo que el tiempo en que ésta será entregada depende de qué tan rápido sean emitidas las respuestas requeridas (Ferster & Skinner, 1957). Por otro lado, en el programa de IF, la primera respuesta después de un intervalo de tiempo designado es seguida de la consecuencia (Ferster & Skinner, 1957).

Además de los programas básicos, existen los programas complejos, en los cuales se combinan dos o más programas de reforzamiento. En particular, en el programa de reforzamiento encadenado, se dice que dos respuestas están “encadenadas”, si la respuesta que cumple con el requisito del primer programa produce un estímulo en presencia del cual la respuesta requisito del segundo programa será reforzada (Ferster & Skinner, 1957). Por ejemplo, un gimnasta debe

emitir una serie de respuestas (movimientos) antes de recibir un estímulo reforzante, como podría ser una calificación o aplausos, cada una de las respuestas que emite podrían ser iguales, por ejemplo, si tuviera que hacer dos “vueltas de carro”, o diferentes, si tuviera que hacer una “vuelta de carro” y finalizar con un *split*. En ambos casos la recompensa se obtiene hasta que finalizó la cadena/rutina y la ejecución del primer movimiento es el estímulo o señal para comenzar el segundo. Además, es común nombrar “eslabón” a cada una de las condiciones a satisfacer para completar la cadena (estas condiciones pueden ser la presentación de un estímulo o la realización una conducta), siendo la primer condición el eslabón inicial y el último -antes de la consecuencia- el eslabón terminal.

Otros programas complejos son los programas de reforzamiento concurrentes, en estos, a diferencia de los encadenados, se presentan dos o más programas simples de reforzamiento al mismo tiempo (Domjan, 2010c). De esta manera, resultan útiles para estudiar la conducta de elección, pues el participante puede elegir, entre las alternativas que se ofrecen, la que él prefiere. Por ejemplo, podríamos dar a elegir entre una alternativa que va a entregar una recompensa si se responde después de 10 segundos contra una donde se recompense cada vez que se responda. Los programas concurrentes pueden tener tantas alternativas como queramos y volverse tan complejos como necesitemos, por ejemplo, al incluir dos programas encadenados que se presentan de forma concurrente.

Elección subóptima

Como indica el título de esta tesis, el objetivo central de este trabajo es evaluar el efecto del valor incentivo de los estímulos (Robinson & Flagel, 2009) en el procedimiento de elección subóptima (Stagner & Zentall, 2010), por lo que esta sección está dedicada a describir el paradigma general, así como los antecedentes y aproximaciones que han guiado la investigación de este fenómeno, antes de detallar en qué consiste la atribución de valor incentivo.

Para ello, me tomaré la libertad de explicar el procedimiento clásico de la elección subóptima, en un análogo humano hipotético, antes de describirlo en términos técnicos.

Imagine que le dan a elegir entre dos cajas idénticas entre sí, una a su derecha, a la que nos referiremos como A, y otra a su izquierda a la que llamaremos B. Si elige la caja A, el 20% de las veces encontrará dentro una caja roja con dinero y el 80% de las veces una caja verde que estará vacía. Por el contrario, si elige la caja B un 20% de las veces encontrará una caja azul que la mitad de las veces contiene dinero y la mitad de las veces está vacía y un 80% de las veces, encontrará una caja amarilla que también la mitad de las veces contendrá dinero y la otra mitad no. ¿Cuál elegiría? (ver Figura 2).

Procedimiento de elección Subóptima

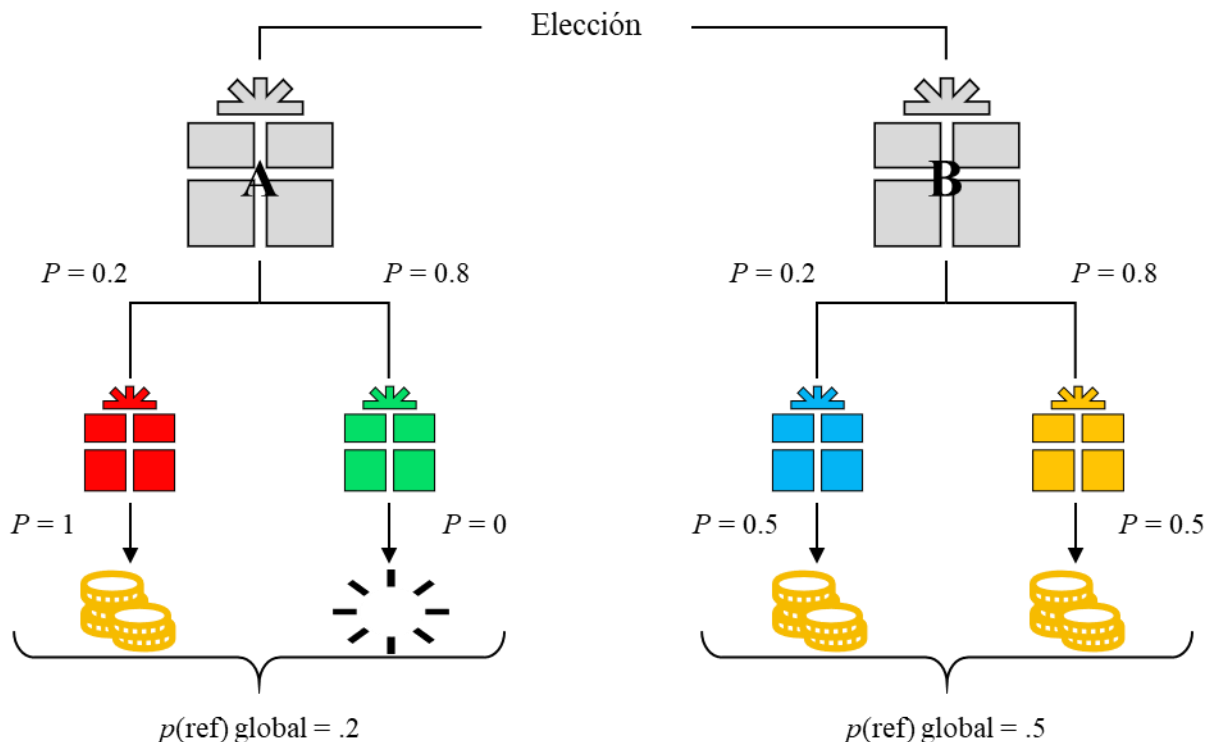


Figura 2. Análogo humano hipotético del procedimiento de elección subóptima de Stagner & Zentall (2010). $p(\text{ref})$ global representa la probabilidad de obtener un reforzador/recompensa según la alternativa elegida (Caja A o B).

Los primeros indicios del estudio del fenómeno de elección subóptima datan de 1974 cuando Stephen B. Kendall encontró que, al dar a elegir a palomas entre una alternativa que el 100% de las veces entregaba comida, contra otra que el 50% de las veces entregaba alimento de forma señalada (100% cuando aparece un estímulo (S+) y 0% cuando aparece un estímulo diferente (S-)), las palomas preferían la alternativa con menor tasa global de reforzamiento (Kendall, 1974). Estos hallazgos parecían contradecir la teoría de la utilidad esperada (Edwards, 1961), y el comportamiento óptimo en las palomas, predicho por dicha teoría (Shimp, 1969).

Desafortunadamente, los hallazgos de Kendall (1974) no pudieron ser replicados y se concluyó que fue un error en el procedimiento lo que dio lugar a esta aparente suboptimalidad (Fantino, Dunn & Meck, 1979). Kendall (1974) utilizó un programa concurrente encadenado poco común, en el eslabón inicial, las teclas que debían picar las palomas para elegir se encontraban oscuras y se iluminaban una vez que picaban alguna de ellas, mientras que la tecla no elegida permanecía apagada. De esta manera, cuando elegían la alternativa asociada a un .5 de probabilidad de reforzamiento y recibían el estímulo que señalaba reforzamiento (S+), se mantenían picando en esa tecla; pero si recibían el estímulo asociado con 0 probabilidad de reforzamiento (S-), picaban la tecla oscura adyacente (inactiva), por lo que la tasa efectiva de reforzamiento en esta última se veía reducida debido a que los animales la picaban y no recibían recompensa, asociando así que responder a esa tecla no siempre era seguido de reforzamiento (Gipson, Alessandri, Miller, & Zentall, 2009). En un intento por explicar los resultados de Kendall (1974), Fantino et al. (1979) realizaron una corrección a su experimento, reduciendo lo que ellos consideraron “ambigüedad del eslabón inicial”. Para eso, utilizaron teclas iluminadas de blanco en lugar de teclas oscuras, las cuales además cambiaban de color en los eslabones finales, y encontraron una preferencia por la alternativa asociada al 100% de reforzamiento, aunque esta

preferencia fue modesta ya que las palomas la elegían cerca del 60% de las veces, es decir, apenas por encima del azar (Fantino et al., 1979).

Años después, Gipson et al., (2009) en un primer experimento, cuestionaron si podrían replicar esta preferencia por la alternativa de mayor tasa de reforzamiento, para ello utilizaron dos grupos de palomas, uno que replicara las condiciones de ambigüedad de Kendall (1974) (grupo ambiguo), con una importante diferencia: en lugar de mantener ambas teclas oscuras en el eslabón inicial, esas teclas permanecían iluminadas. Mientras que el otro grupo replicaba las condiciones de reducción de ambigüedad de Fantino et al. (1979) (grupo sin ambigüedad) señalando las condiciones de reforzamiento de cada ensayo de la alternativa de baja tasa de reforzamiento (un estímulo indicaba cuando había reforzamiento y otro diferente cuando no). En el grupo sin ambigüedad, se observó elección óptima, es decir, los animales prefirieron la alternativa con mayor tasa de reforzamiento, mientras que en el grupo ambiguo no se observó una preferencia estadísticamente significativa para alguna de las opciones, así, aunque el procedimiento de Kendall (1974), pudo no ser acertado, su replicación mostró indicios de conducta subóptima. De esta manera, en un segundo experimento, Gipson et al., (2009) indagaron bajo qué condiciones podría encontrarse una preferencia estadísticamente confiable por la alternativa con menor tasa de reforzamiento.

Para ello, utilizaron un grupo basado en el grupo sin ambigüedad, reduciendo la tasa de reforzamiento de la alternativa de mayor densidad (de 100 a 75%) y manteniendo la tasa de reforzamiento de la alternativa de menor densidad (50%). Además, con el objetivo de descartar que la preferencia por la alternativa del 50% se debiera a que las palomas prefieren la “variabilidad en los estímulos” -ya que se demostró que la preferencia por la alternativa del 50% disminuía cuando la de 100% también contenía dos estímulos en lugar de uno (Mazur, 1996, Parte C)- decidieron incluir también dos estímulos durante el eslabón terminal de la alternativa de

mayor densidad de reforzamiento, ambos señalaban una probabilidad de .75 de reforzamiento (ver Gipson et al., 2009 Design of Experiment 2. [Figure 3]). Sorprendentemente, 13 de los 16 animales utilizados prefirieron de forma estadísticamente significativa la alternativa asociada a una menor tasa global de reforzamiento, es decir, eligieron de forma subóptima.

“Para nuestro conocimiento, esta es la primera evidencia de una preferencia estadísticamente significativa de un 50% de reforzamiento sobre un porcentaje mayor de reforzamiento en donde la demora del reforzamiento y la magnitud del reforzamiento han sido igualados” (Gipson et al., 2009, p.295) Los autores sugirieron que estos resultados se deben al contraste producido por la aparición del estímulo relacionado con 100% de probabilidad de reforzamiento en la alternativa con tasa global de reforzamiento del 50% (ganar, cuando se cree que se va a perder), el cual fue suficiente para sobreponerse a la diferencia de reforzamiento entre ambas alternativas (75% contra 50%).

Más adelante, en el 2010, Stagner y Zentall, condujeron un experimento que pretendía mostrar cómo la preferencia por la alternativa de baja densidad de reforzamiento del experimento de Gipson et al., (2009) podía incrementarse reduciendo la tasa de reforzamiento para esta alternativa, ya que se había demostrado que mientras más baja sea la probabilidad de reforzamiento (12.5%) más prefieren los animales tener estímulos discriminativos que indiquen cuándo habrá o no alimento, en comparación con condiciones donde la probabilidad de reforzamiento es alta (87.5%) (Roper & Zentall, 1999).

En la Figura 3 se presenta un diagrama del procedimiento diseñado por Stagner y Zentall (2010), que se ha convertido en el modelo convencional de elección subóptima. Como se puede observar, ésta es la versión técnica y real del procedimiento hipotético mostrado en la Figura 2, y a partir del cual se han derivado diversas modificaciones, con el propósito de determinar las variables que subyacen al fenómeno.

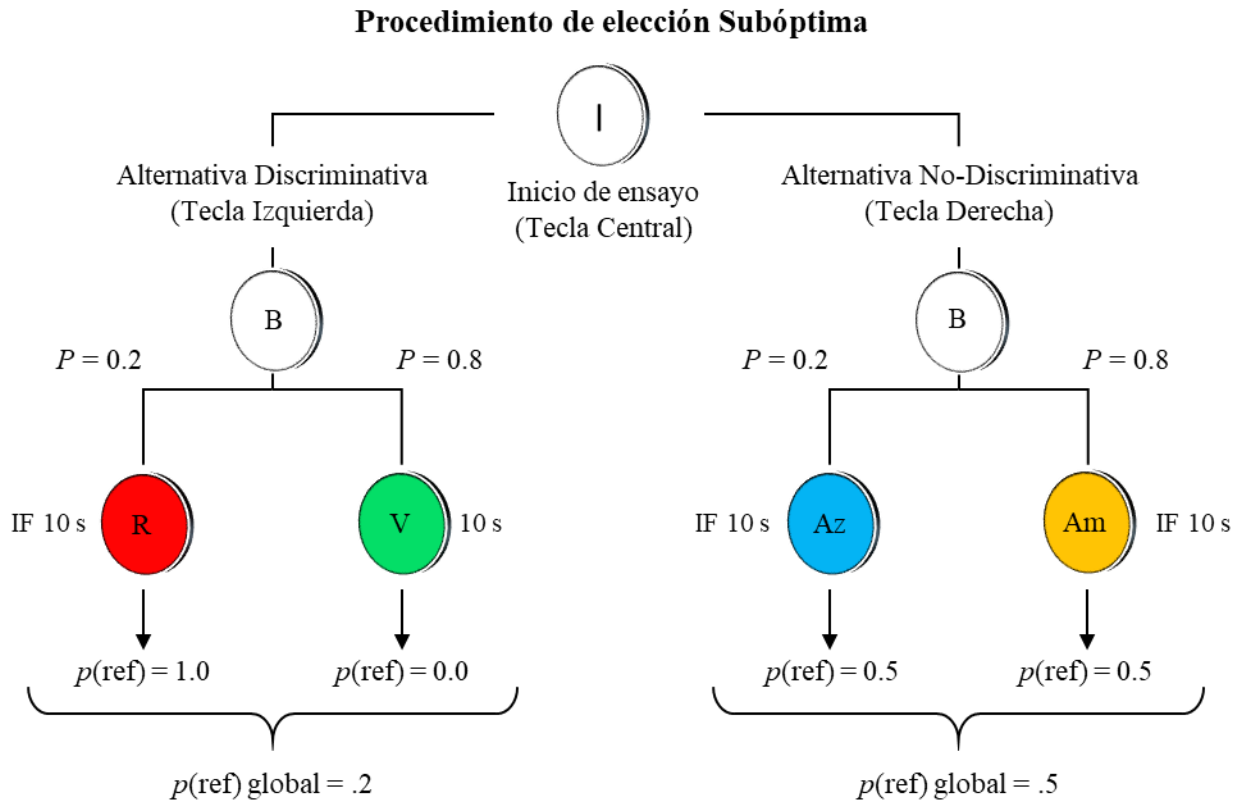


Figura 3. Procedimiento de elección subóptima de Stagner & Zentall (2010). Training [Figure 1 (A)]

Stagner y Zentall (2010), utilizaron un programa concurrente encadenado para evaluar la preferencia de los animales ante dos alternativas con diferente tasa de reforzamiento; cada ensayo comenzaba con una línea vertical proyectada en una tecla central, la cual una vez presionada podía presentar dos posibles alternativas. Estas alternativas eran señaladas mediante la iluminación de una tecla blanca (B), izquierda para la alternativa discriminativa y derecha para la alternativa no discriminativa. Cuando los sujetos elegían la alternativa discriminativa, esa tecla se iluminaba de uno de dos posibles colores, el 20% de las veces se iluminaba de color rojo (R) y la primera respuesta después de 10 segundos (IF10) apagaba el estímulo y entregaba reforzador, mientras que el 80% de las veces se iluminaba de color verde (V) y después de 10 segundos no entregaba reforzador. La tasa global de reforzamiento para esta alternativa fue de .2, es decir, de

optar por esta opción solo el 20% de las veces se obtendría recompensa. A esta alternativa se le llamó discriminativa ya que contenía información segura sobre la presencia o ausencia de reforzamiento en cada ensayo. Por el contrario, cuando los animales elegían la alternativa no discriminativa, la tecla se iluminaba de uno de dos posibles colores, el 20% de las veces de color azul (Az) y el 80% de las veces de color Amarillo (Am), en ambos casos, la primera respuesta después de 10 segundos apagaba el estímulo (IF10) y con un 50% de probabilidad entregaba reforzador. A esta alternativa se le llamó no discriminativa, debido a que sus estímulos no señalaban diferencialmente si se entregaría o no recompensa cuando finalizaran. La ubicación de las alternativas, así como los colores, fueron contrabalanceados entre los sujetos.

En una siguiente condición, para cerciorarse de que la preferencia de los animales no se debiera a un sesgo de posición, la posición de las alternativas fue invertida. Posteriormente para asegurar, que la alternativa era elegida por sus características y no por su posición, se sustituyó la tecla blanca en el eslabón inicial de las alternativas por un círculo para una de las alternativas y por un signo de más para la otra alternativa, de esta manera cualquiera de las alternativas podía aparecer tanto en la tecla derecha como izquierda. El resultado fue contundente, las palomas prefirieron la alternativa con menor recompensa, sin importar dónde se encontraba (Stagner & Zentall, 2010). Estos resultados plantearon un nuevo paradigma para la literatura, al establecer de forma confiable un fenómeno que parecía contradecir teorías como la del forrajeo óptimo (Pyke, 1984). Esta muestra de suboptimalidad, fue explicada como el resultado de una sobreestimación de la ocurrencia de un evento infrecuente que produce ganancias y una subestimación de un evento frecuente que produce pérdidas, sugiriendo que el contraste que genera ganar a pesar de su baja probabilidad de ocurrencia genera un efecto positivo en los animales, superior al efecto negativo generado por perder. Esta explicación dio pie a que se propusiera al fenómeno como un modelo de juego patológico, ya que, de forma análoga, en las situaciones de apuestas, las

personas están dispuestas a invertir en alternativas donde es muy poco probable ganar y aquello que puede mantenerlos eligiendo esas opciones resulta de lo reforzante de ganar en esas situaciones (Zentall, 2011).

Sin embargo, ganar no es el único elemento atractivo para participar en un juego de azar, sino cuánto se gana, es decir, la magnitud de la recompensa, así que en un experimento posterior se trató de emular con mayor cercanía el contexto de apuestas que experimentan las personas (Zentall & Stagner, 2011). Para ello, se dio a elegir a las palomas entre una alternativa asociada con una consecuencia segura, es decir, en todos los ensayos se entregarían 3 pellets -como lo sería conservar el dinero que uno tiene en el bolsillo- y una segunda alternativa asociada a una gran recompensa, es decir, 10 pellets, pero de baja probabilidad de ocurrencia. Para asegurarse de que esta conducta podía llegar a ser subóptima, se equilibraron las probabilidades de presentación de los estímulos y las probabilidades de recompensa, de tal manera que la alternativa de consecuencia segura proveyera en promedio más pellets que la alternativa de gran recompensa incierta.

De forma similar a los resultados observados en el experimento clásico de elección subóptima, las palomas volvieron a preferir la alternativa con menor tasa global de reforzamiento, de esta manera, los autores sugirieron que este comportamiento era parecido al que mostraría una persona con conducta de juego patológico (posteriormente demostrado con personas por Molet et al., (2012)), por lo que es posible que en ambos casos el comportamiento estuviera mediado por los mismos mecanismos (Zentall & Stagner, 2011). Sin embargo, aunque la investigación en este último protocolo, que trata de emular con mayor cercanía el contexto de apuestas no es tan abundante como en la versión prototípica de elección subóptima, ambos casos suponen que las causas subyacentes al fenómeno en palomas podrían aportar información sobre esta forma de conducta maladaptativa en humanos, o viceversa, ya que, al ser considerado como un análogo de

la conducta de juego patológico en humanos, estudiar las variables que influyen en la conducta de juego patológico de las personas, podría ayudar a explicar los resultados encontrados en palomas (Zentall, 2019). Por ejemplo, se ha identificado el ingreso económico de las personas como una de las variables demográficas involucradas en el juego patológico, así, las personas que tienen menos recursos suelen incurrir más en estas conductas que aquellas con más recursos (Zentall, 2019), por lo que en un intento de hacer esta situación análoga para las palomas, se controló el nivel de privación de alimento al que estaban sometidos los animales (Laude, Pattison, & Zentall, 2012), esperando que de forma similar a las personas, aquellas menos privadas (con más recursos), mostraran menos elección subóptima que aquellas más privadas (con menos recursos). En efecto, se encontró que las palomas más privadas eran más propensas a elegir la alternativa subóptima, que aquellas menos privadas, aunque la preferencia de estas últimas comenzó a revertirse conforme avanzaban las sesiones (Laude et al., 2012).

Otra de las variables que se creyó podría revertir el comportamiento subóptimo en las palomas fue el enriquecimiento ambiental, ya que al estudiarse sus efectos, se habían demostrado beneficios en el aprendizaje y, sobre todo, que puede actuar como factor de protección contra el desarrollo de conductas maladaptativas (Stairs & Bardo, 2009). De esta manera, Pattison et al. (2013) realizaron un experimento en el que variaron la presencia de enriquecimiento ambiental al que fueron expuestos dos grupos de animales antes de la tarea, aunque ambos grupos fueron alojados de forma individual en cajas de 30.5 x 28 x 38 cm (alto, largo, ancho) solo al grupo de animales que recibieron enriquecimiento ambiental se les permitió acceder, 4hrs al día durante 5 días a la semana a una “caja de vuelo” de 2.13x 1.22 x 1.84m (alto, largo, ancho), la cual permitía la interacción social con las demás palomas además de proveer un espacio más grande respecto al que estaba restringidas el resto del tiempo. De esta manera Pattison et al. (2013) encontraron que, en efecto, un mayor enriquecimiento ambiental podía “proteger” a los animales de presentar

conducta subóptima, pero sólo durante las primeras sesiones, ya que posteriormente el efecto protector parece desaparecer y la suboptimalidad alcanza niveles similares a los observados en palomas sin enriquecimiento ambiental.

En síntesis, las explicaciones más recurrentes y aceptadas del fenómeno se pueden dividir en dos, que no son excluyentes: la primera, ya antes mencionada, explica la elección subóptima con base en el contraste que genera ganar en un ambiente donde es muy probable perder (Gipson et al., 2009; Stagner & Zentall, 2010; Zentall, Andrews, & Case, 2019) y la segunda, lo explica como una subestimación de la ocurrencia de eventos de pérdidas y sobre valoración de la ocurrencia de eventos de ganancias, sin embargo, ambos casos suponen primero: un pobre control del estímulo predictor de no reforzador (S-) (Laude, Stagner, & Zentall, 2014; Zentall, 2011), y segundo: un control exacerbado del estímulo predictor de reforzador (S+) (Zentall, Laude, Stagner, & Smith, 2015).

En un principio, se sugirió que el bajo control del estímulo inhibitorio (S-) podía deberse a que los animales lo ignoraban, por lo que en un intento de volverlo inescapable este fue presentado de forma ambiental (luz general), en lugar de localizada (tecla) (Stagner, Laude, & Zentall, 2011), sin embargo, esta manipulación no mostró un efecto en la elección de los animales, quienes siguieron eligiendo de forma subóptima. Años después se demostró que aunque el S- sí ejerce un control inhibitorio en las palomas en este procedimiento, este efecto sólo ocurre durante las primeras sesiones y eventualmente desaparece (Laude et al., 2014).

El estudio de esta hipótesis centrada en el deficiente control que ejerce el estímulo inhibitorio, incluso ha inspirado algunos de los modelos que pretenden explicar la elección subóptima, como el modelo ecológico de elección subóptima (Vasconcelos, Monteiro, & Kacelnik, 2015), la SiGN hipótesis (McDevitt, Dunn, Spetch, & Ludvig, 2016), el ADM

(Associability Decay Model) (Daniels & Sanabria, 2018) y el modelo teórico de la información temporal (Cunningham & Shahan, 2018), por mencionar algunos.

Aunque el fenómeno de la elección subóptima había sido ampliamente abordado en palomas, no se había realizado investigación en otras especies (a excepción de humanos con tendencia al juego patológico [Molet et al., 2012]), por lo que en busca de la generalidad de este fenómeno se realizó una adaptación al procedimiento en ratas, sorprendentemente, las ratas mostraron conducta óptima, poniendo en duda la generalidad del procedimiento entre especies (Trujano & Orduña, 2015). En dicho experimento, se sustituyeron las teclas de colores del procedimiento original por luces emitidas a partir de un triple estimulador (tres LED, de diferentes colores, adheridos a una placa metálica) y para que las ratas pudieran elegir entre alternativas se utilizaron palancas.

Sin embargo, este cuestionamiento a la generalidad duró poco, ya que en el 2017, Chow et al., pudieron demostrar que las ratas, como las palomas, también pueden elegir de forma subóptima. Además, con esta demostración se añadió una tercera explicación a la tarea de la elección subóptima: la saliencia incentiva. Chow et al. (2017) usaron palancas como S+, argumentando que estos estímulos poseen mayor valor incentivo que las luces utilizadas por Trujano y Orduña (2015), con lo cual se explicó el fallo que ellos tuvieron en encontrar elección subóptima. Aunque en investigaciones posteriores trató de debatirse esta idea (López, Alba, & Orduña, 2018; Martínez, Alba, Rodríguez, & Orduña, 2017), la saliencia incentiva ha demostrado propiciar la elección subóptima, aunque de forma diferencial dentro los animales de una misma especie, ya que es capaz de influir en la elección de los animales catalogados como seguidores de señales, pero no en los seguidores de objetivos (Orduña & Alba, 2020).

Como se ha mencionado hasta ahora, múltiples hipótesis fueron descartadas en busca de respuestas para explicar el fenómeno de la elección subóptima, dejándonos con 3 posibles

explicaciones: el contraste, la inhibición y ahora, la saliencia incentiva. La primera se ha limitado a explicar los hallazgos con palomas; la segunda, podría explicar las diferencias entre especies como una diferencia en la sensibilidad a la inhibición entre ratas y palomas, pero es incapaz de explicar por qué es posible tener tanto elección óptima como subóptima en la misma especie; mientras que la tercera ha sido capaz de explicar ambas formas de elección en ratas, por lo que la investigación en torno a esta última, podría guiarnos hacia la explicación del fenómeno de la elección subóptima, suponiendo un mecanismo común entre especies.

La atribución de saliencia incentiva a los estímulos

Antes de definir la saliencia incentiva, es necesario definir un mecanismo en el que está basada: el condicionamiento clásico o Pavloviano. El condicionamiento clásico es una forma de aprendizaje en la cual un estímulo neutro (EN) se presenta repetidas veces antes/al mismo tiempo/seguido de un estímulo incondicional (EI) es decir, cualquier estímulo que, sin necesidad de entrenamiento, es capaz de provocar una respuesta refleja, llamada respuesta incondicional (RI). Estas repeticiones terminarán por lograr que el estímulo neutro, ahora llamado estímulo condicionado (EC), provoque la respuesta incondicionada que antes era provocada por el EI, ahora llamada respuesta condicionada (RC).

Por ejemplo, la comida es un EI, que al entrar en contacto con nuestra lengua produce de forma incondicionada (RI) salivación, sin que esta respuesta tuviera que ser entrenada; simplemente surge como un reflejo. Ahora, supongamos que tuviéramos que alimentar perros y antes de hacerlo siempre hacemos sonar una campana, de esta forma, con el paso del tiempo, veremos que la campana que antes no evocaba una respuesta en el animal se vuelve una señal de alimento y provoca la respuesta de salivación que antes sólo ocurría en presencia de la comida. Entonces decimos que la campana, antes EN, es ahora un EC, que evoca una respuesta condicionada (RC), y además de provocar esa respuesta sirve como señal certera sobre la

presentación del EI, por eso también se le puede llamar estímulo excitatorio (S+). Lo que acabo de describir de manera parafraseada y extremadamente resumida representa algunos de los hallazgos de Iván Pavlov (citado en ([Domjan, 2010a]), aunque sus investigaciones sugirieron que el condicionamiento clásico sólo ocurre en respuestas reflejas, una nueva concepción incluye formas de conducta más complejas, como lo es el paradigma de seguimiento de señales, también llamado automoldeamiento (Hollis, 1997; Rescorla, 1988).

Las señales, son una parte importante en nuestras vidas, tener una señal que nos indique cuando algo ocurrirá o no, nos permite tomar mejores decisiones, nos aleja de eventos potencialmente peligrosos pero también nos pueden guiar hacia decisiones maladaptativas, por ejemplo, haciéndonos comer cuando en realidad no tenemos hambre (Colagiuri & Lovibond, 2015), provocar recaídas en alguna adicción (Crombag, Bossert, Koya, & Shaham, 2008), incluso, pueden evocar conductas en animales que parecen tan extrañas como preferir picotear un trozo de madera que predice una pareja sexual en lugar de acercarse a la ubicación de la pareja (Burns & Domjan, 1996, 2000), y en el caso de las palomas, estas señales podrían llevarlas a elegir un alternativa por los estímulos que contiene en lugar de las ganancias asociadas a ella (Stagner & Zentall, 2010). La información es valiosa para nosotros, esta nos permite tomar decisiones, sean acertadas o no.

Cuando un estímulo señala confiablemente la entrega de un EI, por ejemplo comida, los animales usualmente responden ante este estímulo de dos formas, se acercan e interactúan con él o se dirigen a la zona donde se entregará el alimento; los primeros son llamados seguidores de señales, y los segundos, seguidores de objetivos (Flagel, Akil, & Robinson, 2009). Los primeros hallazgos del comportamiento de seguimiento de señales en laboratorio vienen de los experimentos de Brown y Jenkins (1968), en un experimento con palomas presentaron una tecla iluminada durante algunos segundos, que al apagarse era seguida de comida y observaron que los

animales comenzaron a picarla, aun cuando sus picotazos no eran necesarios para la obtención de alimento. Este comportamiento también es conocido como automoldeamiento o aproximación Pavloviana condicionada y es una forma de condicionamiento clásico donde la respuesta condicionada (RC) es una respuesta involuntaria (reflejo adquirido) que es provocada de forma automática por la presentación de un estímulo condicionado (EC) (Tomie, 2018) y en ocasiones esta conducta, continua ocurriendo aún si tiene como consecuencia la pérdida de la recompensa (Williams & Williams, 1969).

A partir del descubrimiento de Brown y Jenkins (1968) se han realizado múltiples experimentos sobre el seguimiento de señales o automoldeamiento en diversas especies, como pollos, codornices, peces de colores, lagartijas, ratas, monos Rhesus y monos ardilla, así como en adultos y niños humanos (Tomie, Brooks & Zito, 1989 citado en Domjan, 2010b). Algunos investigadores proponen el seguimiento de señales como un modelo para estudiar conductas maladaptativas, y destacan su papel motivacional en la adicción a las drogas, nombrando valor incentivo al “deseo” exacerbado que provocan esas señales (Robinson & Flagel, 2009).

Los estímulos con valor incentivo/saliencia incentiva no son sino una forma de EC que puede volverse un poderoso controlador de la conducta, y potencialmente promover conductas maladaptativas (Meyer et al., 2012). Aunque estos estímulos sólo han sido caracterizados en ratas, a partir de una recopilación de evidencia indirecta, podemos hipotetizar que las teclas del procedimiento convencional de elección subóptima en palomas (Stagner & Zentall, 2010) poseen estas propiedades. Los estímulos con saliencia incentiva o estímulos incentivos Pavlovianos cuentan con 3 características (Meyer et al., 2012):

1) Sesgan la atención hacia ellos. Provocan seguimiento de señales

Provocan que los individuos se les aproximen, es decir, elicitan conducta de seguimiento de señales. Esta característica se cumple si el estímulo provoca conducta de seguimiento de señales en un procedimiento de automoldeamiento (Everitt, Dickinson, & Robbins, 2001; Meyer et al., 2012; Robinson, Carr, & Kawa, 2018).

En los experimentos de automoldeamiento observamos cómo las palomas, sin necesidad de entrenarlas a picar, presionan una tecla cuando su iluminación siempre es seguida de alimento, aun cuando la entrega del reforzador no depende de su respuesta (Brown & Jenkins, 1968), la ocurrencia de conducta de seguimiento de señales, es clave en la atribución de saliencia incentiva ya que son solo los seguidores de señales quienes realizan esta atribución (Flagel et al., 2009).

2) Son deseados en sí mismos. Funcionan como reforzadores condicionados

Los individuos trabajan para obtener el EC, incluso ante la ausencia de un EI como recompensa, y pueden incluso servir como reforzadores para el aprendizaje de nuevas conductas. Está característica se cumple si el EC puede funcionar como reforzador condicionado (Everitt et al., 2001; Meyer et al., 2012; Robinson et al., 2018).

En el experimento de respuestas de observación de Wyckoff Jr., (1952), una paloma podía pararse sobre un pedal, que sólo tenía como consecuencia la presentación de una tecla iluminada que había sido asociada con comida, señalando el tipo de componente en que se encontraba el animal, es decir convertía un programa de reforzamiento mixto en múltiple. Esta respuesta de pararse sobre el pedal no fue entrenada por el investigador, sino que surgió a partir del efecto que ocasionaba, en este caso, provocaba la presentación de la tecla iluminada que anteriormente había sido asociada con comida, funcionando ahora como reforzador condicionado, esta tarea es muy similar a la descrita por Robinson y Flagel (2009) quienes demostraron que un estímulo con alto valor incentivo, en su caso una palanca, podía reforzar nuevas conductas en ratas.

3) Pueden generar estados motivacionales condicionados. Potencian acciones instrumentales

Pueden instigar o vigorizar la búsqueda de la recompensa asociada y potenciar las acciones instrumentales como en el efecto de transferencia Pavloviana-instrumental (PIT en inglés). En el caso de señales asociadas con drogas, estos estímulos pueden producir craving o recaídas. Esta característica se cumple si el estímulo provoca PIT (Everitt et al., 2001; Meyer et al., 2012; Robinson et al., 2018). En el procedimiento de PIT se parea un EC con un EI, formando una asociación Pavloviana, posteriormente el animal se entrena en un procedimiento operante donde requiere realizar una respuesta a cierta manipulando (aparato en el que se tiene que responder) para obtener un EI, por último, se demuestra que hubo una transferencia si al presentarse el EC de la primera fase incrementa las respuestas a la manipulanda de la segunda fase (Cartoni, Balleine, & Baldassarre, 2016).

Lolordo (1971) demostró que las palomas incrementaban su tasa de respuestas en una tecla en la que al responder se les entregaba comida, cuando ésta se iluminaba de un color que anteriormente había sido empleado como un estímulo predictor de reforzamiento independiente de respuesta. Aunque en ese momento, este procedimiento no fue definido como PIT, cumple con las características del proceso.

Esta evidencia, en conjunto, permite suponer que las teclas iluminadas predictoras de reforzamiento pueden ser potencialmente atribuidas con saliencia incentiva por las palomas, incluso, el control que pueden llegar a ejercer estos estímulos es tal, que las palomas pueden continuar respondiendo ante ellos aun cuando esto implique la cancelación de entrega de alimento (Williams & Williams, 1969).

Las operantes y su relación con la atribución de saliencia incentiva

Como se mencionó en la sección anterior, Brown y Jenkins (1968) encontraron que la presentación repetida de una tecla iluminada seguida de acceso a comida resultaba en picotazos a la tecla aunque estas respuestas no hubieran sido entrenadas, sin embargo la topografía presente en la conducta de seguimiento de señales, es decir, los picotazos, no son una casualidad.

Las características de los estímulos excitatorios o predictores de la presentación de un estímulo incondicionado (S+) evocan y mantienen un subconjunto particular de las respuestas comúnmente relacionadas con esa recompensa (Timberlake, 1993). Dependiendo de las características del estímulo incondicionado (EI) diferentes clases de respuesta pueden ser activadas (Bertram O Ploog, 2014; Timberlake & Grant, 1975). Si le presentamos a las palomas un estímulo de forma constante, por ejemplo una tecla iluminada, cuya presentación es seguida de alimento, es muy probable que las palomas presenten, ante la tecla, una respuesta similar a la conducta consumatoria de la especie, es decir, a la forma en la que consumirían el EI, como se observa en el procedimiento de automoldeamiento (Brown & Jenkins, 1968); de forma similar, si la recompensa fuera agua, la topografía de respuesta con la que responderían al S+ se asemejaría a la de beber (Jenkins & Moore, 1973). Sin embargo, la respuesta también requiere ser compatible con el S+, de forma que, aunque el reforzador sea comida, si el S+ fuera otro animal, no se activarán los sistemas de conducta asociados a la conducta de comer (Timberlake & Grant, 1975).

Así, un estímulo que provoca seguimiento de señales, está reflejando su habilidad para elicitare respuestas consumatorias apropiadas, las cuales estarían relacionadas con la atribución de saliencia incentiva (Zentall, Smith, & Beckmann, 2019). Dicho esto, es posible que un estímulo que no genera respuestas consumatorias apropiadas no adquiera saliencia incentiva y por lo tanto no genere conducta subóptima.

El presente experimento supone que, si anulamos esta tendencia biológica al seguimiento de señales provocada por los estímulos utilizados típicamente en elección subóptima, podríamos prevenir la atribución de saliencia incentiva, y en consecuencia, evitar que los animales sesguen su preferencia hacia la alternativa con estos estímulos, propiciando una elección óptima.

Por lo anterior, si la saliencia incentiva es un factor involucrado en la suboptimalidad en las palomas, es necesario realizar ajustes al paradigma convencional (Stagner & Zentall, 2010). De esta manera, en lugar de utilizar diferentes colores en las teclas que parecen poseer un alto valor incentivo para las palomas, se utilizarán luces ambientales, ya que se ha demostrado que el seguimiento de señales sólo ocurre en situaciones en las que el S+ es localizado, mientras que el uso de estímulos ambientales, termina propiciando la aparición de conducta de seguimiento de objetivos (Cleland & Davey, 1983; Meyer, Cogan, & Robinson, 2014).

Además de reducir el valor incentivo de los estímulos, también es importante evitar que los animales puedan realizar conducta de seguimiento de señales hacia otro estímulo y adjudicarle saliencia incentiva, por lo que es necesario un aparato en el cual puedan responder para elegir entre las alternativas sin usar teclas y sin que la respuesta operante sea picar, así que se utilizarán palancas que los animales deberán pisar para emitir su respuesta, ya que de otra forma podría activarse su respuesta natural de picoteo.

Experimento 1. Condición de bajo valor incentivo

Método

Sujetos

Se utilizaron 14 palomas híbridas, experimentalmente ingenuas con edad entre 1 y 2 años y de sexo indistinto, las cuales fueron privadas de alimento, hasta alcanzar y mantener el 80% de su peso *ad libitum*. Se les mantuvo en jaulas individuales, con ciclos de iluminación de 12 horas luz, 12 horas oscuridad, a una temperatura aproximada de 22°C con acceso libre a agua, y en caso de requerirlo, alimento después de la sesión para mantener su peso. El manejo de los animales y los protocolos experimentales empleados estuvieron apegados a la Norma Oficial Mexicana NOM-123 062-ZOO-1999 sobre especificaciones técnicas para la producción, cuidado y uso de los animales de laboratorio.

Aparatos

El experimento se llevó a cabo en 2 cajas modulares de condicionamiento operante rectangulares (MED Associates, Inc., Model ENV-008) (ver Figura 4), con un área de trabajo de 29.53 cm (largo) x 23.5 cm (ancho) x 27.31 cm (alto). Cada caja tenía 2 paneles operativos idénticos (izquierdo y derecho) de 3 canales cada uno y el piso de la caja estaba compuesto por 19 barras de acero inoxidable de 0.48 cm de diámetro. En la parte inferior del canal central a 1.3 cm del piso se ubicaba una abertura cuadrada de 5.5 cm que proveía acceso a un comedero para palomas (MED Associates, Inc., Model ENV-205M) que contenía una mezcla de granos. El comedero era iluminado por una lámpara de 28V, 0.04A cada vez que se activaba. En el canal derecho, a 2cm del piso, se colocó una palanca retráctil (MED Associates, Inc., Model ENV-112CM) que se extendía 4.8 centímetros dentro de la caja y en la parte superior del canal a 24.5cm de la palanca se colocó una luz general con canastilla (MED Associates, Inc., Model

ENV-215M). Finalmente, se colocaron en el techo 4 tiras de luces LED ultrabrillantes de 5mm; cada una compuesta por 5 LED del mismo color (rojo, verde, azul y blanco) separados 4 cm entre ellos. La Figura 4 muestra un modelo tridimensional del equipo utilizado.

Las cajas estaban dentro de cubículos sonoamortiguadores fabricados de Multi-Density Fiberboard (MDF) (MED Associates, Inc., Model ENV-022MD) con un ventilador acoplado para atenuar sonidos externos. Todo el equipo estaba conectado a una interfaz de 28-V (MED Associates, Inc., Model SG-6510D), permitiendo que la presentación de estímulos, así como la recolección de datos fueran controlados por computadora usando programación en lenguaje Medstate (Med-PC IV, MED Associates, Inc.).

Adicionalmente, las sesiones experimentales fueron monitoreadas mediante webcams (Logitech C170) a través de una perforación circular de 5cm de diámetro detrás de los canales centrales a 17 cm del piso.

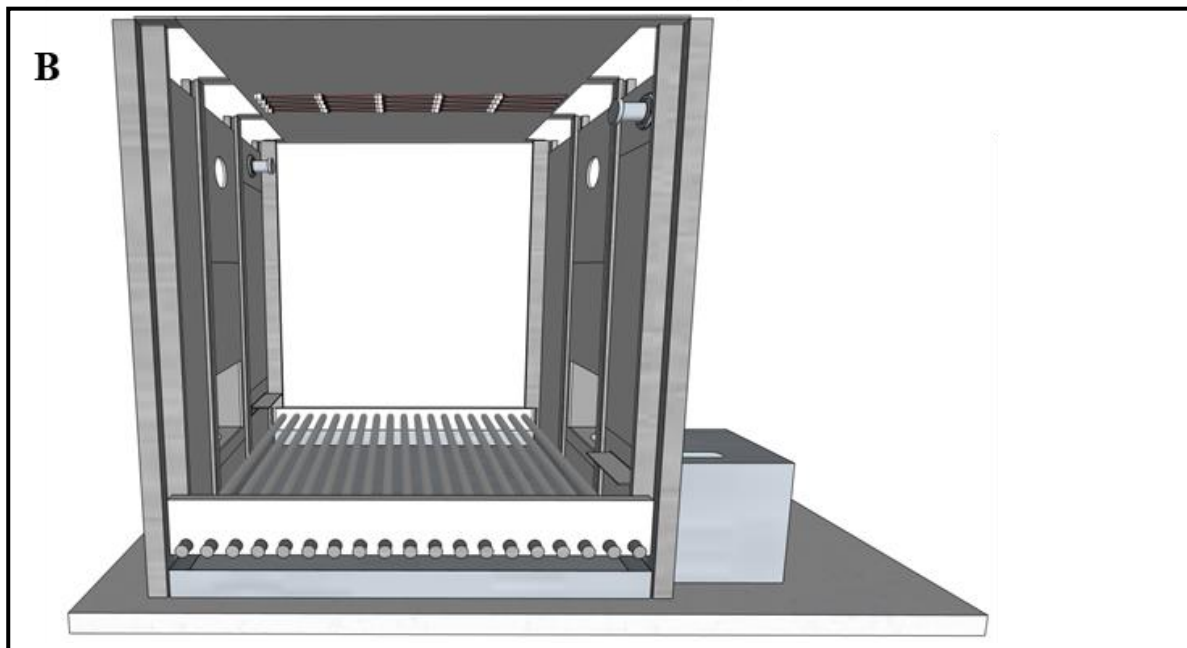
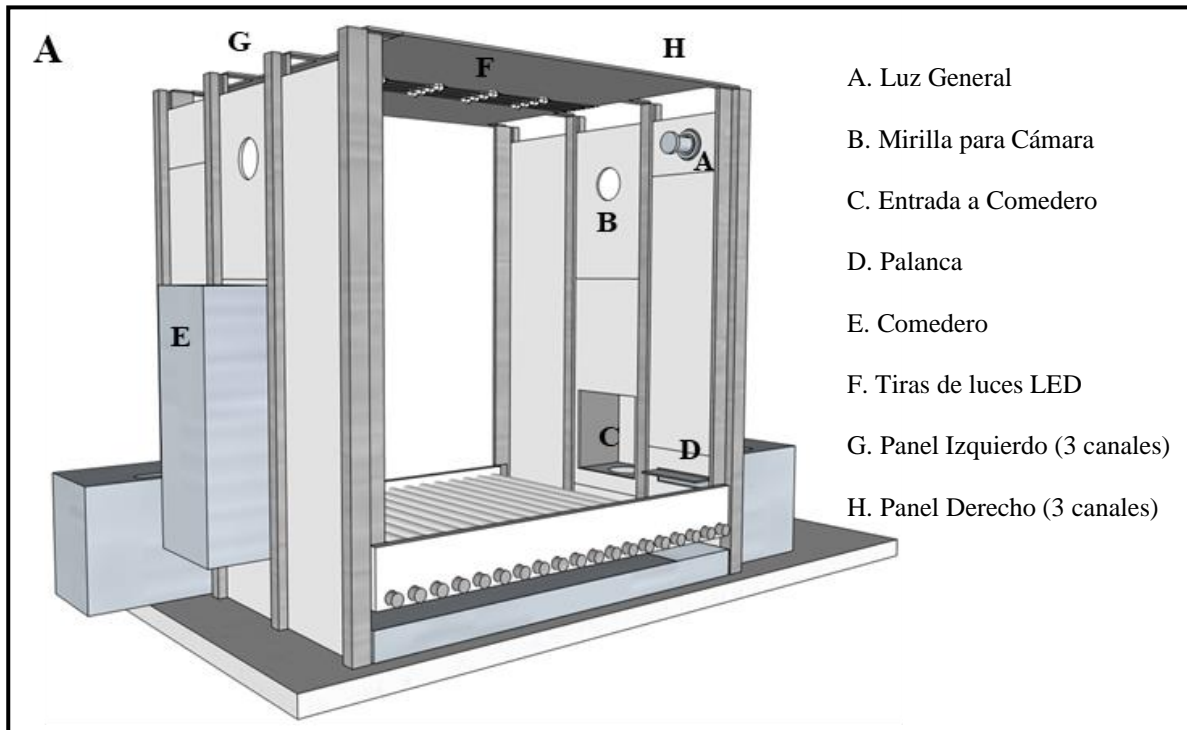


Figura 4. Modelo 3D realizado en SketchUp de las cajas operantes utilizadas en el Experimento 1. El panel A señala y enumera los componentes utilizados mientras que el panel B muestra la caja desde la perspectiva del experimentador, con todos sus componentes

Procedimiento

Habitación

Cuando los sujetos alcanzaron el 80% de su peso *ad libitum*, se realizó una sola sesión de habituación de 15 minutos, durante la cual, los animales fueron introducidos a la caja de condicionamiento operante donde se realizarían todas las demás fases del experimento para que así se familiarizaran con su entorno. Para ello, se colocaron algunos granos de alimento alrededor de cada comedero y se realizó una observación remota del interior de la caja por medio de cámaras web. Al inicio de la sesión se encendieron ambas luces generales para que el animal pudiera ver el alimento y después de 3 segundos de haberse acercado a alguno de los comederos, se apagaron ambas luces y encendió el comedero más cercano al animal, después de 5 segundos de que el animal realizara movimientos introductorios a éste, se apagó y encendieron las luces generales nuevamente hasta que el animal se orientara hacia el comedero restante y cumpliera el mismo criterio. Una vez que el animal se había dado una respuesta en cada comedero, comenzó un programa automatizado de IF 20s durante el cual se presentó de forma cuasi aleatoria la luz general de uno de los paneles durante 20 s y al apagarse se activó el comedero del mismo panel durante 3 s (el panel operativo no podía ser el mismo más de 3 veces seguidas).

Moldeamiento de presión a la palanca

Se moldeó la respuesta de pisar una palanca manualmente mediante el método de aproximaciones sucesivas (Skinner, 1951). Para reforzar las conductas similares a pisar una palanca que era la conducta objetivo, se permitió el acceso al alimento del comedero más cercano durante aproximadamente 2 segundos, cada que el animal realizara una conducta parecida a la de pisar. Primero, se reforzó cuando el animal adoptó una posición erguida, posteriormente la orientación hacia alguna de las palancas, después las aproximaciones a la palanca con movimientos de péndulo (hacia atrás y hacia adelante) y por último, la entrega del reforzador fue

contingente a las presiones a la palanca. Cada sesión de moldeamiento duró 25 minutos y fueron requeridas de 1 a 5 sesiones para cada animal.

Entrenamiento en presión a la palanca

Con el propósito de mantener la conducta de presión a la palanca, fue necesario el entrenamiento de ésta para reducir su latencia de respuesta, es decir, el tiempo que el animal tardaba en pisar la palanca, desde que ésta se presentaba, debía volverse más corto.

Para ello, se expuso a las palomas a un programa de reforzamiento RF1 durante el cual se presentaban 60 ensayos, divididos en 4 bloques de 15 ensayos, durante los ensayos de cada bloque, se presentaba exclusivamente la palanca de uno de los paneles con su respectiva luz general. Para que una respuesta fuera considerada efectiva debía tener una duración mayor a 0.2 s para así evitar que los animales respondieran picando las palancas en lugar de pisarlas, este criterio se mantuvo en las siguientes fases del experimento. Cada presión exitosa iba seguida de 2 s de acceso al comedero del mismo panel en el que se encontraba la palanca presionada y posteriormente comenzaba un intervalo entre ensayos (IEE) de 8 s, durante el cual todos los estímulos permanecían apagados. Los bloques fueron presentados de forma alternada. El criterio para pasar a la siguiente fase fue que los animales presentaran una mediana de latencias menor a 5 s y que la diferencia entre las medianas de cada palanca fuera menor a 1. Esta fase duro de 1 a 5 sesiones. Se tomaron en cuenta las medianas, para evitar que los valores de latencia fueran acarreados por valores extremos comúnmente presentados durante los primeros ensayos.

Pre-exposición a los estímulos

Durante una sola sesión los sujetos fueron pre-expuestos estímulos que se utilizarían en la siguiente fase. Cada ensayo comenzaba con la presentación de la palanca de alguno de los paneles asociada a un programa de RF1 y su respectiva luz general como eslabón inicial, si el animal presionaba la palanca, avanzaba al eslabón terminal, durante el cual la luz general se

apagaba y se encendían una de las 4 tiras de luces ambientales instaladas en el techo de la caja (roja, verde, azul, blanca) asociadas a un programa de reforzamiento RF 3 después del cual las luces junto con la palanca se desactivaban y se activaba el comedero durante 3 s. Cada color se presentaba 5 veces para cada una de las palancas, siendo un total de 40 ensayos durante la sesión.

Entrenamiento. Procedimiento de elección subóptima

En esta fase, se usó un programa encadenado de dos eslabones que presentaba dos tipos de ensayos: forzados y de elección (ver Figura 5). Durante el eslabón inicial de los ensayos forzados, sólo una palanca y su respectiva luz general se encendían en alguno de los paneles, mientras el otro permanecía oscuro y la palanca inactiva. Una respuesta en la palanca activa, que durara más de 0.2 segundos, permitía la transición al eslabón terminal, durante el cual se presentaba un estímulo de iluminación ambiental por 10s. Si la palanca correspondía a la alternativa discriminativa, en el 20% de los ensayos, la cámara experimental era iluminada de color rojo, durante un IF 10 s que funcionaría como estímulo excitatorio (S+) al señalar confiablemente la presentación del reforzador (3 segundos de acceso al comedero); por otro lado, en el 80% de los ensayos, una presión a esa misma palanca iluminaba la cámara de color verde, el cual finalizaba después de 10 segundos sin que fuera necesaria una respuesta, funcionando como estímulo inhibitorio (S-), o predictor de la ausencia del EI. En esta alternativa la tasa global de reforzamiento era del 20%.

Cuando la palanca correspondía a la alternativa no discriminativa, en el 20% de los ensayos, una pisada a la palanca iluminaba la cámara, de color azul (S1), durante un IF 10 s que el 50% de las veces era seguido por reforzamiento; y en el 80% de los ensayos en esa alternativa, una presión de la palanca iluminaba la cámara de color blanco (S2), durante un IF 10 s que el 50% de las veces fue seguido por reforzamiento. La tasa global de reforzamiento en esta alternativa era del 50%. Cada sesión constaba de 60 ensayos que fueron divididos en 20 ensayos

forzados para cada alternativa y 20 ensayos de elección, todos presentados de forma cuasi aleatoria (no se podía presentar más de 3 veces seguidas el mismo tipo de ensayo ni más de 3 veces seguidas la misma alternativa en el caso de los ensayos forzados). Los ensayos fueron separados por un IEE de 10 s, durante el cual todos los estímulos permanecían inactivos. En los ensayos libres ambas luces generales y ambas palancas se activaban hasta que el animal respondía en alguna de ellas, la opción no elegida fue inactivada, retrayendo la palanca e inactivando la respectiva luz general. El número de sesiones requeridas para dar por concluida esta fase, fue diferente para cada sujeto, los sujetos debían haber completado por lo menos diez sesiones y no mostrar alguna tendencia a incrementar o disminuir su proporción de elección por la alternativa discriminativa en las últimas 5 sesiones.

La posición de la alternativa discriminativa, así como los colores asociados al S+ y al S- fueron contrabalanceados entre los sujetos.

Entrenamiento en reversión

Durante esta fase, con el objetivo de verificar que la elección se debiera a las contingencias asociadas a cada alternativa y descartar un posible sesgo de posición, la localización de la alternativa discriminativa y no discriminativa fue invertida. El criterio de finalización fue el mismo que en la fase anterior. En caso de que la preferencia obtenida fuera inconsistente con la fase anterior, la ubicación de las alternativas era revertida a su posición durante el entrenamiento y el animal debía continuar bajo los mismos criterios de esa fase, para identificar si su elección se debía a un sesgo o a la adquisición de una nueva preferencia.

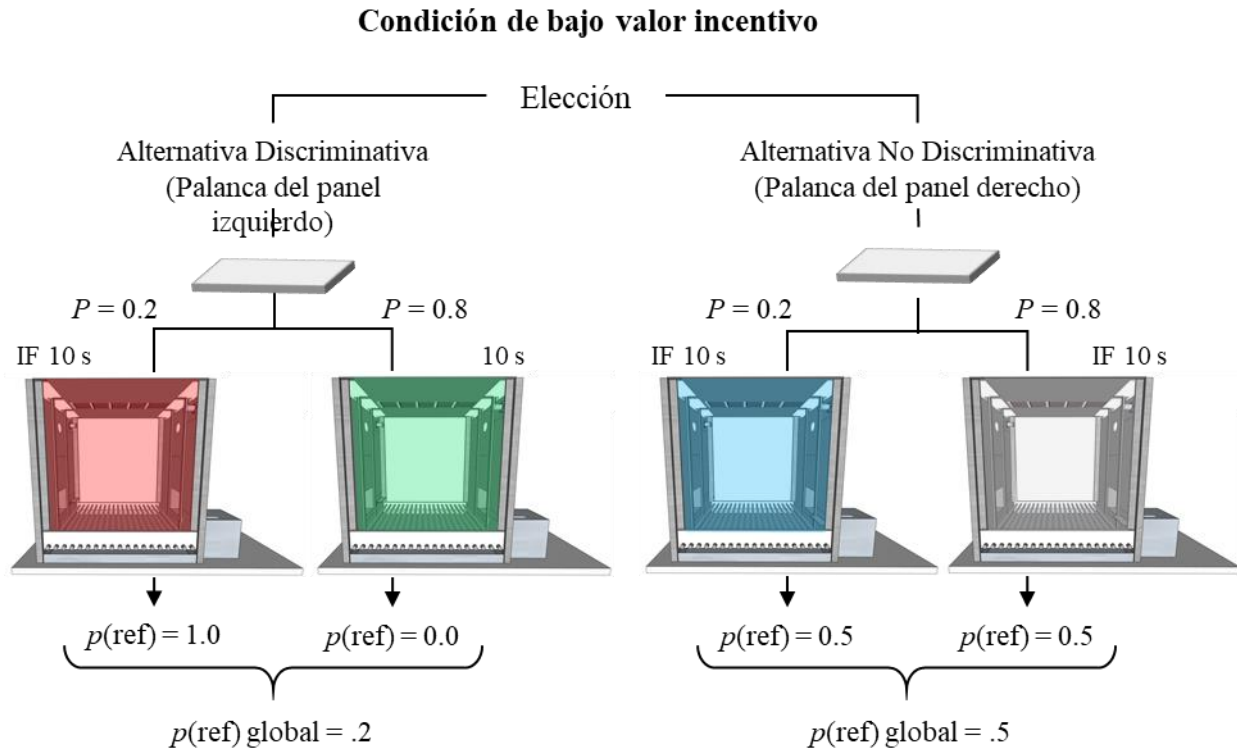


Figura 5. Esquemización del Entrenamiento. Procedimiento de elección subóptima del Experimento 1. Condición de bajo valor incentivo. El color asociado a $p(\text{ref}) = 1$ y $p(\text{ref}) = 0$ fue contrabalanceado entre los sujetos, así como la palanca asociada a la alternativa discriminativa y no discriminativa. Durante el entrenamiento en reversión la posición de las alternativas fue invertido. $p(\text{ref})$ = probabilidad de reforzamiento.

Resultados

Se analizó la proporción de elección de cada sujeto por la alternativa discriminativa, tal como suele reportarse en la literatura de elección subóptima desde sus inicios (Stagner & Zentall, 2010) durante el entrenamiento y el entrenamiento en reversión, esta proporción resulta del número de veces que se eligió la alternativa discriminativa entre el número total de ensayos de elección realizados. Como era necesario un criterio de estabilidad en las últimas cinco sesiones para dar por concluida estas fases, cada sujeto terminó por realizar un número diferente de sesiones, además, dado que la preferencia de los sujetos debía ser consistente y libre de sesgo, se consideraron las primeras 2 fases consecutivas donde la preferencia por alguna alternativa se mantuviera, por ejemplo, si un animal mostraba conducta subóptima durante el entrenamiento,

era necesario que volviera a mostrar conducta subóptima durante la reversión.

Desafortunadamente, 3 de los 14 sujetos invertían su preferencia cada vez que se realizaba una reversión, es decir, su elección estaba sesgada por la posición de las alternativas más que por las consecuencias asociadas a estas, por esta razón, tuvieron que ser descartados, por lo que sus datos no fueron considerados para este experimento y ya no participaron en el Experimento 2.

La Figura 6 muestra, para cada sujeto, la proporción de elección para la alternativa discriminativa durante cada una de las sesiones de las fases de entrenamiento y reversión. Dado que cada sujeto hizo un diferente número de sesiones en cada fase, sólo se consideraron las últimas cinco sesiones para el análisis estadístico de los datos ya que estas reflejan el momento donde se encontró estabilidad en la preferencia de los animales. El promedio de proporción de elección de los animales, durante las últimas 5 sesiones de la fase de entrenamiento fue de $.23 \pm .10$ (Media \pm SEM), el cual fue estadísticamente diferente de la indiferencia de acuerdo con una prueba t de dos colas contra .5 ($t(10) = -2.58, p = .03$). Para la fase de reversión, la mayoría de los sujetos recuperaron su preferencia con excepción de dos sujetos que se mostraron indiferentes (PAL2 y PAL5), en esta fase el promedio de proporción de elección fue de $.295 \pm .11$, pero no fue significativamente diferente de la indiferencia ($t(10) = -1.9, p = .09$)

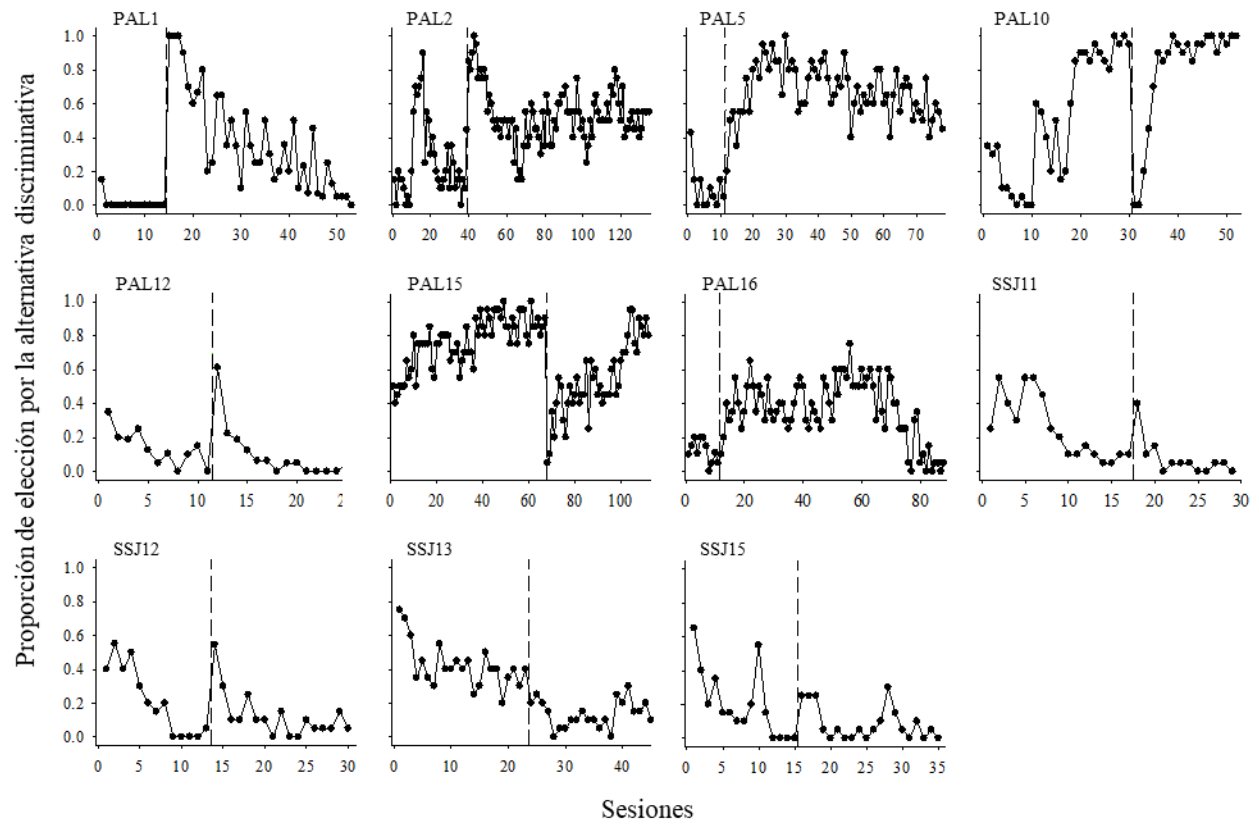


Figura 6. Promedio de proporción de elección por la alternativa discriminativa para cada sujeto a lo largo de sus sesiones, en la condición de bajo valor incentivo. La línea punteada indica el final del entrenamiento e inicio de la fase de reversión.

Además, para demostrar que los animales fueron sensibles a las consecuencias asociadas a los estímulos de cada alternativa y por lo tanto su elección no pudo deberse a la ausencia de discriminación; se registró la cantidad de respuestas realizadas a la palanca durante la presentación de cada uno de los estímulos. En el panel A de la Figura 7, se muestra el promedio de respuestas de cada sujeto en los ensayos forzados de la alternativa discriminativa durante el S+ y durante el S-. Si los animales, pudieron diferenciar la función de los estímulos en la alternativa discriminativa, se esperaría que respondieran más veces ante el estímulo que predice reforzador y menos veces ante el estímulo que predice no reforzador, para ello se realizó un ANOVA de medidas repetidas de 2 (estímulo: S+ vs S-) x 2 (fase: entrenamiento vs reversión), con las

últimas cinco sesiones de los sujetos en cada fase el cual mostró diferencias estadísticamente significativas entre estímulos ($F(1,10) = 42.30, p < .001$); pero no entre fases ($F(1,10) = 0.75, p = .41$) ni en la interacción ($F(1,10) = 0.16, p = .70$). El promedio de respuestas de los sujetos al S+ durante el entrenamiento y la reversión fue de 7.66 ± 6.97 y 6.97 ± 1.04 respectivamente, mientras que para el S- fue de $.80 \pm .14$ y $.53 \pm .09$

Por otro lado, la Figura 7 en su panel B, muestra el promedio de respuestas de cada sujeto en los ensayos forzados de la alternativa no discriminativa para los estímulos 1 y 2 y dado que ambos están asociados a la misma probabilidad de reforzamiento, no se esperaban diferencias en la cantidad de respuestas; comprobado estadísticamente mediante un ANOVA de medidas repetidas de 2 (estímulo: S1 vs S2) x 2 (fase: entrenamiento vs reversión), el cual no mostró diferencias estadísticamente significativas entre estímulos ($F(1,10) = 1.07, p = .33$); fase ($F(1,10) = 0.04, p = .85$) ni en la interacción ($F(1,10) = 3.03, p = .11$). El promedio de respuestas de los sujetos al S1 durante el entrenamiento y la reversión fue de $7.01 \pm .1.02$ y $6.34 \pm .8$ respectivamente, mientras que para el S2 fue de $5.89 \pm .61$ y $6.31 \pm .38$.

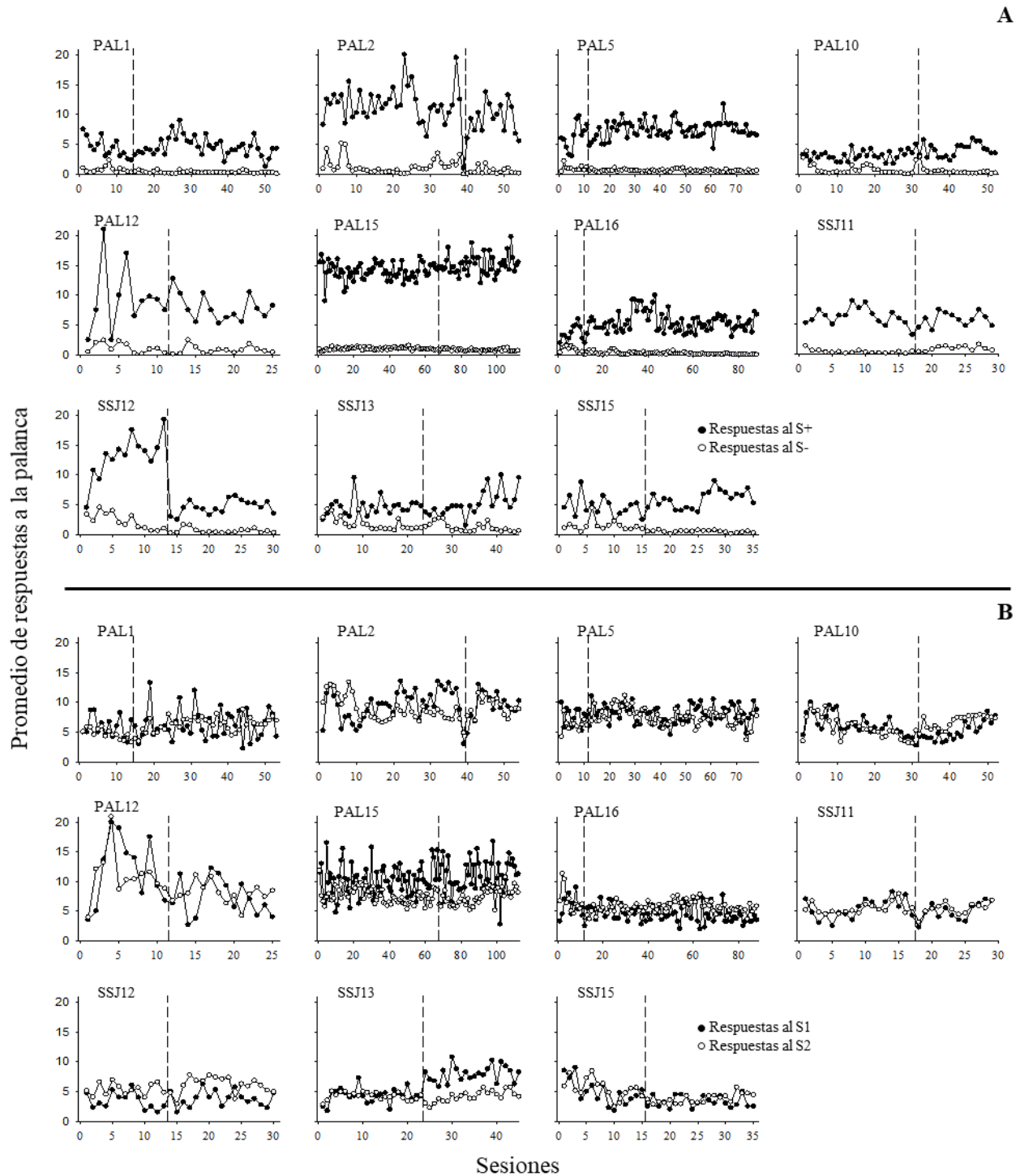


Figura 7. Panel A. Promedio de respuestas a la palanca durante los eslabones terminales de los estímulos de la alternativa discriminativa (S+, S-), para cada sujeto a lo largo de sus sesiones, en la condición de bajo valor incentivo. Panel B. Promedio de respuestas para los estímulos de la alternativa no discriminativa (S1, S2). La línea punteada indica el final del entrenamiento e inicio de la fase de reversión.

Experimento 2. Condición de alto valor incentivo

Se realizó un segundo experimento basado en el procedimiento clásico de elección subóptima (Stagner & Zentall, 2010), como condición control, para demostrar si ha sido la saliencia incentiva responsable de los hallazgos del Experimento 1, de esta manera, si la ausencia de estímulos con alto valor incentivo es responsable en la optimalidad mostrada en el Experimento 1, su presencia en este experimento (Experimento 2) debería guiar a conducta subóptima.

Método

Sujetos

Se utilizaron los 11 sujetos del Experimento 1 que mostraron una preferencia estable por alguna de las alternativas tanto en entrenamiento como en reversión. Las condiciones de privación y cuidado se mantuvieron iguales.

Aparatos

Se utilizaron cuatro cajas de condicionamiento operante (ver Figura 8) de 35 cm (largo) x 30 cm (ancho) x 37 cm (alto). El piso era una malla de acero inoxidable que cubría el total del área de trabajo (35 x 30 cm) y 5 cm de alto. Las cajas sólo tenían un panel del lado derecho que constaba de una lámina sólida con una perforación cuadrada de 5 cm localizada a 9.5 cm de la base de la caja en la cual se proveía acceso a un comedero para aves que era iluminado por una lámpara miniatura de 1 W cada vez que el comedero se levantaba, además, estaba equipado con un fotorreceptor, para detectar las entradas al comedero. El panel además contaba con tres perforaciones para teclas de respuesta de 2.5 cm de diámetro a 23 cm del piso y a 7.5 cm de distancia entre ellas. Cada tecla requería una fuerza de 0.15 N para ser activada y podían ser

iluminadas de color blanco, rojo o verde. Se colocó un foco de 1 W para luz general al centro del techo.

Las sesiones fueron monitoreadas mediante una cámara web (Logitech C170) colocada en la parte superior central del panel izquierdo. Las cajas estaban dentro de cubículos sonoamortiguadores con un ventilador acoplado para atenuar sonidos externos. La presentación de estímulos, así como la recolección de datos fueron controlados por computadora usando programación en lenguaje Medstate (Med-PC IV, MED Associates, Inc.).

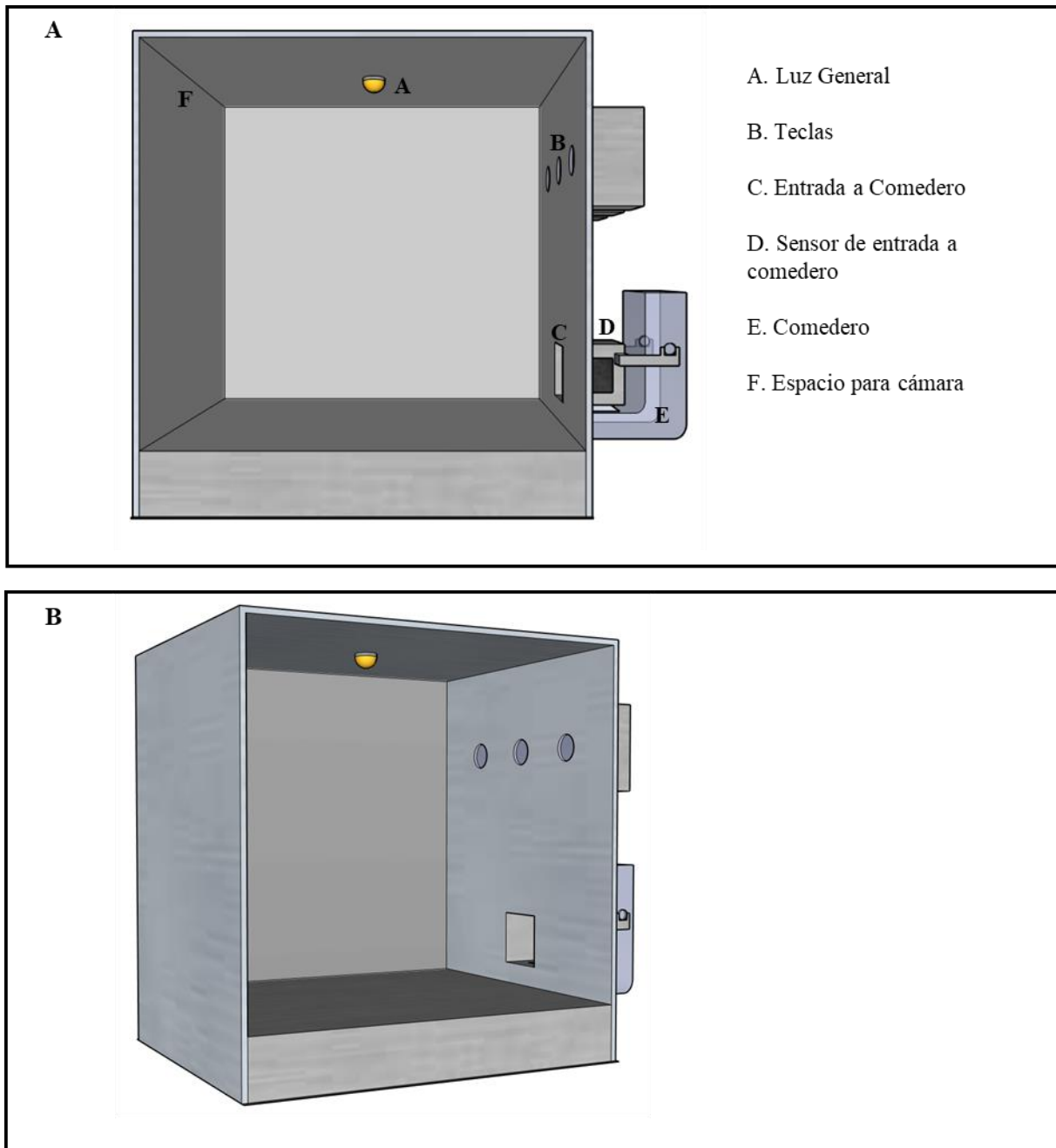


Figura 8. Modelo 3D realizado en SketchUp de las cajas operantes utilizadas en el experimento 2 (condición de alto valor incentivo). La figura A señala y enumera los componentes utilizados desde la perspectiva del experimentador, mientras que la figura B muestra la vista lateral de la caja, con todos sus componentes.

Procedimiento

Habitación

Debido a que este experimento se realizó en un entorno con instrumentación diferente, los animales, recibieron de 1 a 5 sesiones de habitación a la caja operante, cada sesión consistía en un programa de TF 25 s, que otorgaba acceso al comedero por 5 segundos, todos los demás estímulos permanecían apagados. La sesión finalizaba después de 50 minutos o al realizarse 30 entradas al comedero. Esta fase se dio por concluida cuando el animal completaba el programa por el criterio de entradas al comedero.

Automoldeamiento

Con el fin de facilitar el picoteo, que sería la operante requerida para responder en este experimento, los animales ingresaron a un programa basado en automoldeamiento (Brown & Jenkins, 1968). Al inicio de cada ensayo se iluminaba la tecla central de color blanco durante 8 s, seguida de 3 s de acceso a comedero, todos los demás estímulos permanecían apagados y un nuevo ensayo comenzaba después de un IEE de 52 s, los picotazos a la tecla no tenían ningún efecto en esta. El programa finalizaba después de 1 hora y se realizó durante 7 sesiones.

Moldeamiento en picoteo a la tecla

Debido a que algunos animales no adquirieron la conducta de picoteo en el programa de automoldeamiento, se realizó un nuevo entrenamiento para todos mediante el método de aproximaciones sucesivas (Skinner, 1951). Para esta fase, se utilizaron 2 s de acceso al comedero como reforzador y fueron requeridas de 3 a 7 sesiones, hasta que la conducta de picoteo se presentara de manera confiable, cada que la tecla central era encendida.

Entrenamiento en picoteo a la tecla

Para garantizar que los animales no tuvieran preferencia por alguna de las teclas o por alguno de los estímulos; se entrenó a los sujetos a responder en las dos teclas laterales, las cuales

podían iluminarse de 3 posibles colores (blanco, rojo o verde), asociadas a un programa de RF1 que proveía 3 s de acceso al comedero seguido de un IEE en el cual todos los estímulos permanecían apagados por 10 s. Las teclas se presentaron de forma cuasi aleatoria con la restricción de que la misma tecla no podía presentarse más de 3 veces seguidas. El programa finalizaba después de una hora o al completar 72 ensayos, 36 en la tecla derecha y 36 en la izquierda (12 ensayos de cada color). El criterio para pasar a la siguiente fase fue que los animales presentaran una mediana de latencias menor a 5 s y que la diferencia entre las medianas de cada tecla fuera menor a 1. Esta fase duró de 1 a 3 sesiones.

Entrenamiento. Procedimiento clásico de elección subóptima

Durante este entrenamiento, de la misma forma que en el Experimento 1, se presentaron dos tipos de ensayos: forzados y de elección, asociados a un programa encadenado de dos eslabones. Durante el eslabón inicial de los ensayos forzados, sólo se iluminaba una de las teclas laterales (izquierda o derecha) de color blanco, mientras que la otra permanecía inactiva. Por otro lado, en los ensayos libres ambas teclas se iluminaban hasta que el animal respondiera en alguna de ellas, la opción no elegida permanecía inactiva. En ambos casos un solo picotazo en la tecla (RF1) permitía la transición al eslabón terminal señalado con un cambio de color en la tecla presionada. Si la tecla correspondía a la alternativa discriminativa, en el 20% de los ensayos, la tecla se iluminaba de color rojo durante un IF10 s que siempre era seguido de reforzamiento (3 segundos de acceso al comedero) y en el 80% de los ensayos, un picotazo en esa misma tecla la iluminaba de color verde durante 10 segundos y finalizaba sin que fuera necesaria una respuesta. En esta alternativa la tasa global de reforzamiento era del 20%.

Cuando la tecla elegida correspondía a la alternativa no discriminativa, en el 20% de los ensayos, se iluminaba del mismo color presentado el 20% de las veces en la alternativa discriminativa (e.g. rojo), durante un IF 10 s que el 50% de las veces era seguido por

reforzamiento, mientras que el 80% de las veces se iluminaba del color presentado el 80% de las veces en la alternativa discriminativa (e.g. verde), durante un IF 10 s proveyendo reforzamiento el 50% de las veces. Para esta alternativa la tasa global de reforzamiento fue de del 50%.

Cada sesión constaba de 60 ensayos que fueron divididos en 20 ensayos forzados para cada alternativa y 20 ensayos de elección, presentados de forma aleatoria cuasi aleatoria (no se podían presentar más de 3 veces seguidas el mismo tipo de ensayo ni más de 3 veces seguidas la misma alternativa en el caso de los ensayos forzados). Los ensayos fueron separados por un IEE de 10s, durante el cual todos los estímulos permanecieron inactivos. La posición de la alternativa discriminativa, así como los colores asociados a presencia o ausencia de reforzamiento fueron contrabalanceados entre los sujetos. La Figura 9 muestra un esquema de esta fase. El número de sesiones requeridas para dar por concluida esta fase, fue diferente para cada sujeto, los sujetos debían haber completado por lo menos diez sesiones y no mostrar alguna tendencia a incrementar o disminuir su proporción de elección por la alternativa discriminativa en las últimas 5 sesiones.

Entrenamiento en reversión

Durante esta fase la localización de la alternativa discriminativa y no discriminativa fue invertida, con el objetivo de verificar que la elección se debía a las contingencias asociadas a cada alternativa y no a la posición de la tecla iluminada. El criterio de finalización fue el mismo que en la fase anterior. En caso de que la preferencia obtenida fuera inconsistente con la fase anterior, la ubicación de las alternativas era revertida a su posición durante el entrenamiento y el animal debía continuar bajo los mismos criterios de esa fase, para identificar si su elección se debía a un sesgo o a la adquisición de una nueva preferencia.

Procedimiento de elección Subóptima

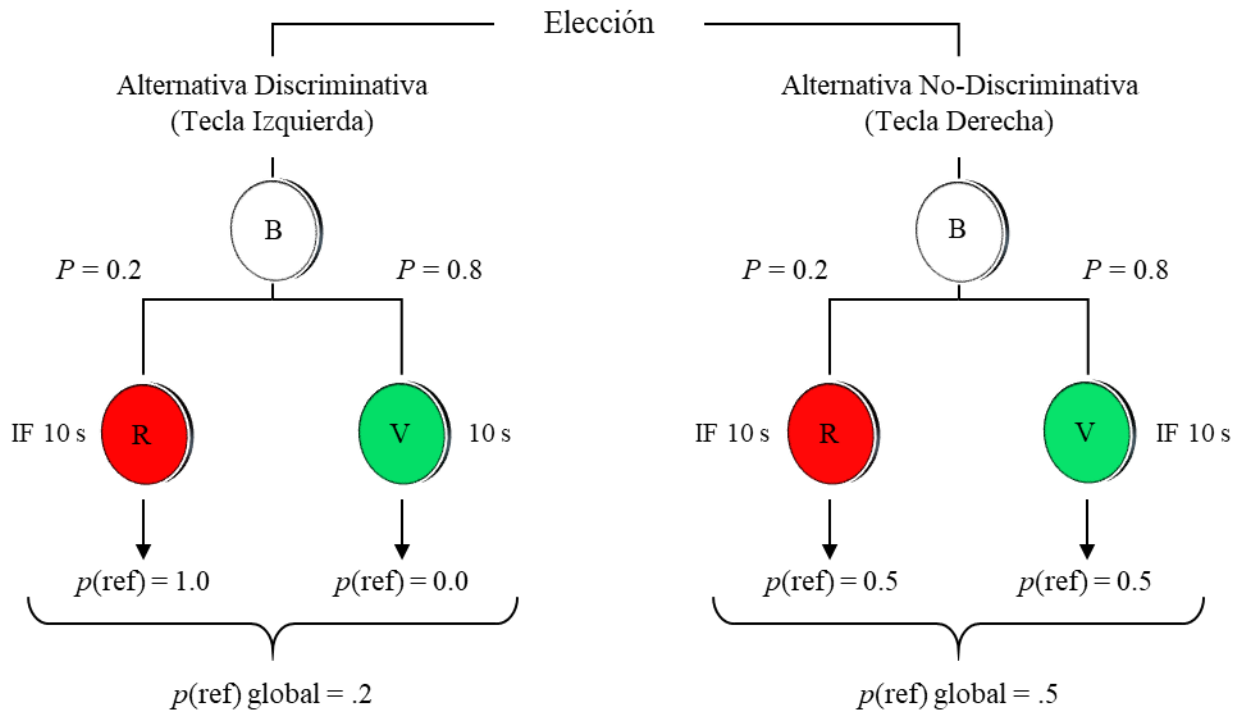


Figura 9. Esquematación del Entrenamiento. Procedimiento clásico elección subóptima del Experimento 2. Condición de alto valor incentivo. El color asociado a $p(\text{ref}) = 1$ y $p(\text{ref}) = 0$ fue contrabalanceado entre los sujetos, así como la tecla asociada a la alternativa discriminativa y no discriminativa. Durante el entrenamiento en reversión la posición de las alternativas fue invertido. $p(\text{ref})$ = probabilidad de reforzamiento.

Resultados

Como en la condición anterior, la principal variable de análisis fue la proporción de elección para la alternativa discriminativa, además, debido a que los criterios para la fase de entrenamiento y reversión fueron los mismos que en el Experimento 1, cada sujeto realizó un número diferente de sesiones y reversiones, por lo que nuevamente se consideraron las primeras 2 condiciones consecutivas donde la preferencia por alguna alternativa se mantuviera. En este experimento, ninguno de los sujetos fue excluido. La Figura 10 muestra la proporción de elección para la alternativa discriminativa de cada sujeto durante cada una de las sesiones de las fases de entrenamiento y reversión. Al igual que en el experimento anterior, se utilizaron las últimas cinco

sesiones de cada sujeto en cada fase para los análisis estadísticos ya que estas reflejan el momento en que se encontró estabilidad en la preferencia de los animales. De esta manera, el promedio de proporción de elección obtenido, durante las últimas 5 sesiones de la fase de entrenamiento, fue de $.87 \pm .07$, el cual fue estadísticamente diferente de la indiferencia de acuerdo con una prueba *t* de dos colas contra $.5$ ($t(10) = 5.58, p < .001$). Por otro lado, el promedio de proporción de elección que en la fase de reversión fue de $.76 \pm .11$ también fue estadísticamente diferente de la indiferencia ($t(10) = 2.35, p = .04$).

Nuevamente, para demostrar que los animales fueron sensibles a las consecuencias asociadas a los estímulos de cada alternativa, se registró la cantidad de picotazos realizados a la tecla durante la presentación de cada uno de los estímulos. En el panel A de la Figura 11, se muestra el promedio de respuestas de cada sujeto en los ensayos forzados de la alternativa discriminativa para el S+ y para el S-. Para el caso del S+ el promedio de respuestas obtenido a partir de las últimas cinco sesiones de cada sujeto durante el entrenamiento fue de 18.35 ± 2.12 y de 16.09 ± 1.86 durante la reversión; mientras que el promedio para el S- fue de $1.37 \pm .46$ durante el entrenamiento y de $1.60 \pm .34$ durante la reversión. Se realizó un ANOVA de medidas repetidas de 2 (estímulo: S+ vs S-) x 2 (fase: entrenamiento vs reversión), el cual mostró diferencias estadísticamente significativas entre estímulos ($F(1, 10) = 122.00, p < .001$), pero no entre fases ($F(1, 10) = 0.57, p = .47$) ni en la interacción ($F(1, 10) = 1.30, p = .28$).

Por otro lado, la Figura 11 en su panel B, muestra el promedio de respuestas de cada sujeto en los ensayos forzados de la alternativa no discriminativa para los estímulos 1 y 2. En este caso, dado que ambos estímulos están asociados a la misma probabilidad de reforzamiento, no se esperaban diferencias en la cantidad de respuestas, lo cual fue comprobado estadísticamente mediante un ANOVA de medidas repetidas de 2 (estímulo: S1 vs S2) x 2 (fase: entrenamiento vs reversión), el cual en efecto, no mostró diferencias estadísticamente significativas entre estímulos

($F(1, 10) = 0.05, p = .83$); fase ($F(1, 10) = 0.08, p = .79$), ni en la interacción ($F(1, 10) = 0.72, p = .42$). El promedio de respuestas de los sujetos al S1 durante el entrenamiento y la reversión fue de 16.27 ± 1.53 y de 16.05 ± 2.57 , respectivamente, mientras que para el S2 fue de 15.64 ± 2.02 y de 16.99 ± 2.86

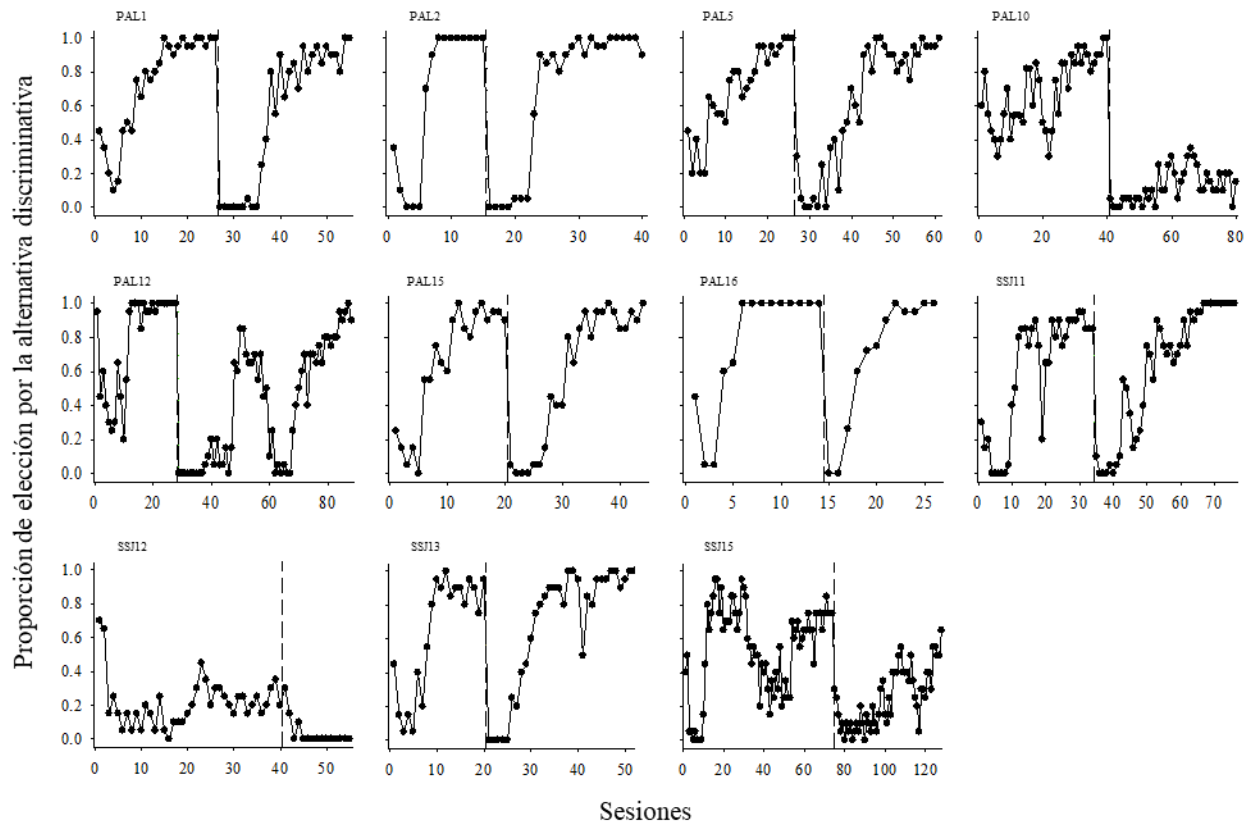


Figura 10. Promedio de proporción de elección por la alternativa discriminativa para cada sujeto a lo largo de sus sesiones, en la condición de alto valor incentivo. La línea punteada indica el final del entrenamiento e inicio de la fase de reversión.

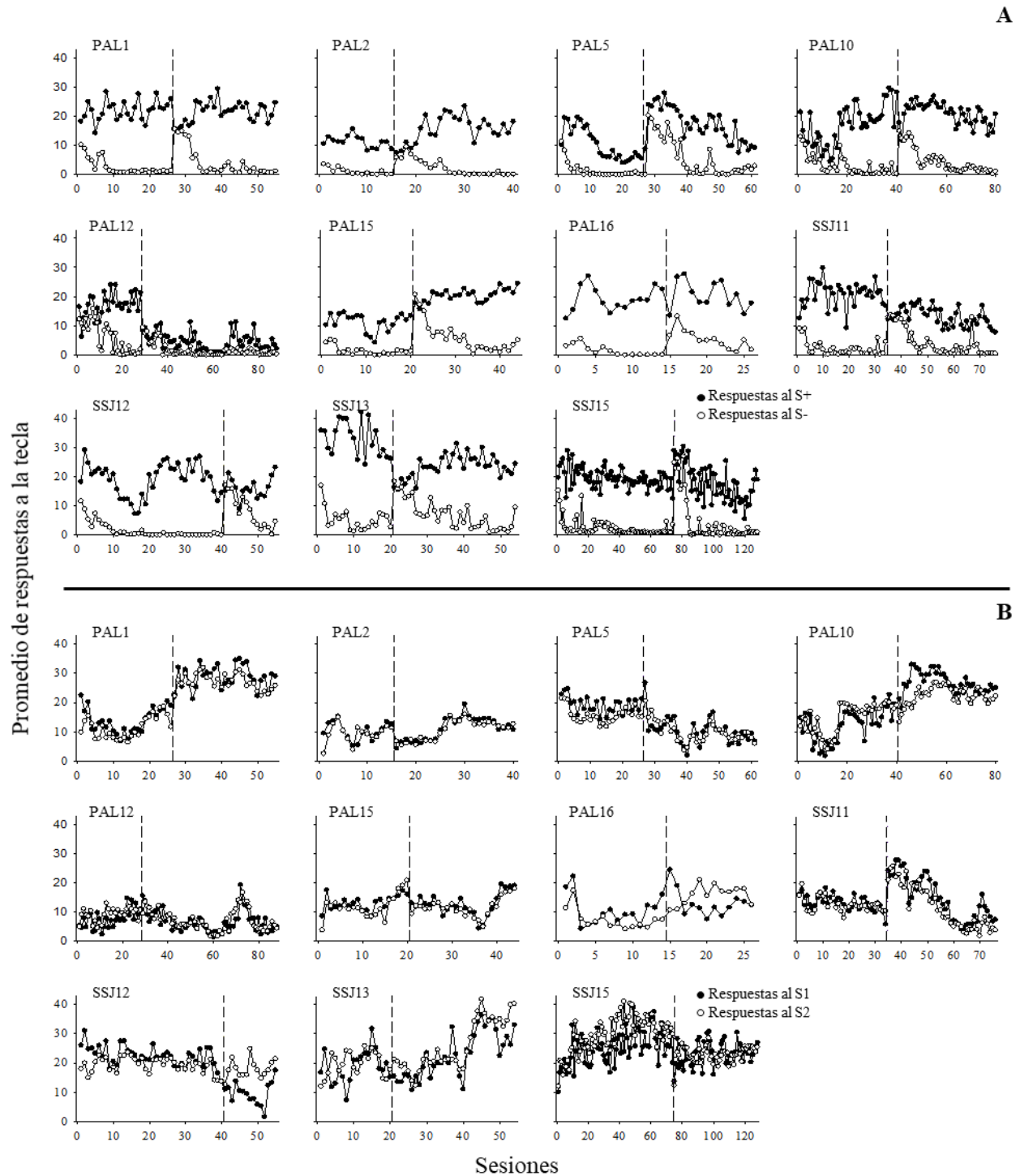


Figura 11. Panel A. Promedio de respuestas a la tecla durante los eslabones terminales de los estímulos de la alternativa discriminativa (S+, S-), para cada sujeto a lo largo de sus sesiones, en la condición de alto valor incentivo. Panel B. Promedio de respuestas para los estímulos de la alternativa no discriminativa (S1, S2). La línea punteada indica el final del entrenamiento e inicio de la fase de reversión.

Comparación entre condiciones

Resumiendo, los hallazgos muestran que en el Experimento 1, condición de bajo valor incentivo, se encontró un promedio de proporción de elección por la alternativa discriminativa en el entrenamiento de $.23 \pm .10$ y de $.295 \pm .11$ durante la reversión, mientras que en el Experimento 2, condición de alto valor incentivo, obtuvieron un promedio de proporción de elección de $.87 \pm .07$ durante el entrenamiento y $.76 \pm .11$ durante la reversión. Estas aparentes diferencias entre condiciones fueron también apoyadas por un ANOVA de medidas repetidas de 2 (condición: experimento 1 vs experimento 2) x 2 (fase: entrenamiento vs reversión) que mostró diferencias estadísticamente significativas entre condiciones ($F(1, 10) = 16.62, p < .01$); pero no entre fases ($F(1, 10) = 0.31, p = .59$); ni en la interacción ($F(1, 10) = 3.41, p = .09$). La Figura 12 ilustra gráficamente las diferencias entre condiciones y fases mostrando el promedio de proporción de elección de los animales durante las últimas 5 sesiones de cada caso.

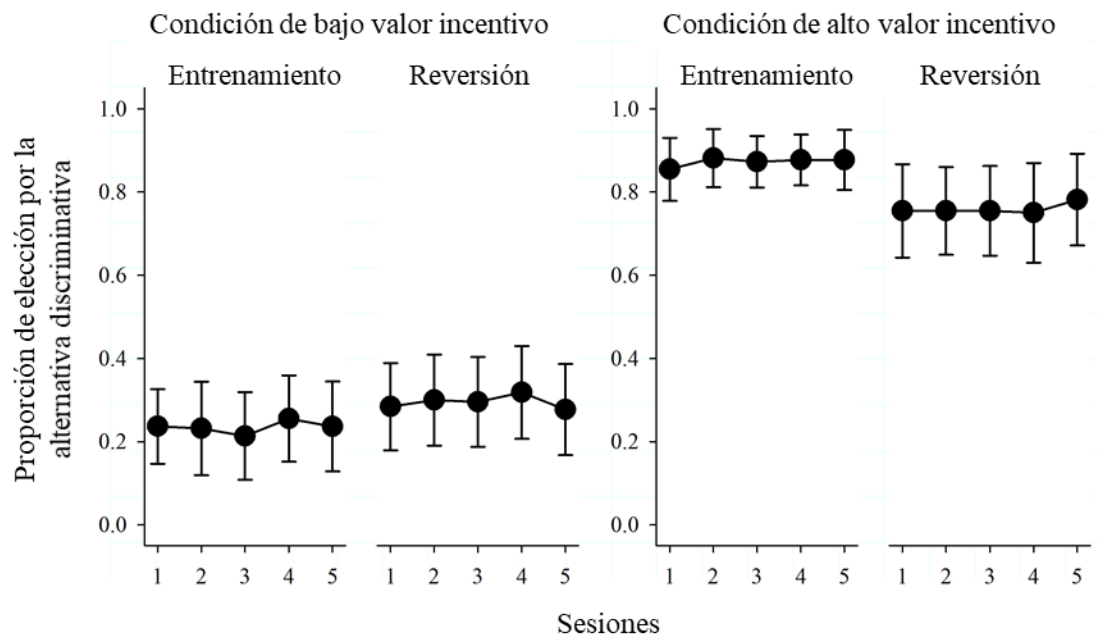


Figura 12. Promedio de proporción de elección por la alternativa discriminativa en las últimas 5 sesiones de cada fase, durante la condición de bajo valor incentivo (gráfica izquierda) y de alto valor incentivo (gráfica derecha).

Por otro lado, la Figura 13 compara la discriminación de los animales entre experimentos durante las últimas cinco sesiones de cada fase; para ello, en lugar de utilizar el promedio de respuestas, se utiliza un índice de discriminación (idx), esto, debido a que en general se observó un mayor número de respuestas a los estímulos en la condición de alto valor incentivo cuando la operante era picar con respecto a la de bajo valor incentivo cuando la operante era pisar, sin embargo, estas diferencias en la cantidad de respuestas ya se han reportado cuando se utilizan estas dos topografías (Green & Holt, 2003), lo cual no necesariamente indica una mejor discriminación, ya que ésta depende de la diferencia de respuestas entre los estímulos y no sólo de la cantidad de respuestas realizadas. El índice de discriminación para la alternativa discriminativa es calculado de la siguiente manera: $\text{Respuestas al S+} / ((\text{Respuestas al S+}) + (\text{Respuestas al S-}))$, siendo los valores cercanos a 1, indicadores de una buena discriminación y los cercanos a .5, indicadores de indiferencia. Los índices de discriminación para la alternativa discriminativa en la condición de bajo valor incentivo fueron de $.89 \pm .02$ en entrenamiento y de $.92 \pm .01$ reversión; mientras que en la condición de alto valor incentivo fueron de $.94 \pm .01$ y $.91 \pm .02$ (entrenamiento y reversión, respectivamente). Un ANOVA de medidas repetidas de 2 (condición) x 2 (fase), no mostró diferencias estadísticamente significativas entre condiciones ($F(1, 10) = 1.14, p = .31$); fases ($F(1, 10) < .01, p = .96$); ni en la interacción ($F(1, 10) = 4.65, p = .06$).

También se obtuvieron los índices de discriminación para la alternativa no discriminativa, para este caso, la fórmula fue la siguiente: $\text{Respuestas al S1} / ((\text{Respuestas al S1}) + (\text{Respuestas al S2}))$. Así, los índices para la condición de bajo valor incentivo en el entrenamiento y reversión fueron de $.52 \pm .03$ y $.48 \pm .03$ respectivamente, y durante la condición de alto valor incentivo, fueron de $.52 \pm .01$ y $.49 \pm .03$. Un ANOVA de medidas repetidas de 2 (condición) x 2 (fase), no

mostró diferencias estadísticamente significativas entre condiciones ($F(1, 10) = < .01, p = .98$); fases ($F(1, 10) = 3.59, p = .08$; ni en la interacción ($F(1, 10) = .09, p = .77$).

Por último, se incluye en el anexo, un resumen del promedio de proporción de elección, respuestas e índice de discriminación de cada sujeto, en los dos experimentos durante las últimas cinco sesiones de cada fase.

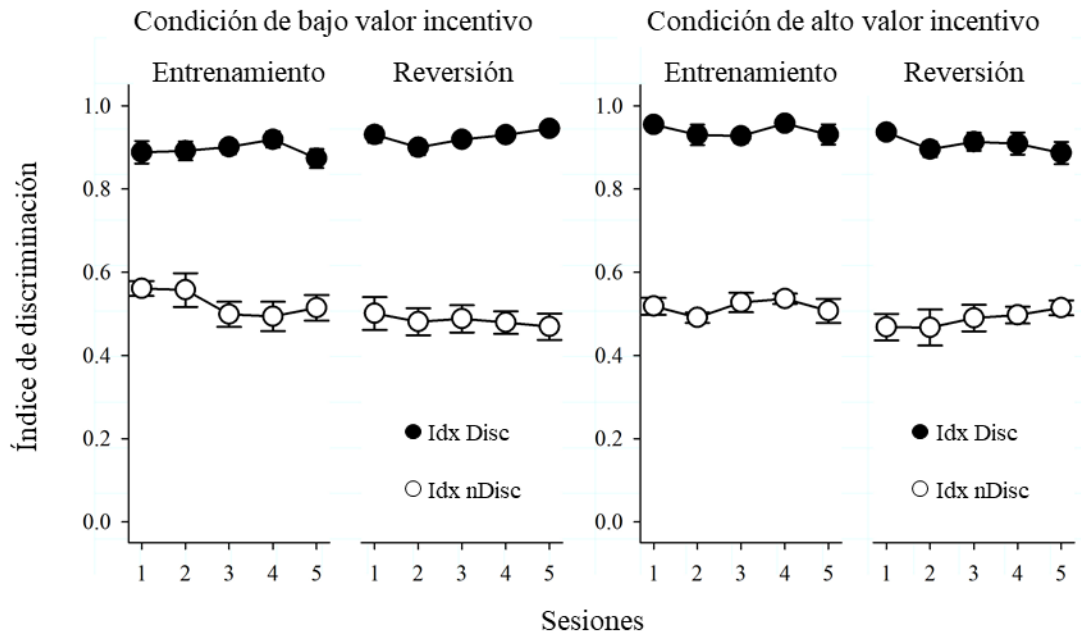


Figura 13. Promedio de índice de discriminación (idx) en las últimas 5 sesiones de cada fase, para la alternativa discriminativa (círculos llenos) y para la alternativa no discriminativa (círculos vacíos); durante la condición de bajo valor incentivo (gráfica izquierda) y de alto valor incentivo (gráfica derecha).

Discusión

El presente estudio trató de aportar evidencia a favor de la saliencia incentiva de los estímulos como posible explicación del fenómeno de elección subóptima en palomas, para ello, en un primer experimento se redujo el supuesto valor incentivo de las teclas usadas en el procedimiento de elección subóptima, haciendo dos modificaciones. Primero, se cambiaron estos estímulos por luces ambientales; y segundo, se utilizaron palancas para que los animales pudieran responder en el programa, las cuales debían ser pisadas, ya que era posible que la atribución de saliencia incentiva se debiera a la similitud de la operante con la topografía de respuesta consumatoria típica de la especie (es decir, picar) (Zentall, Smith, et al., 2019).

Los resultados del primer experimento, condición de bajo valor incentivo, mostraron que los animales prefirieron la alternativa no discriminativa, misma que estaba asociada a una mayor tasa global de reforzamiento, en pocas palabras, eligieron de forma óptima. Además, mostraron altos niveles de discriminación para los estímulos de la alternativa discriminativa, sustentando así que su preferencia no pudo deberse a un fallo en la identificación de la función de los estímulos, lo cual es importante ya que una primera condición para la atribución de saliencia incentiva es que los animales identifiquen un estímulo predictor de reforzamiento, por lo que, en este caso, pese a haberlo identificado, es posible que no ocurriera esta atribución. Además, se ha demostrado que cuando los animales eligen entre dos alternativas no discriminativas tienden a elegir aquella con mayor tasa de reforzamiento, pero si una de ellas contiene estímulos discriminativos, su preferencia se sesga hacia esta última, sin importar su tasa de reforzamiento (Stagner & Zentall, 2010; Zentall & Stagner, 2011).

Por otro lado, el Experimento 2, condición de alto valor incentivo, fungió como control del Experimento 1, de esta manera, si la atribución de valor incentivo estaba jugando un papel importante en el fenómeno de elección subóptima, su presencia debía mostrar conducta

subóptima en los mismos sujetos, tal como su ausencia en el Experimento 1 mostró conducta óptima. Probablemente un mejor diseño experimental habría contrabalanceado la presentación de las condiciones entre los animales, sin embargo, las pruebas piloto de este experimento mostraron que una vez que la conducta de picar se establece en los animales, esta interfiere con el moldeamiento de pisar a la palanca, ya que los animales comienzan a picar la palanca en lugar de pisarla.

Los resultados del Experimento 2 cumplieron satisfactoriamente su propósito, los animales invirtieron la preferencia mostrada en el Experimento 1 y en este caso, tal como se ha reportado por años en la literatura del tema, mostraron conducta subóptima, siendo su proporción de elección por la alternativa discriminativa estadísticamente diferente de la encontrada en el Experimento 1. Adicionalmente, los análisis de discriminación mostraron que sus preferencias no pudieron deberse a la ausencia de ésta.

En resumen, la conducta de elección óptima mostrada en la condición de bajo valor incentivo, junto con el cambio a elección subóptima en la condición de alto valor incentivo, apoya la hipótesis de que la atribución de saliencia incentiva puede ser un fuerte determinante de las preferencias de los animales en el procedimiento de elección subóptima, al menos en palomas, ya que en las ratas se ha encontrado evidencia contradictoria. Por un lado, Chow et al. (2017) usando una palanca como estímulo predictor de reforzador, por el supuesto alto valor incentivo para la especie y un apagón como estímulo predictor de no reforzador, encontraron conducta de elección subóptima en ratas, mientras que Martínez, et. al. (2017) demostraron conducta óptima en la mayoría de sus ratas aun cuando el estímulo predictor de reforzador fue una palanca, aunque, algunos de sus animales, mostraron conducta subóptima.

Es importante mencionar que la atribución de saliencia incentiva no necesariamente se expresa en todos los organismos de una población; tal como se mencionó en la introducción de

este trabajo, al presentar un estímulo predictor de alimento, los animales suelen responder a él de dos maneras, acercándose a él e interactuando recreado las respuestas consumatorias de su especie (seguimiento de señales), o aproximándose a la ubicación donde aparecerá el alimento (seguimiento de objetivos), siendo los primeros quienes atribuyen saliencia incentiva a los estímulos (Robinson & Flagel, 2009) (Meyer et al., 2012). Para el caso de las ratas, el “fenotipo conductual” de seguimiento de señales tiene una prevalencia cercana al 35% (Meyer et al., 2012), lo que podría explicar por qué solo algunos animales del experimento de Martínez et al. (2017), mostraron conducta subóptima aunque no la gran mayoría.

Recientemente, se han encontrado diferencias en la preferencia por la alternativa discriminativa y subóptima, en ratas clasificadas como seguidoras de señales y seguidoras de objetivos, siendo las seguidoras de señales quienes muestran mayor suboptimalidad respecto a las seguidoras de objetivos, las cuales tienden a elegir la alternativa óptima (Orduña & Alba, 2020), por lo que para el caso de las palomas, no es de sorprender que la gran mayoría muestre conducta subóptima dado que en esta especie prevalece el fenotipo de seguidoras de señales (Brown & Jenkins, 1968).

Finalmente, aunque parece que la atribución de saliencia incentiva puede explicar adecuadamente los resultados de este trabajo, como se mencionó al principio, el diseño experimental no constó de una única manipulación por lo que dedicaré esta última sección a identificar y abordar las posibles explicaciones alternas que puedan dar cuenta de los resultados.

La elección óptima dependió del uso de luces ambientales

El uso de estímulos difusos como lo son las luces ambientales en lugar de teclas iluminadas pretendía reducir la atribución de saliencia incentiva, previniendo la aparición de conducta de seguimiento de señales ya que se ha visto que cuando los estímulos predictores de reforzamiento son difusos, los animales son más propensos a mostrar conducta de seguimiento de

objetivos que cuando son focalizados (Holland, 1977; Meyer et al., 2014), posiblemente, porque de esta manera los animales no tienen un objeto tangible al cual responder, lo que además termina por prevenir la atribución de saliencia incentiva. Sin embargo, se ha visto que la conducta de seguimiento de señales puede aparecer aun cuando los estímulos predictores de reforzador sean difusos, ya que su presencia genera conducta de picoteo si hay una tecla disponible (Patterson & Winokur, 1973), por lo tanto, el cambio de estímulos no parece suficiente para evitar la atribución de saliencia incentiva.

La elección óptima dependió del cambio de operante

Una de las principales condiciones para la atribución de saliencia incentiva, es que un estímulo predictor de reforzador sea capaz de generar conducta de seguimiento de señales (Robinson & Flagel, 2009), y la topografía que suele acompañar a esta conducta depende en cierta medida de cómo los animales interactúan con el EI predicho por el S+ (Bertram O Ploog, 2014; Timberlake & Grant, 1975; Zentall, Smith, et al., 2019). Cuando el EI es comida, las ratas responden al S+ mordisqueándolo o lamiéndolo (Peterson, Ackilt, Frommer, & Hearst, 1972), mientras que las palomas lo picotearán (Brown & Jenkins, 1968), y esta topografía puede variar incluso si el EI es agua (Jenkins & Moore, 1973) o según sus dimensiones, si este es grande o pequeño (Ploog & Zeigler, 1996). Por ello, establecer pisar una palanca como operante, sin cambiar además los estímulos de la tarea, es decir las teclas iluminadas, la tendencia natural de picar en las palomas habría prevalecido sobre pisar las palancas, como se observó durante un experimento piloto, similar a los casos reportados por Breland y Breland (1961) cuando las conductas aprendidas debían competir con las biológicamente establecidas, lo que difícilmente habría permitido la realización del experimento o modificado la preferencia de los animales, debido a la potencial atribución de saliencia incentiva de las teclas iluminadas.

La interacción entre el uso de luces ambientales y el cambio en la operante, como responsable de la elección óptima

Por último, considero que no ha sido una de las modificaciones sino la interacción entre ambas, la que ha disminuido la atribución de saliencia incentiva y ha resultado en una elección óptima. Como se acaba de mencionar, un primer paso para la atribución de saliencia, requiere que los organismos muestren conducta de seguimiento de señales (Flagel et al., 2009), que de acuerdo con una reciente propuesta teórica (Zentall, Smith, et al., 2019) basada en la idea general de la teoría de los sistemas de conducta (Timberlake, 1993), este comportamiento no es sino la activación del sistema de forrajeo natural de esa especie (si la recompensa es alimento). Es decir, un estímulo que provoca seguimiento de señales está reflejando su habilidad para evocar las respuestas consumatorias propias de una especie, y es entonces que éste puede ser atribuido con saliencia incentiva, y por tanto guiar la elección hacia la alternativa que contenga este estímulo (Zentall, Smith, et al., 2019); por lo que, si un estímulo no activa el sistema adecuado (o se previene al animal de poder activarlo), se prevendrá la atribución de saliencia incentiva y por lo tanto la aparición de conducta subóptima. Se presenta una ejemplificación de esta cadena de eventos en la Figura 14.

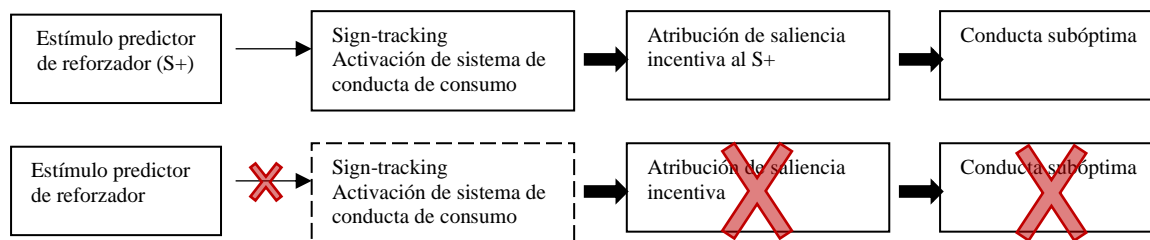


Figura 14. Receta para una elección subóptima. Esquematación de los eventos involucrados en el procedimiento de elección subóptima, cuando el estímulo predictor de reforzador activa conducta de seguimiento de señales (vertiente superior) y cuando no lo activa (vertiente inferior). Los tachos indican la anulación de alguno de los eventos.

En el Experimento 1, fue necesario tanto el cambio de picar a pisar, como la sustitución de teclas por luces ambientales, para poder anular toda forma de picoteo en las palomas, y de esta manera, al impedir la conducta de seguimiento de señales, prevenir la elección subóptima.

En conclusión, el presente trabajo representa la primera evidencia de elección óptima en palomas, cuando se modifica el valor incentivo de los estímulos utilizados en la tarea de elección subóptima, sin embargo, aunque la hipótesis de saliencia incentiva puede explicar satisfactoriamente estos hallazgos, aún tiene como desafío explicar los resultados con otras especies, ya que su generalidad podría dar solución a un problema que lleva años en la literatura.

Referencias

- Baum, W. M. (1981). Optimization and the matching law as accounts of instrumental behavior. *Journal of the Experimental Analysis of Behavior*, 36(3), 387–403.
<https://doi.org/10.1901/jeab.1981.36-387>
- Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, 16(11), 681–684. <https://doi.org/10.1037/h0040090>
- 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. <https://doi.org/10.1901/jeab.1968.11-1>
- Burns, M., & Domjan, M. (1996). Sign tracking versus goal tracking in the sexual conditioning of male Japanese quail (*Coturnix japonica*). *Journal of Experimental Psychology: Animal Behavior Processes*, Vol. 22, pp. 297–306. <https://doi.org/10.1037/0097-7403.22.3.297>
- Burns, M., & Domjan, M. (2000). Sign tracking in domesticated quail with one trial a day: Generality across CS and US parameters. *Animal Learning & Behavior*, 28(1), 109–119.
<https://doi.org/10.3758/BF03199776>
- Cartoni, E., Balleine, B., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental Transfer: A review. *Neuroscience & Biobehavioral Reviews*, 71, 829–848.
<https://doi.org/https://doi.org/10.1016/j.neubiorev.2016.09.020>
- 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. <https://doi.org/10.1016/j.bbr.2016.12.013>
- Cleland, G. G., & Davey, G. C. L. (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. <https://doi.org/10.1901/jeab.1983.40-47>
- Colagiuri, B., & Lovibond, P. F. (2015). How food cues can enhance and inhibit motivation to

obtain and consume food. *Appetite*, 84, 79–87.

<https://doi.org/https://doi.org/10.1016/j.appet.2014.09.023>

Crombag, H. S., Bossert, J. M., Koya, E., & Shaham, Y. (2008). Context-induced relapse to drug seeking: a review. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1507), 3233–3243. <https://doi.org/10.1098/rstb.2008.0090>

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. <https://doi.org/10.1037/xan0000160>

Daniels, C. W., & Sanabria, F. (2018). An associability decay model of paradoxical choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44(3), 258–271. <https://doi.org/10.1037/xan0000179>

Domjan, M. (2010a). Condicionamiento clásico: fundamentos. En *PRINCIPIOS DE APRENDIZAJE Y CONDUCTA* (6a ed., pp. 67–101). D.F.: Cengage Learning Editores S.A. de C.V.

Domjan, M. (2010b). Condicionamiento clásico: mecanismos. En *PRINCIPIOS DE APRENDIZAJE Y CONDUCTA* (6a ed., pp. 103–141). D.F.: Cengage Learning Editores S.A. de C.V.

Domjan, M. (2010c). Programas de reforzamiento y conducta de elección. En *PRINCIPIOS DE APRENDIZAJE Y CONDUCTA* (6a ed., pp. 198–215). D.F.: Cengage Learning Editores S.A. de C.V.

Edwards, W. (1961). Behavioral Decision Theory. *Annual Review of Psychology*, 12(1), 473–498. <https://doi.org/10.1146/annurev.ps.12.020161.002353>

Everitt, B. J., Dickinson, A., & Robbins, T. W. (2001). The neuropsychological basis of addictive behaviour. *Brain Research Reviews*, 36(2), 129–138.

[https://doi.org/https://doi.org/10.1016/S0165-0173\(01\)00088-1](https://doi.org/https://doi.org/10.1016/S0165-0173(01)00088-1)

Fantino, E., Dunn, R., & Meck, W. (1979). Percentage reinforcement and choice. *Journal of the Experimental Analysis of Behavior*, 32(3), 335–340. <https://doi.org/10.1901/jeab.1979.32-335>

Ferster, C. B., & Skinner, B. F. (1957). Schedules of reinforcement. In *Schedules of reinforcement*. <https://doi.org/10.1037/10627-000>

Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology*, 56, 139–148. <https://doi.org/https://doi.org/10.1016/j.neuropharm.2008.06.027>

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. <https://doi.org/10.3758/LB.37.4.289>

Green, L., & Holt, D. D. (2003). Economic and biological influences on key pecking and treadle pressing in pigeons. *Journal of the Experimental Analysis of Behavior*, Vol. 80, pp. 43–58. <https://doi.org/10.1901/jeab.2003.80-43>

Holland, P. C. (1977). Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *Journal of Experimental Psychology: Animal Behavior Processes*, Vol. 3, pp. 77–104. <https://doi.org/10.1037/0097-7403.3.1.77>

Hollis, K. L. (1997). Contemporary research on Pavlovian conditioning: A “new” functional analysis. *American Psychologist*, Vol. 52, pp. 956–965. <https://doi.org/10.1037/0003-066X.52.9.956>

Jenkins, H. M., & Moore, B. R. (1973). The form of the auto-shaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior*, 20(2), 163–181. <https://doi.org/10.1901/jeab.1973.20-163>

- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, *21*(3), 463–473. <https://doi.org/10.1901/jeab.1974.21-463>
- Laude, J. R., Pattison, K. F., & Zentall, T. R. (2012). Hungry pigeons make suboptimal choices, less hungry pigeons do not. *Psychonomic Bulletin & Review*, *19*(5), 884–891. <https://doi.org/10.3758/s13423-012-0282-2>
- 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), 12–21. <https://doi.org/10.1037/xan0000010>
- Lolordo, V. M. (1971). Facilitation of food-reinforced responding by a signal for response-independent food. *Journal of the Experimental Analysis of Behavior*, *15*(1), 49–55. <https://doi.org/10.1901/jeab.1971.15-49>
- 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. <https://doi.org/https://doi.org/10.1016/j.bbr.2017.12.028>
- 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. <https://doi.org/https://doi.org/10.1016/j.beproc.2017.06.012>
- Mazur, J. E. (1996). Choice with certain and uncertain reinforcers in an adjusting-delay procedure. *Journal of the Experimental Analysis of Behavior*, *66*(1), 63–73. <https://doi.org/10.1901/jeab.1996.66-63>
- Mazur, J. E. (2016). Reinforcement Schedules. Experimental Analyses and Applications. In *Learning and Behavior* (8th ed., pp. 142–167). New York: Routledge, Taylor & Francis Group.

- 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.
<https://doi.org/10.1002/jeab.192>
- Meyer, P. J., Cogan, E. S., & Robinson, T. E. (2014). The Form of a Conditioned Stimulus Can Influence the Degree to Which It Acquires Incentive Motivational Properties. *PLOS ONE*, *9*(6), e98163. Retrieved from <https://doi.org/10.1371/journal.pone.0098163>
- 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. Retrieved from <https://doi.org/10.1371/journal.pone.0038987>
- 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. <https://doi.org/10.3758/s13420-012-0065-7>
- Orduña, V., & Alba, R. (2020). The incentive salience of the stimuli biases rats' preferences in the “suboptimal choice” procedure. *Behavioural Processes*, *172*, 104057.
<https://doi.org/https://doi.org/10.1016/j.beproc.2020.104057>
- Patterson, D. D., & Winokur, S. (1973). Autoshaping pigeons' keypecking with a conditioned reinforcer. *Bulletin of the Psychonomic Society*, *1*(4), 247–249.
<https://doi.org/10.3758/BF03333358>
- Pattison, K. F., Laude, J. R., & Zentall, T. R. (2013). Environmental enrichment affects suboptimal, risky, gambling-like choice by pigeons. *Animal Cognition*, *16*(3), 429–434.
<https://doi.org/10.1007/s10071-012-0583-x>
- Peterson, G. B., Ackilt, J. E., Frommer, G. P., & Hearst, E. S. (1972). Conditioned Approach and Contact Behavior toward Signals for Food or Brain-Stimulation Reinforcement. *Science*,

177(4053), 1009 LP – 1011. <https://doi.org/10.1126/science.177.4053.1009>

Ploog, B O, & Zeigler, H. P. (1996). Effects of food-pellet size on rate, latency, and topography of autoshaped key pecks and gapes in pigeons. *Journal of the Experimental Analysis of Behavior*, 65(1), 21–35. <https://doi.org/10.1901/jeab.1996.65-21>

Ploog, Bertram O. (2014). Species-specific response-topography of chickens' and pigeons' water-induced autoshaped responding. *Behavioural Processes*, 106, 5–11. <https://doi.org/https://doi.org/10.1016/j.beproc.2014.03.007>

Pyke, G. H. (1984). Optimal Foraging Theory: A Critical Review. *Annual Review of Ecology and Systematics*, 15, 523–575. Retrieved from <http://www.jstor.org/stable/2096959>

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, Vol. 43, pp. 151–160. <https://doi.org/10.1037/0003-066X.43.3.151>

Robinson, T. E., Carr, C., & Kawa, A. B. (2018). The Propensity to Attribute Incentive Salience to Drug Cues and Poor Cognitive Control Combine to Render Sign-Trackers Susceptible to Addiction. In A. Tomie & J. Morrow (Eds.), *Sign-Tracking and Drug Addiction* (1a ed., pp. 75–105). Michigan: Michigan Publishing, University of Michigan Library.

Robinson, T. E., & Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues Through the Study of Individual Differences. *Biological Psychiatry*, 65(10), 869–873. <https://doi.org/10.1016/j.biopsych.2008.09.006>

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. <https://doi.org/https://doi.org/10.1006/lmot.1999.1030>

Shimp, C. P. (1969). Optimal behavior in free-operant experiments. *Psychological Review*, Vol. 76, pp. 97–112. <https://doi.org/10.1037/h0027311>

- Skinner, B. F. (1951). How to teach animals. *Scientific American*, 185(6), 26–29.
- 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. <https://doi.org/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. <https://doi.org/10.3758/PBR.17.3.412>
- Stairs, D. J., & Bardo, M. T. (2009). Neurobehavioral effects of environmental enrichment and drug abuse vulnerability. *Pharmacology, Biochemistry, and Behavior*, 92(3), 377–382. <https://doi.org/10.1016/j.pbb.2009.01.016>
- Timberlake, W. (1993). Behavior systems and reinforcement: an integrative approach. *Journal of the Experimental Analysis of Behavior*, 60(1), 105–128. <https://doi.org/10.1901/jeab.1993.60-105>
- Timberlake, W., & Grant, D. L. (1975). Auto-Shaping in Rats to the Presentation of Another Rat Predicting Food. *Science*, 190(4215), 690 LP – 692. <https://doi.org/10.1126/science.190.4215.690>
- Tomie, A. (2018). Introduction: The Role of Sign-Tracking in Drug Addiction. En A. Tomie & J. Morrow (Eds.), *Sign-Tracking and Drug Addiction* (1st ed., pp. 1–7). Recuperado de: <https://quod.lib.umich.edu/m/maize/mpub10215070/1:4/--sign-tracking-and-drug-addiction?rgn=div1;view=toc>
- Trujano, R. E., & Orduña, V. (2015). Rats are optimal in a choice task in which pigeons are not. *Behavioural Processes*, Vol. 119, pp. 22–27. <https://doi.org/10.1016/j.beproc.2015.07.010>
- Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information. *Scientific Reports*, 5(1), 13874. <https://doi.org/10.1038/srep13874>
- 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. <https://doi.org/10.1901/jeab.1969.12-511>
- Wyckoff Jr., L. B. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review*, Vol. 59, pp. 431–442. <https://doi.org/10.1037/h0053932>
- Zentall, T. R. (2011). Maladaptive “gambling” by pigeons. *Behavioural Processes*, 87(1), 50–56. <https://doi.org/10.1016/j.beproc.2010.12.017>
- Zentall, T. R., Andrews, D. M., & Case, J. P. (2019). Contrast between what is expected and what occurs increases pigeon’s suboptimal choice. *Animal Cognition*, 22(1), 81–87. <https://doi.org/10.1007/s10071-018-1223-x>
- Zentall, T. R., Laude, J. R., Stagner, J. P., & Smith, A. P. (2015). Suboptimal Choice by Pigeons: Evidence that the Value of the Conditioned Reinforcer Rather than its Frequency Determines Choice. *The Psychological Record*, 65(2), 223–229. <https://doi.org/10.1007/s40732-015-0119-2>
- Zentall, T. R., Smith, A. P., & Beckmann, J. (2019). Differences in rats and pigeons suboptimal choice may depend on where those stimuli are in their behavior system. *Behavioural Processes*, 159, 37–41. <https://doi.org/10.1016/j.beproc.2018.11.012>
- 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. Biological Sciences*, 278(1709), 1203–1208. <https://doi.org/10.1098/rspb.2010.1607>

Apéndice

Tabla 1.

Datos individuales durante el entrenamiento y reversión de la condición de bajo valor incentivo

Sujeto	Proporción de elección	Entrenamiento						Reversión						
		A.D.+	A.D. -	A.n.D.1	A.n.D.2	Idx Disc	Idx nDisc	Proporción de elección	A.D.+	A.D. -	A.n.D.1	A.n.D.2	Idx Disc	Idx nDisc
PAL1	0 ± 0	3.37 ± .57	0.51 ± .13	5.7 ± .93	3.92 ± .27	0.87	0.59	0.06 ± .02	3.05 ± .57	0.23 ± .04	6.55 ± .97	6.82 ± .25	0.93	0.49
PAL2	0.12 ± .03	12.4 ± 1.91	1.83 ± .39	11.55 ± .76	7.59 ± .39	0.87	0.60	0.53 ± .02	8.8 ± 1.04	0.83 ± .13	5.75 ± .29	7.21 ± .39	0.91	0.44
PAL5	0.07 ± .03	8.2 ± .6	0.9 ± .15	7.7 ± .64	6.11 ± .3	0.90	0.56	0.5 ± .04	6.95 ± .33	0.48 ± .07	7.8 ± .91	5.85 ± .88	0.94	0.57
PAL10	0.94 ± .04	3.55 ± .27	0.13 ± .05	4.5 ± .26	3.69 ± .34	0.97	0.55	0.97 ± .02	4.15 ± .37	0.26 ± .07	7.1 ± .48	7.61 ± .11	0.94	0.48
PAL12	0.07 ± .03	8.4 ± .61	0.55 ± .19	11.1 ± 2.02	10.32 ± .51	0.94	0.52	0.01 ± .01	7.7 ± .85	0.92 ± .24	6.15 ± 1.01	7.37 ± .83	0.89	0.46
PAL15	0.86 ± .02	14.8 ± .18	0.9 ± .05	12.75 ± 1.31	7.8 ± .59	0.94	0.62	0.85 ± .02	16.1 ± .98	0.61 ± .02	12.25 ± .51	8.45 ± .4	0.96	0.59
PAL16	0.07 ± .03	4.63 ± .53	0.63 ± .18	6.17 ± .79	5.59 ± .29	0.88	0.52	0.03 ± .01	5.75 ± .61	0.03 ± .02	3.55 ± .18	5.19 ± .45	1.00	0.41
SSJ11	0.08 ± .01	5.25 ± .63	0.24 ± .07	6.5 ± .72	5.98 ± .59	0.96	0.52	0.02 ± .01	5.8 ± .52	1.03 ± .2	5.8 ± .67	5.79 ± .36	0.85	0.50
SSJ12	0.01 ± .01	14.95 ± 1.16	0.93 ± .13	2.45 ± .44	5.43 ± .46	0.94	0.31	0.07 ± .02	4.8 ± .37	0.59 ± .14	3.35 ± .43	5.76 ± .33	0.89	0.37
SSJ13	0.33 ± .04	4.65 ± .28	1.06 ± .03	4.55 ± .44	3.74 ± .19	0.81	0.55	0.18 ± .03	7.1 ± 1.11	0.59 ± .08	8.45 ± .63	4.76 ± .38	0.92	0.64
SSJ15	0.03 ± .03	4.05 ± .5	1.1 ± .13	4.15 ± .43	4.63 ± .47	0.79	0.47	0.03 ± .02	6.45 ± .41	0.31 ± .07	3 ± .51	4.55 ± .39	0.95	0.40

Los datos corresponden al promedio de las últimas cinco sesiones de cada sujeto durante el entrenamiento (izquierda) y la reversión (derecha) de la condición de bajo valor incentivo, cada valor es acompañado de su respectivo SEM (error estándar de la media, por sus siglas en inglés). A.D.+ señala el promedio de respuestas al S+ de la alternativa discriminativa; A.D.- el promedio de respuestas al S- de la alternativa discriminativa; A.n.D.1 señala el promedio de respuestas al S1 de la alternativa no discriminativa y A.n.D.2 señala el promedio de respuestas al S2 de la alternativa no discriminativa. Idx Disc e Idx nDisc representan el índice de discriminación para la alternativa discriminativa y no discriminativa, respectivamente.

Tabla 2.

Datos individuales durante el entrenamiento y reversión de la condición de alto valor incentivo

Sujeto	Entrenamiento							Reversión						
	Proporción de elección	A.D.+	A.D. -	A.n.D.1	A.n.D.2	Idx Disc	Idx nDisc	Proporción de elección	A.D.+	A.D. -	A.n.D.1	A.n.D.2	Idx Disc	Idx nDisc
PAL1	0.99 ± .01	24.65 ± 1.03	1.09 ± .11	18.05 ± .84	15.79 ± 1.29	0.96	0.53	0.92 ± .04	21.95 ± 1.35	0.94 ± .12	27.6 ± 1.27	23.88 ± .59	0.96	0.54
PAL2	1 ± 0	9.8 ± .59	0.3 ± .24	10.65 ± 1.22	10.06 ± .74	0.97	0.51	0.98 ± .02	15.5 ± .94	0.25 ± .22	11.7 ± .58	12.09 ± .32	0.98	0.49
PAL5	0.97 ± .02	5.95 ± .46	0.33 ± .16	19.45 ± .9	16.4 ± .81	0.95	0.54	0.97 ± .01	10 ± .62	2.13 ± .38	8.35 ± .6	8.26 ± .7	0.83	0.50
PAL10	0.93 ± .03	26.05 ± 2.49	1.2 ± .7	18.7 ± 1.44	19.28 ± .49	0.96	0.49	0.13 ± .04	17 ± 1.37	1.05 ± .25	24.7 ± .4	20.94 ± .42	0.94	0.54
PAL12	1 ± 0	19.5 ± 1.39	0.86 ± .29	10.85 ± 1.24	9.76 ± .76	0.96	0.53	0.94 ± .02	3 ± .75	0.48 ± .17	4.05 ± .44	4.85 ± .55	0.86	0.46
PAL15	0.94 ± .02	12.5 ± .66	1.13 ± .23	15.55 ± .52	17.26 ± 1.31	0.92	0.47	0.91 ± .03	23 ± .61	2.78 ± .75	18.05 ± .9	15.89 ± 1.09	0.89	0.53
PAL16	1 ± 0	20.6 ± 1.17	0.06 ± 0	10.8 ± 1.81	5.94 ± .59	1.00	0.65	0.98 ± .01	19.05 ± 1.94	2.69 ± .74	12.05 ± 1.03	16.44 ± 1.01	0.88	0.42
SSJ11	0.89 ± .02	20.85 ± 1.79	2.44 ± 1.16	10.15 ± 1.27	11.29 ± .54	0.90	0.47	1 ± 0	9.95 ± .95	0.53 ± .07	8.85 ± .94	4.99 ± .94	0.95	0.64
SSJ12	0.24 ± .04	16.55 ± 1.82	0.2 ± .15	20.05 ± 2.28	16.44 ± 1.57	0.99	0.55	0 ± 0	17.45 ± 1.94	2.55 ± .77	9.95 ± 2.89	18 ± 1.23	0.87	0.36
SSJ13	0.87 ± .04	29 ± 1.76	5.33 ± .99	19.2 ± 1.69	18.54 ± 1.92	0.85	0.51	0.97 ± .02	22.05 ± .87	3.44 ± 1.54	27.35 ± 1.75	36.43 ± 1.58	0.87	0.43
SSJ15	0.77 ± .02	16.45 ± 1.35	2.18 ± .78	25.55 ± 1.63	31.25 ± .82	0.88	0.45	0.55 ± .03	18.05 ± 1.38	0.83 ± .11	23.9 ± 1.15	25.18 ± .37	0.96	0.49

Los datos corresponden al promedio de las últimas cinco sesiones de cada sujeto durante el entrenamiento (izquierda) y la reversión (derecha) de la condición de alto valor incentivo, cada valor es acompañado de su respectivo SEM (error estándar de la media, por sus siglas en inglés). A.D.+ señala el promedio de respuestas al S+ de la alternativa discriminativa; A.D.- el promedio de respuestas al S- de la alternativa discriminativa; A.n.D.1 señala el promedio de respuestas al S1 de la alternativa no discriminativa y A.n.D.2 señala el promedio de respuestas al S2 de la alternativa no discriminativa. Idx Disc e Idx nDisc representan el índice de discriminación para la alternativa discriminativa y no discriminativa, respectivamente.